

open

EUPHORIA

programming



Euphoria v4.0 svn3117

Table of Contents

1 Euphoria Programming Language v4.0.....	1
1.1 Introduction.....	1
1.1.1 Yet Another Programming Language?.....	1
1.1.2 First, EUPHORIA delivers the "expected" features of a modern language:.....	2
1.1.3 EUPHORIA is unique.....	2
1.1.4 EUPHORIA has qualities that go beyond the elegance of sequences.....	2
1.1.5 As a first programming language:.....	3
1.1.6 As an extension to languages you already know:.....	3
1.1.7 But, my favorite language is	3
1.1.8 Products.....	3
1.1.9 Requirements.....	3
1.1.10 Disclaimer.....	4
1.1.11 Discover EUPHORIA.....	4
1.1.12 Conventions used in the manual.....	4
1.2 An important message for all C/C++ programmers.....	4
1.2.1 24 Reasons Why You Are Going to Write Your Next Program in EUPHORIA!.....	4
1.3 Installing EUPHORIA.....	6
1.3.1 Possible Problems.....	6
1.3.2 Manually editing Environment Variables.....	6
1.3.3 How to Uninstall EUPHORIA.....	7
1.4 What to Do?.....	7
1.4.1 Run the Demo Programs.....	7
1.4.2 Edit Sample Files.....	7
1.4.3 Benchmark.....	7
1.4.4 Read the Manual.....	8
1.4.5 Trace a Demo.....	8
1.4.6 Run the Tutorial Programs.....	8
1.4.7 Modify the Tutorial Programs.....	8
1.4.8 Write Your Own.....	9
2 EUPHORIA License.....	10
2.1 EUPHORIA License.....	10
2.2 Licensing.....	10
2.3 EUPHORIA Credits.....	11
2.3.1 Current Authors.....	11
2.3.2 Past Authors.....	11
2.3.3 Contributors.....	11
3 Quick Start.....	13
3.1 Introduction.....	13
3.1.1 Example Programs.....	14
3.1.2 Installation.....	16
3.1.3 EUPHORIA programs.....	17
3.1.4 Running a Program.....	17
3.1.5 Running under Windows.....	19
3.1.6 Editing a Program.....	19

Table of Contents

3 Quick Start

3.1.7 Distributing a Program.....	20
3.2 Using EUPHORIA.....	21
3.2.1 Command line switches.....	21
3.2.2 Further notes.....	22
3.2.3 eu.cfg.....	23

4 Language Reference.....25

4.1 Definition.....	25
4.1.1 Objects.....	25
4.1.2 Identifiers.....	31
4.1.3 Comments.....	32
4.1.4 Expressions.....	32
4.1.5 Precedence Chart.....	41
4.2 Declarations.....	42
4.2.1 Identifiers.....	42
4.2.2 Specifying the type of a variable.....	47
4.2.3 Scope.....	51
4.3 Assignment statement.....	59
4.3.1 Assignment with Operator.....	60
4.4 Branching Statements.....	61
4.4.1 if statement.....	61
4.4.2 switch statement.....	62
4.4.3 ifdef statement.....	65
4.5 Loop statements.....	69
4.5.1 while statement.....	69
4.5.2 loop until statement.....	70
4.5.3 for statement.....	71
4.6 Flow control statements.....	72
4.6.1 exit statement.....	73
4.6.2 break statement.....	75
4.6.3 continue statement.....	75
4.6.4 retry statement.....	76
4.6.5 with entry statement.....	76
4.6.6 goto statement.....	77
4.6.7 Header Labels.....	78
4.7 Short-Circuit Evaluation.....	79
4.8 Special Top-Level Statements.....	80
4.8.1 include.....	81
4.8.2 with / without.....	83

5 Formal Syntax.....88

5.1 basics.....	88
5.2 statements.....	89
5.2.1 flow control.....	89
5.3 slice.....	89

Table of Contents

5 Formal Syntax

5.4 if.....	90
5.5 switch.....	90
5.6 break.....	90
5.7 continue.....	90
5.8 retry.....	90
5.9 exit.....	90
5.10 fallthru.....	90
5.11 for.....	91
5.12 while.....	91
5.13 loop.....	91
5.14 goto.....	91
5.15 EUPHORIA Internals.....	91
5.15.1 The EUPHORIA Data Structures.....	91
5.15.2 MAKE_INT.....	93
5.15.3 MAKE_SEQ.....	93
5.15.4 IS_ATOM_INT.....	93
5.15.5 DBL_PTR.....	93
5.15.6 MAKE_SEQ.....	94
5.15.7 IS_ATOM_DBL.....	94
5.15.8 IS_ATOM.....	95
5.15.9 IS_SEQUENCE.....	95
5.15.10 IS_DBL_OR_SEQUENCE.....	95
5.16 The C Representations of a EUPHORIA Sequence and a EUPHORIA Double.....	96

6 Mini-Guides.....97

6.1 Debugging and Profiling.....	98
6.1.1 Debugging.....	98
6.1.2 The Trace Screen.....	100
6.1.3 The Trace File.....	101
6.1.4 Profiling.....	102
6.1.5 Some Further Notes on Time Profiling.....	103
6.2 Binding and Shrouding.....	104
6.2.1 The Shroud Command.....	104
6.2.2 The Bind Command.....	105
6.3 EUPHORIA To C Translator.....	106
6.3.1 Introduction.....	106
6.3.2 C Compilers Supported.....	106
6.3.3 How to Run the Translator.....	107
6.3.4 Command-Line Options.....	108
6.3.5 Dynamic Link Libraries (Shared Libraries).....	109
6.3.6 Executable Size and Compression.....	110
6.3.7 Interpreter vs. Translator.....	111
6.3.8 Legal Restrictions.....	111
6.3.9 Disclaimer.....	112
6.3.10 Frequently Asked Questions.....	112

Table of Contents

6 Mini-Guides

6.3.11 Common Problems.....	113
6.4 Indirect routine calling.....	113
6.4.1 Indirect calling a routine coded in EUPHORIA.....	114
6.4.2 Calling EUPHORIA's internals.....	116
6.5 Multitasking in EUPHORIA.....	116
6.5.1 Introduction.....	116
6.5.2 Why Multitask?.....	116
6.5.3 Types of Tasks.....	117
6.5.4 A Small Example.....	117
6.5.5 Comparison with earlier multitasking schemes.....	118
6.5.6 Comparison with multithreading.....	119
6.5.7 Summary.....	120
6.6 EUPHORIA Database System (EDS).....	120
6.6.1 Introduction.....	120
6.6.2 Structure of an EDS database.....	120
6.6.3 How to access the data.....	121
6.6.4 How does storage get recycled?.....	121
6.6.5 Security / Multi-user Access.....	122
6.6.6 Scalability.....	122
6.6.7 Disclaimer.....	122
6.6.8 Warning: Use the right file mode.....	122
6.7 The User Defined Pre-Processor.....	122
6.7.1 A Quick Example.....	123
6.7.2 Pre-process Details.....	124
6.7.3 Command Line Options.....	125
6.7.4 DLL/Shared Library Interface.....	126
6.7.5 Advanced Examples.....	127
6.8 EUPHORIA Trouble-Shooting Guide.....	129
6.8.1 Common Problems and Solutions.....	129
6.9 Platform Specific Issues.....	134
6.9.1 Introduction.....	134
6.9.2 The WIN32 Platform.....	135
6.9.3 The Unix Platforms.....	136
6.9.4 Interfacing with C Code (WIN32, Linux, FreeBSD).....	136
6.10 Performance Tips.....	140
6.10.1 General Tips.....	141
6.10.2 Measuring Performance.....	142
6.10.3 How to Speed-Up Loops.....	143
6.10.4 Converting Multiplies to Adds in a Loop.....	143
6.10.5 Saving Results in Variables.....	143
6.10.6 In-lining of Routine Calls.....	144
6.10.7 Operations on Sequences.....	144
6.10.8 Some Special Case Optimizations.....	144
6.10.9 Assignment with Operators.....	145
6.10.10 Pixel-Graphics Tips.....	145

Table of Contents

6 Mini-Guides

6.10.11 Library Routines.....	145
6.10.12 Searching.....	146
6.10.13 Sorting.....	146
6.10.14 Taking Advantage of Cache Memory.....	147
6.10.15 Using Machine Code and C.....	147
6.10.16 Using The EUPHORIA To C Translator.....	147

7 Included Tools.....148

7.1 EuTEST - Unit Testing.....	148
7.1.1 Introduction.....	149
7.1.2 The eutest Program.....	149
7.1.3 The Unit Test Files.....	150
7.1.4 The Error Control Files.....	150
7.2 EuDOC - Source Documentation Tool.....	151
7.2.1 Documentation tags.....	151
7.2.2 Generic documentation.....	151
7.2.3 Source documentation.....	152
7.2.4 Assembly file.....	153
7.2.5 Creole markup.....	153
7.2.6 Documentation software.....	154
7.3 Ed - EUPHORIA Editor.....	154
7.3.1 Introduction.....	155
7.3.2 Summary.....	155
7.3.3 Special Keys.....	156
7.3.4 Escape Commands.....	156
7.3.5 Recalling Previous Strings.....	157
7.3.6 Cutting and Pasting.....	157
7.3.7 Use of Tabs.....	158
7.3.8 Long Lines.....	158
7.3.9 Maximum File Size.....	158
7.3.10 Non-text Files.....	158
7.3.11 Line Terminator.....	158
7.3.12 Source Code.....	159
7.3.13 Platform Issues.....	159

8 General Routine Reference.....160

8.1 Command Line Handling.....	161
8.1.1 Constants.....	162
8.1.2 Routines.....	166
8.2 Console.....	175
8.2.1 Cursor Style Constants.....	175
8.2.2 Keyboard related routines.....	176
8.2.3 Cross Platform Text Graphics.....	182
8.3 Date/Time.....	192
8.3.1 Localized Variables.....	192

Table of Contents

8 General Routine Reference

8.3.2 Constants.....	193
8.3.3 Types.....	194
8.3.4 Routines.....	195
8.4 File System.....	208
8.4.1 Constants.....	210
8.4.2 Directory Handling.....	211
8.4.3 File name parsing.....	222
8.4.4 File Types.....	230
8.4.5 File Handling.....	231
8.5 I/O.....	240
8.5.1 Constants.....	241
8.5.2 Read/Write Routines.....	242
8.5.3 Low Level File/Device Handling.....	254
8.5.4 File Reading/Writing.....	263
8.6 Math.....	274
8.6.1 Sign and comparisons.....	275
8.6.2 Roundings and remainders.....	279
8.6.3 Trigonometry.....	286
8.6.4 Logarithms and powers.....	293
8.6.5 Hyperbolic trigonometry.....	298
8.6.6 Accumulation.....	303
8.6.7 Bitwise operations.....	305
8.7 Math Constants.....	315
8.7.1 Constants.....	316
8.8 Random Numbers.....	319
8.8.1 rand ;rand;.....	320
8.9 Mouse.....	327
8.9.1 Requirements.....	327
8.9.2 Constants.....	327
8.9.3 Routines.....	329
8.10 Operating System Helpers.....	332
8.10.1 CMD_SWITCHES ;CMD_SWITCHES;.....	333
8.10.2 Operating System Constants.....	333
8.10.3 Environment.....	334
8.10.4 Interacting with the OS.....	339
8.10.5 Miscellaneous.....	341
8.10.6 Pipe Input/Output.....	343
8.10.7 Accessor Constants.....	343
8.10.8 Opening/Closing.....	344
8.10.9 Read/Write Process.....	345
8.11 Pretty Printing.....	347
8.11.1 PRETTY_DEFAULT ;PRETTY_DEFAULT;.....	348
8.11.2 Routines.....	349
8.12 Statistics.....	352
8.12.1 Routines.....	353

Table of Contents

8 General Routine Reference

8.13 Multi-tasking.....	374
8.13.1 General Notes.....	374
8.13.2 Warning.....	374
8.13.3 Routines.....	374
8.14 Types - Extended.....	383
8.14.1 OBJ_UNASSIGNED ;OBJ_UNASSIGNED;.....	384
8.14.2 Support Functions.....	388
8.14.3 Types.....	390

9 Sequence Centric Routines.....405

9.1 Data type conversion.....	405
9.1.1 Routines.....	406
9.2 Input Routines.....	416
9.2.1 Error Status Constants.....	416
9.2.2 Answer Types.....	417
9.2.3 Routines.....	417
9.3 Searching.....	422
9.3.1 Equality.....	423
9.3.2 Finding.....	425
9.3.3 Matching.....	434
9.4 Sequence Manipulation.....	441
9.4.1 Constants.....	443
9.4.2 Basic routines.....	443
9.4.3 Building sequences.....	453
9.4.4 Adding to sequences.....	455
9.4.5 Extracting, removing, replacing from/into a sequence.....	463
9.4.6 Changing the shape of a sequence.....	477
9.5 Serialization of EUPHORIA Objects.....	490
9.5.1 Routines.....	490
9.6 Sorting.....	495
9.6.1 Constants.....	495
9.6.2 Routines.....	496

10 String Centric Routines.....503

10.1 Locale Routines.....	503
10.1.1 Message translation functions.....	504
10.1.2 Time/Number Translation.....	509
10.1.3 Constants.....	511
10.1.4 Locale Name Translation.....	515
10.2 Regular Expressions.....	517
10.2.1 Introduction.....	518
10.2.2 General Use.....	519
10.2.3 Option Constants.....	519
10.2.4 Error Constants.....	523
10.2.5 Create/Destroy.....	526

Table of Contents

10 String Centric Routines

10.2.6 Utility Routines.....	528
10.2.7 Find/Match.....	529
10.2.8 Splitting.....	534
10.2.9 Replacement.....	535
10.3 Text Manipulation.....	537
10.3.1 Routines.....	538
10.4 Unicode.....	554
10.4.1 isUChar ;isUChar;.....	555
10.5 Wildcard Matching.....	576
10.5.1 Routines.....	577

11 Data Structures.....580

11.1 EUPHORIA Database (EDS).....	580
11.1.1 Database File Format.....	582
11.1.2 Error Status Constants.....	583
11.1.3 Lock Type Constants.....	584
11.1.4 Error Code Constants.....	585
11.1.5 Variables.....	586
11.1.6 Routines.....	587
11.1.7 Managing databases.....	589
11.2 Prime Numbers.....	609
11.2.1 Routines.....	610
11.3 Map (hash table).....	612
11.3.1 hash ;hash;.....	614
11.3.2 Hashing Algorithms.....	614
11.3.3 Operation codes for put.....	615
11.3.4 Types of Maps.....	616
11.3.5 Types.....	617
11.3.6 Routines.....	617
11.4 Stack.....	639
11.4.1 Constants.....	640
11.4.2 Types.....	640
11.4.3 Routines.....	641
11.5 Sets.....	654
11.5.1 Notes:.....	655
11.5.2 Types.....	655
11.5.3 Inclusion and belonging.....	657
11.5.4 Basic set-theoretic operations.....	663
11.5.5 Maps between sets.....	666
11.5.6 Reverse mappings.....	676
11.5.7 Products.....	678
11.5.8 Constants.....	680
11.5.9 Operations on sets.....	681

Table of Contents

12 Networking Routines.....	688
12.1 Core Sockets.....	688
12.1.1 Error Information.....	691
12.1.2 Socket Type Constants.....	699
12.1.3 Select Accessor Constants.....	701
12.1.4 Shutdown Options.....	702
12.1.5 Socket Options.....	702
12.1.6 Send Flags.....	707
12.1.7 Server and Client sides.....	711
12.1.8 Client side only.....	716
12.1.9 Server side only.....	717
12.1.10 UDP only.....	719
12.1.11 Information.....	720
12.2 Common Internet Routines.....	721
12.2.1 IP Address Handling.....	721
12.2.2 URL Parsing.....	723
12.3 DNS.....	725
12.3.1 Constants.....	726
12.3.2 General Routines.....	731
12.4 HTTP.....	733
12.4.1 Constants.....	734
12.4.2 Header management.....	737
12.4.3 Web interface.....	739
12.5 URL handling.....	741
12.5.1 Parsing.....	742
12.5.2 URL encoding and decoding.....	744
 13 Low Level Routines.....	 746
13.1 Dynamic Linking to external code.....	746
13.1.1 C Type Constants.....	747
13.1.2 External EUPHORIA Type Constants.....	751
13.1.3 Constants.....	752
13.1.4 Routines.....	752
13.2 Errors and Warnings.....	762
13.2.1 Routines.....	762
13.3 Pseudo Memory.....	768
13.3.1 ram_space ;ram_space;.....	768
13.4 Indirect Routine Calling.....	772
13.4.1 SAFE mode.....	772
13.4.2 Data Execute mode.....	772
13.4.3 Accessing EUPHORIA coded routines.....	773
13.4.4 Accessing EUPHORIA internals.....	776
13.5 Types supporting Memory.....	778
13.5.1 valid_memory_protection_constant ;valid_memory_protection_constant;.....	778
13.5.2 Allocating and Writing to memory:.....	779
13.5.3 Memory disposal.....	784

Table of Contents

13 Low Level Routines

13.6 Memory Management - Low-Level.....	786
13.6.1 Usage Notes.....	789
13.6.2 Memory allocation.....	790
13.6.3 Reading from, Writing to, and Calling into Memory.....	792
13.6.4 Safe memory access.....	805
13.6.5 safe.e.....	809
13.6.6 Microsoft Windows Memory Protection Constants.....	815
13.6.7 Standard Library Memory Protection Constants.....	817

14 Graphics.....825

14.1 Error Code Constants.....	825
14.1.1 BMP_SUCCESS ;BMP_SUCCESS;.....	825
14.1.2 video_config sequence accessors.....	826
14.1.3 Routines.....	830
14.2 Graphics - Cross Platform.....	831
14.2.1 Routines.....	832
14.2.2 Graphics Modes.....	836
14.3 Graphics - Image Routines.....	836
14.3.1 graphics_point ;graphics_point;.....	836
14.3.2 Bitmap handling.....	837

15 EUPHORIA Routines.....840

15.1 EUPHORIA Information.....	840
15.1.1 Numeric Version Information.....	840
15.1.2 Compiled Platform Information.....	840
15.1.3 String Version Information.....	842
15.1.4 Copyright Information.....	844
15.2 Keyword Data.....	845
15.2.1 keywords ;keywords;.....	845
15.3 Syntax Coloring.....	846
15.3.1 set_colors ;set_colors;.....	846
15.4 Source Tokenizer.....	846
15.4.1 T_EOF ;T_EOF;.....	846
15.5 Unit Testing Framework.....	849
15.5.1 Background.....	850
15.5.2 Constants.....	851
15.5.3 Setup Routines.....	851
15.5.4 Reporting.....	854
15.5.5 Tests.....	854

16 Windows Routines.....859

16.1 Windows Message Box.....	859
16.1.1 Style Constants.....	860
16.1.2 Return Value Constants.....	864
16.1.3 Routines.....	866

Table of Contents

16 Windows Routines	
16.2 Windows Sound.....	866
16.2.1 SND_DEFAULT ;SND_DEFAULT;.....	867
17 Release Notes.....	869
17.1 Version 4.0.....	869
17.1.1 Bug Fixes.....	869
17.1.2 Changes.....	869
17.1.3 New Features.....	870
17.1.4 New Routines/Constants.....	871
18 Index.....	873

1 Euphoria Programming Language v4.0

Introduction

Yet Another Programming Language?

First, EUPHORIA delivers the "expected" features of a modern language:

EUPHORIA is unique

EUPHORIA has qualities that go beyond the elegance of sequences

As a first programming language:

As an extension to languages you already know:

But, my favorite language is ...

Products

Requirements

Disclaimer

Discover EUPHORIA

Conventions used in the manual

An important message for all C/C++ programmers...

24 Reasons Why You Are Going to Write Your Next Program in EUPHORIA!

Installing EUPHORIA

Possible Problems

Manually editing Environment Variables

How to Uninstall EUPHORIA

What to Do?

Run the Demo Programs

Edit Sample Files

Benchmark

Read the Manual

Trace a Demo

Run the Tutorial Programs

Modify the Tutorial Programs

Write Your Own

1.1 Introduction

Welcome to EUPHORIA! *End User Programming with Hierarchical Objects for Robust Interpreted Applications.*

EUPHORIA has come a long way since v1.0 was released in July 1993 by **Rapid Deployment Software (RDS)**. There are now enthusiastic users around the world.

1.1.1 Yet Another Programming Language?

EUPHORIA is a very high-level programming language. It is unique among a crowd of conventional languages.

1.1.2 First, EUPHORIA delivers the "expected" features of a modern language:

- open source
- free for personal and commercial use
- produces royalty-free, stand-alone, programs
- multi-platform--Windows, Linux, FreeBSD, and OSX
- provides a choice of multi-platform gui toolkits: IUP, GTK, wxWindows
- features a Windows IDE and gui library written in EUPHORIA
- colored-syntax editors, profiling, and tracing of code
- dynamic memory allocation and efficient garbage collection
- interfacing to existing C libraries and databases
- well-documented, lots of example source-code, and an enthusiastic forum
- edit and run convenience

1.1.3 EUPHORIA is unique

What makes EUPHORIA unique is a design that uses just two basic data-types -- *atom* and *sequence*, and two 'helper' data-types--*object* and *integer*.

- An **atom** is single numeric value (either an integer or floating point)
- A **sequence** is a list of zero or more *objects*.
- An **object** is a *variant* type in that it can hold an atom or a sequence.
- An *integer* is just a special form of atom that can only hold integers. You can use the *integer* type for a performance advantage in situations where floating point values are not required.

What follows from this design are some advantages over conventional languages:

- the language syntax is smaller--and thus easier to learn
- the language syntax is consistent--and thus easier to program
- routines are more generic--a routine used for strings may also be applied to any data structure
- a higher level view of programming--because sequences encompass conventional lists, arrays, tables, tuples, ..., and all other data-structures.
- sequences are dynamic--you may create and destroy at will--and modify them to any size and complexity
- it supports both *static* data typing and *dynamic* data typing.

1.1.4 EUPHORIA has qualities that go beyond the elegance of sequences

- EUPHORIA programs are considerably faster than conventional interpreted languages--EUPHORIA makes a better website server
- EUPHORIA programs can be translated then compiled as C programs--fast programs become even faster
- EUPHORIA lets you write multi-tasking programs--independent of the platform you are using
- EUPHORIA has a coherent design--EUPHORIA programmers enjoy programming in EUPHORIA

1.1.5 As a first programming language:

- Easy to learn, easy to program
- No limits as to what you can program
- EUPHORIA programming skills will enhance learning other languages

1.1.6 As an extension to languages you already know:

- A fast, flexible, and powerful language
- EUPHORIA, the language you will prefer to program in

1.1.7 But, my favorite language is ...

You will find that EUPHORIA programmers are also knowledgeable in other languages. I find that the more tools you have (saws and hammers, or programming languages) the richer you are. Picking the correct tool is part of the art of programming. It will remain true that some people can program better in their favorite language rather than an arguably superior language.

Give EUPHORIA a try, and discover why it has enthusiastic supporters.

1.1.8 Products

This **EUPHORIA Interpreter, Binder, and Translator** package is free for anyone to use.

Using the *EUPHORIA Binder* is used to create stand-alone programs.

The *EUPHORIA Translator* converts EUPHORIA-source into C-source. This allows EUPHORIA programs to be compiled by a standard C compiler to make even faster stand-alone programs.

You can freely distribute the EUPHORIA interpreter, and any other files contained in this package, in whole or in part, so anyone can run a EUPHORIA program that you have developed. You are completely free to distribute any EUPHORIA programs that you write.

1.1.9 Requirements

To run the *WIN32* version of EUPHORIA, you need Windows 95 or any later version of Windows. It runs fine on XP and Vista.

To run the *Linux* version of EUPHORIA you need any reasonably up-to-date Linux distribution, that has libc6 or later. For example, Red Hat 5.2 or later will work fine.

To run the *FreeBSD* version of EUPHORIA you need any reasonably up-to-date FreeBSD distribution.

To run the *Mac OS X* version of EUPHORIA, you need any reasonably up-to-date Intel based Mac.

1.1.10 Disclaimer

EUPHORIA is provided "as is" without warranty of any kind. In no event shall Rapid Deployment Software be held liable for any damages arising from the use of, or inability to use, this product.

1.1.11 Discover EUPHORIA

More information on EUPHORIA may be found at OpenEUPHORIA.org, along with an active [discussion forum](#).

Download EUPHORIA from the [EUPHORIA Web site](#). You will also find source-code for example programs and library files.

1.1.12 Conventions used in the manual

EUPHORIA has multiple interpreters, the main one being **eui**.

- On windows platforms you have two choices. If you run **eui** then a console window is created. If you run **euiw** then no console is created--making it suitable for GUI applications.

The manual will only reference **eui** in examples and instructions; the reader is left to choose the correct interpreter.

EUPHORIA runs on many platforms. When operating specific issues must be described you will see these descriptions:

- "Windows" will be a general reference to a family of operating systems that includes WIN/95/ME/98/XP/NT/Vista/7/... You will see the constant **WIN32** used for Windows specific code.
- "Unix" will be a general reference to the family operating systems that includes Linux/FreeBSD,Mac OS X/... You will see the constant **UNIX** used for Unix specific code. You may also see **"*nix"** used to describe this group of operating systems.

Directory names in Windows use \ separators, while Unix systems use /. Unix users should substitute / when they examine example code.

Operating system names are often trademarks. There is no intent to infringe on their owner rights.

1.2 An important message for all C/C++ programmers...

1.2.1 24 Reasons Why You Are Going to Write Your Next Program in EUPHORIA!

1. because you are tired of having to re-invent dynamic storage allocation for each program that you write
2. because you have spent too many frustrating hours tracking down malloc arena corruption bugs
3. because you were once plagued for several days by an on-again/off-again "flaky" bug that eventually was traced to an uninitialized variable
4. because no matter how hard you try to eliminate them, there is always one more storage "leak"
5. because you are tired of having the machine "lock up", or your program come crashing down in flames with no indication of what the error was
6. because you know that **subscript checking** would have saved you from hours of debugging
7. because your program should not be allowed to overwrite random areas in memory via "wild" pointers
8. because you know it would be bad to overflow your fixed-size stack area but you have no idea of how close you are
9. because one time you had this weird bug, where you called a function, that didn't actually return a value, but instead fell off the end and some random garbage was "returned"
10. because you wish that library routines would stop you from passing in bad arguments, rather than just setting "errno" or whatever (who looks at errno after every call?)
11. because you would like to "recompile the world" in a fraction of a second rather than several minutes -- you can work much faster with a cycle of edit/run rather than edit/compile/link/run.
12. because *The C++ Programming Language* 3rd Ed. by Bjarne Stroustrup is 911 very dense pages, (and doesn't even discuss platform-specific programming for Windows, Linux or any other system).
13. because you have been programming in C/C++ for a long time now, but there are still a lot of weird features in the language that you don't fully understand
14. because portability is not as easy to achieve as it should be
15. because you know the range of legitimate values for each of your variables, but you have no way of enforcing this at runtime
16. because you would like to pass variable numbers of arguments, but you are put off by the complicated way of doing it in C
17. because you would like a *clean way* of returning multiple values from a function
18. because you want an integrated **full-screen source-level debugger** that is so easy to use that you don't have to search through the manual each time, (or give up and recompile with printf statements)
19. because you hate it when your program starts working just because you added a debug print statement or compiled with the debug option
20. because you would like a reliable, accurate **statement-level profile** to understand the internal dynamics of your program, and to boost performance
21. because very few of your programs have to squeeze every cycle of performance out of your machine. The speed difference between EUPHORIA and C/C++ is not that great, especially when you use the EUPHORIA to C Translator. Try some benchmark tests. We bet you'll be surprised!
22. because you'd rather not clutter up your hard disk with .obj and .exe files
23. because you'd rather be running your program, than wading through several hundred pages of documentation to decide what compiler and linker options you need
24. because your C/C++ package has 57 different routines for memory allocation, and 67 different routines for manipulating strings and blocks of memory. How many of these routines does EUPHORIA need? **Answer: zero.** In EUPHORIA, memory allocation happens automatically and strings are manipulated just like any other sequences.

1.3 Installing EUPHORIA

Download and run the latest EUPHORIA setup program. It will install EUPHORIA on any *Windows* system from *Windows 95* up.

After installing, see [What to do next](#) for ideas on how to use this package.

1.3.1 Possible Problems

- If the install appeared to run ok, but example programs are not working, did you remember to shut down and restart your computer?
- On *Windows XP/2000*, be careful that your `PATH` and `EUDIR` do not conflict with `autoexec.nt`, which can also be used to set environment variables.
- On *WinME/98/95* if the install procedure fails to edit your `autoexec.bat` file, you will have to do it yourself. Follow the manual procedure described below.

1.3.2 Manually editing Environment Variables

On *Windows 95/98/ME* systems, you will need to edit the `autoexec.bat` file to update the environment variables before using EUPHORIA.

1. In the file `c:\autoexec.bat` add `C:\EUPHORIA\BIN` to the list of directories in your `PATH` command. You might use the Windows Notepad or any other text editor to do this. You can also go to the *Start Menu*, select *Run*, type in `sysedit` and press *Enter*. `autoexec.bat` should appear as one of the system files that you can edit and save.
2. In the same `autoexec.bat` file add a new line: `SET EUDIR=C:\EUPHORIA`
3. The `EUDIR` environment variable indicates the full path to the main EUPHORIA directory.
4. Reboot (restart) your machine. This will define your new `PATH` and `EUDIR` environment variables.

On *WinNT/2000/XP* and higher, if for some reason your `EUDIR` and `PATH` variables are not set correctly, then set them in whatever way your system allows. For example on *Windows XP* select: *Start Menu* -> *Control Panel* -> *Performance & Maintenance* -> *System* -> *Advanced* then click the *Environment Variables* button. Click the top *New...* button then enter `EUDIR` as the *Variable Name* and `c:\euphoria` (or whatever is correct) for the value, then click *OK*. Find `PATH` in the list of your variables, select it, then click *Edit...*. Add `;c:\euphoria\bin` at the end and click *OK*. The change will be effective next time you open a console window, or after a computer reboot.

Some systems, such as *Windows ME*, have an `AUTOEXEC.BAT` file, but it's a hidden file that might not show up in a directory listing. Nevertheless it's there, and you can view it and edit it if necessary by typing, for example: `notepad c:\autoexec.bat` in a console window.

If you have an `autoexec.bat` file, but it doesn't contain a `PATH` command, you will have to create one that includes `C:\EUPHORIA\BIN`.

There is another, optional, environment variable used by some experienced users of EUPHORIA. It is called `EUINC` (see a [Guru Search](#)). It determines the search path for included files. You might want to add it, or

change it, when you install a new version of EUPHORIA.

1.3.3 How to Uninstall EUPHORIA

1. If you wish to recover your previous version of EUPHORIA, or parts of it, the "backup" subdirectory contains a backup copy of each of your previous standard EUPHORIA subdirectories, plus any files that you may have added. However it does not contain any additional subdirectories that you may have created on your own.
2. If there are no files that you need, you can delete the EUPHORIA directory that you installed into.
3. Delete the EUDIR environment variable, and remove the EUPHORIA directory from your PATH, either in C:\AUTOEXEC.BAT, or in Control Panel/System/Advanced.
4. Delete the references to EUPHORIA in your Start Menu.

1.4 What to Do?

Now that you have installed EUPHORIA, here are some things you can try:

1.4.1 Run the Demo Programs

Run each of the demo programs in the demo directory. You just type `eui <program name>`. An example of running the demos in a console

```
eui buzz
```

You can also double-click on a `.ex` or `.exw` file from *Windows* as file associations have been setup during the installation process.

1.4.2 Edit Sample Files

Use the EUPHORIA editor, `ed`, to edit a EUPHORIA file. Notice the use of colors. You can adjust these colors along with the cursor size and many other "user-modifiable" parameters by editing constant declarations in `ed.ex`. Use *Esc q* to quit the editor or *Esc h* for help. There are several, even better, EUPHORIA-oriented editors in [The Archive](#). If you use a more sophisticated text editor, many have a highlighter file for EUPHORIA. You will find it either on the Archive or on the community page for that editor.

1.4.3 Benchmark

Create some new benchmark tests. See `demo\bench`. Do you get the same speed ratios as we did in comparison with other popular languages?

1.4.4 Read the Manual

Read the manual in `html\index.html` by double-clicking it. The simple expressive power of EUPHORIA makes this manual much shorter than manuals for other languages. If you have a specific question, type at the console:

```
guru word
```

The `guru` program will search all the `.doc` files, example programs, and other files, and will present you with a *sorted* list of the most relevant chunks of text that might answer your enquiry.

1.4.5 Trace a Demo

Try running a EUPHORIA program with `tracing` turned on. Add:

```
with trace
trace(1)
```

at the beginning of any EUPHORIA source file.

1.4.6 Run the Tutorial Programs

Run some of the tutorial programs in `euphoria\tutorial`.

1.4.7 Modify the Tutorial Programs

Try modifying some of the demo programs.

First some *simple* modifications (takes less than a minute):

1.4.7.1 Simple

What if there were 100 C++ ships in **Language Wars**? What if `sb.ex` had to move 1000 balls instead of 125? Change some parameters in `polygon.ex`. Can you get prettier pictures to appear? Add some funny phrases to `buzz.ex`.

1.4.7.2 Harder

Then, some slightly harder ones (takes a few minutes):

Define a new function of `x` and `y` in `plot3d.ex`.

1.4.7.3 Challenging

Then a challenging one (takes an hour or more):

Set up your own customized database by defining the fields in `mydata.ex`.

1.4.7.4 Major

Then a major project (several days or weeks):

Write a *smarter* 3D TicTacToe algorithm.

1.4.8 Write Your Own

Try writing your own program in EUPHORIA. A program can be as simple as:

```
? 2+2
```

Remember that after any error you can simply type: `ed` to jump into the editor at the offending file and line.

Once you get used to it, you'll be developing programs *much* faster in EUPHORIA than you could in Perl, Java, C/C++ or any other language that we are aware of.

2 EUPHORIA License

[EUPHORIA License](#)

[Licensing](#)

[EUPHORIA Credits](#)

[Current Authors](#)

[Past Authors](#)

[Contributors](#)

2.1 EUPHORIA License

Copyright (c) 2007–2010 by OpenEUPHORIA Group

Copyright (c) 1993–2006 Rapid Deployment Software (RDS)

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

RDS requests, but does not require, that you:

1. Acknowledge RDS and others who contribute to this software,
 2. Provide a link to www.RapidEUPHORIA.com, if possible, from your Web site.
-

2.2 Licensing

This product is free and open source, and has benefited from the contributions of many people. You have complete royalty-free rights to distribute any EUPHORIA programs that you develop. You are also free to distribute the interpreter, backend and even translator. You can shroud or bind your program and distribute the resulting files royalty-free.

You may incorporate any EUPHORIA source files from this package into your program, either "as is" or with your modifications. (You will probably need at least a few of the standard `euphoria\include` files in any large program).

We would appreciate it if you told people that your program was developed using EUPHORIA, and gave them the address: <http://www.rapideuphoria.com> of our Web page, but we do not require any such acknowledgment.

Icon files, such as `euphoria.ico` in `euphoria\bin`, may be distributed with or without your changes.

The high-speed version of the EUPHORIA Interpreter back-end is written in ANSI C, and can be compiled with many different C compilers. The complete source code is in `euphoria\source`, along with `execute.e`, the alternate, EUPHORIA-coded back-end. The generous [Open source License](#) allows both personal and commercial use, and unlike many other open source licenses, your changes do not have to be made open source.

Some additional 3rd-party legal restrictions might apply when you use the [EUPHORIA To C Translator](#).

2.3 EUPHORIA Credits

EUPHORIA has been continuously developed since it was started in 1993 by Robert Craig. In 2006 version 3.0 was released as open source. Various releases were made to the 3.x series and then in 2009 version 4.0 was released... EUPHORIA's biggest release ever.

It has taken quite a few people to get thus far and we would like to recognize them here. Authors/Contributors are listed in alphabetical order by their last name.

2.3.1 Current Authors

- Jim Brown
- Tom Ciplijauskas
- Jeremy Cowgar
- Matthew Lewis
- Derek Parnell
- Shawn Pringle

2.3.2 Past Authors

- Robert Craig
- Chris Cuvier
- Junko Miura

2.3.3 Contributors

- Jiri Babor
- Chris Bensler
- CoJaBo
- Jason Gade

- Ryan Johnson
- Christopher Lester
- Marco Antonio Achury Palma
- Michael Sabal
- Kathy Smith
- Yuku (Aku)

3 Quick Start

[Introduction](#)
[Example Programs](#)
[Installation](#)
[EUPHORIA programs](#)
[Running a Program](#)
[Running under Windows](#)
[Editing a Program](#)
[Distributing a Program](#)
[Using EUPHORIA](#)
[Command line switches](#)
[Further notes](#)
[eu.cfg](#)

3.1 Introduction

EUPHORIA is a programming language with the following advantages over conventional languages:

- Euphoric
 - ◆ A remarkably simple, flexible, powerful language definition that is easy to learn and use.
- Dynamic
 - ◆ Variables grow or shrink without the programmer having to worry about allocating and freeing chunks of memory. Objects of any size can be assigned to an element of a EUPHORIA sequence (array).
- Fast
 - ◆ a high-performance, state-of-the-art interpreter that's significantly faster than conventional interpreters such as Perl and Python.
- Compiles
 - ◆ an optimizing [EUPHORIA To C Translator](#), that can boost your speed even further, often by a factor of 2x to 5x versus the already-fast interpreter.
- Safe
 - ◆ extensive run-time checking for: out-of-bounds subscripts, uninitialized variables, bad parameter values for library routines, illegal value assigned to a variable and many more. There are no mysterious machine exceptions--you will always get a full English description of any problem that occurs with your program at run-time, along with a call-stack trace-back and a dump of all of your variable values. Programs can be debugged quickly, easily and more thoroughly.
- High level
 - ◆ features of the underlying hardware are completely hidden. Programs are not aware of word-lengths, underlying bit-level representation of values, byte-order etc.
- Debugger
 - ◆ a full-screen source debugger and an execution profiler are included, along with a full-screen, multi-file editor. On a color monitor, the editor displays EUPHORIA programs in multiple colors, to highlight comments, reserved words, built-in functions, strings, and level of nesting of brackets. It optionally performs auto-completion of statements, saving you typing effort

and reducing syntax errors. This editor is written in EUPHORIA, and the source code is provided to you without restrictions. You are free to modify it, add features, and redistribute it as you wish.

- Multi-platform
 - ◆ EUPHORIA programs run under Linux, FreeBSD, 32-bit *Windows*
- Stand-alone
 - ◆ You can make a single, stand-alone executable file from your program.
- Generic
 - ◆ EUPHORIA routines are naturally generic. The example program below shows a single routine that will sort any type of data--integers, floating-point numbers, strings etc. EUPHORIA is not an "object-oriented" language, yet it achieves many of the benefits of these languages in a much simpler way.
- Free
 - ◆ EUPHORIA is completely free and open source.

3.1.1 Example Programs

3.1.1.1 Hello, World

The *mandatory* 'hello world' program is a one-liner in EUPHORIA.

```
-----  
puts(1, "Hello, World\n")
```

This simply calls the built-in routine `puts` ('put string'), passing it two parameters:- 1, which is the file number to output to (STDOUT - the console, in this case) and a string of text.

3.1.1.2 Sorting

The following is an example of a more useful EUPHORIA program.

```
-----  
include std/console.e  
sequence original_list  
  
function merge_sort(sequence x)  
-- put x into ascending order using a recursive merge sort  
  integer n, mid  
  sequence merged, a, b  
  
  n = length(x)  
  if n = 0 or n = 1 then  
    return x -- trivial case  
  end if  
  
  mid = floor(n/2)  
  a = merge_sort(x[1..mid])      -- sort first half of x  
  b = merge_sort(x[mid+1..n])    -- sort second half of x  
  
  -- merge the two sorted halves into one
```

```

merged = {}
while length(a) > 0 and length(b) > 0 do
    if compare(a[1], b[1]) < 0 then
        merged = append(merged, a[1])
        a = a[2..length(a)]
    else
        merged = append(merged, b[1])
        b = b[2..length(b)]
    end if
end while
return merged & a & b -- merged data plus leftovers
end function

procedure print_sorted_list()
-- generate sorted_list from original_list
sequence sorted_list

original_list = {19, 10, 23, 41, 84, 55, 98, 67, 76, 32}
sorted_list = merge_sort(original_list)
for i = 1 to length(sorted_list) do
    display ("Number [] was at position [:2], now at [:2]",
        { sorted_list[i], find(sorted_list[i], original_list), i}
    )
end for
end procedure

print_sorted_list() -- this command starts the program

```

The above example contains a number of statements that are processed in order.

```
include std/console.e
```

This tells the EUPHORIA that this application needs access to the public symbols declared in the file 'std/console.e'. This is referred to as a *library* file. In our case here, the application will be using the `display()` routine from `console.e`.

```
sequence original_list
```

This declares a variable that is not public but is accessible from anywhere in this file. The datatype for the variable is a sequence, which is a variable-length array, and whose symbol name is `original_list`.

```
function merge_sort(sequence x) ... end function
```

This declares and defines a function routine. Functions return values when called. This function must be passed a single parameter when called - a sequence.

```
procedure print_sorted_list() ... end procedure
```

This declares and defines a procedure routine. Procedures never return values when called. This procedure must not be passed any parameters when called.

```
print_sorted_list()
```

This calls the routine called `print_sorted_list`.

The output from the program will be:

```
Number 10 was at position 2, now at 1
Number 19 was at position 1, now at 2
Number 23 was at position 3, now at 3
Number 32 was at position 10, now at 4
Number 41 was at position 4, now at 5
Number 55 was at position 6, now at 6
Number 67 was at position 8, now at 7
Number 76 was at position 9, now at 8
Number 84 was at position 5, now at 9
Number 98 was at position 7, now at 10
```

Note that `merge_sort()` will just as easily sort any list of data items ...

```
{1.5, -9, 1e6, 100}
{"oranges", "apples", "bananas"}
```

This example is stored as `euphoria\tutorial\example.ex`. This is not the fastest way to sort in EUPHORIA. Go to the `euphoria\demo` directory and type

```
eui allsorts
```

to compare timings on several different sorting algorithms for increasing numbers of objects.

For a quick tutorial example of EUPHORIA programming, see `euphoria\demo\bench\filesort.ex`.

3.1.2 Installation

To install EUPHORIA on your machine, first read the file `install.txt`.

Installation simply involves copying the *euphoria* files to your hard disk under a directory named `euphoria`, and then appending the search path of your operating system to this directory.

When installed, the `euphoria` directory will look something like this:

- \euphoria
 - ◆ readme.doc
 - ◆ readme.htm
 - ◆ License.txt
 - ◆ \bin - Interpreters `eui.exe` and `euiw.exe`. Translator `euc.exe`. Or on *Linux/FreeBSD*, Interpreter `eui` and Translator `euc`. There are also utility programs such as `ed.bat`, `guru.bat`, etc.
 - ◆ \include - standard include files
 - ◇ \include\std - the standard EUPHORIA library include files, e.g. `file.e`, `sequence.e`
 - ◆ \source - the complete source code (interpreter, translator, binder)
 - ◆ \docs - plain text, man pages, html documentation files
 - ◆ \tutorial - small tutorial programs to help you learn EUPHORIA
 - ◆ \demo - generic demo programs that run on all platforms
 - ◇ \win32 - WIN32-specific demo programs (optional)

- ◇ \unix - *Linux/FreeBSD/OS X*-specific demo programs (optional)
- ◇ \langwar - language war game for *Linux/FreeBSD/OS X*
- ◇ \bench - benchmark programs

The *Linux* subdirectory is not included in the *Windows* distribution, and the `win32` subdirectories are not included in the *Linux/FreeBSD* distribution. In this manual, directory names are shown using backslash (\). *Linux/FreeBSD* users should substitute forward slash (/).

3.1.2.1 Windows

The installer copies the required files, and modifies the `autoexec.bat` for you. The `euphoria\bin` is then on your search path, and the environment variable **EUDIR** is set to the euphoria directory.

3.1.2.2 Linux and FreeBSD

EUPHORIA may be installed using either the gzipped tarball, or from a distribution specific package, if available.

The gzipped tarball is laid out similarly to the windows directory structure. You'll need to manually edit `/etc/profile` so the `PATH` contains `euphoria\bin`, and either create a `eu.cfg` file or set up **EUDIR** and **EUINC**.

The packaged version installs EUPHORIA in a more *Unix*-like way, putting the executables into `/usr/bin`, `/usr/share/euphoria` and `/usr/share/doc/euphoria`. **Man** pages for `eui`, `euc`, `eub`, `shroud` and `bind` are also installed. It will also create `/etc/euphoria/eu.cfg`, which will point to the standard euphoria include directory in `/usr/share/euphoria/include`.

3.1.3 EUPHORIA programs

EUPHORIA programs can be written with *any* plain text editor. As a convenience EUPHORIA comes with `ed`, an editor written in EUPHORIA, that is handy for editing and executing EUPHORIA programs. Take a look at `\euphoria\demo` and `euphoria\tutorial` to see many example programs.

3.1.4 Running a Program

To run a EUPHORIA program you type the name of the interpreter followed by the filename of the program you want to run. Such as:

```
eui example.ex
```

What you just typed is known as the *command-line*.

Depending on the platform you are using the interpreter could be called:

Executable**Purpose**

eui	General interpreter on Windows and *nix variants
euiw	Console-less Windows interpreter

The command-line may contain extra information. Following your program filename you may add extra words (known as *arguments*) that can be used in your program to customize its behavior. These arguments are read within your program by the built-in function `command_line()`.

Optionally, you may also use [command line switches](#) that are typed between the interpreter name and the program name. Command line switches customize how the interpreter itself behaves.

Unlike many other compilers and interpreters, there is no obligation for any special command-line options for `eui` or `euiw`. Only the name of your EUPHORIA file is expected, and if you don't supply it, EUPHORIA will display all the command line options available.

EUPHORIA doesn't care about your choice of file extensions. By convention, however, console-based applications have an extension of `.ex`, GUI-based applications have an extension of `.exw` and include files have an extension of `.e`. Note that a GUI application is not necessarily a Microsoft Windows (tm) program. A GUI application can exist on Linux, OS X, FreeBSD, etc...

You can redirect standard input and standard output when you run a EUPHORIA program, for example:

```
eui filesort.ex < raw.txt > sorted.txt
```

or simply,

```
eui filesort < raw.txt > sorted.txt
```

For frequently-used programs under *Windows* you might want to make a small `.bat` (batch) file, perhaps called `myprog.bat`, containing two statements like:

```
@echo off
eui myprog.ex %1 %2 %3 %4 %5 %6 %7 %8 %9
```

The first statement turns off echoing of commands to the screen. The second runs `eui myprog.ex` with up to 9 command-line arguments. See [command_line\(\)](#) for an example of how to read these arguments. Having a `.bat` file will save you the minor inconvenience of typing `eui` all the time; i.e., you can just type:

```
myprog
```

instead of:

```
eui myprog
```

Under modern *Unix* variants, you can use `#!/usr/bin/env eui` as the first line of your script file. On older *Unix* variants, you may need to use the full path to `eui`, `#!/usr/local/bin/eui`.

If your program is called `foo.ex`:

```
#!/usr/bin/env eui
```

```
procedure foo()  
  ? 2+2  
end procedure
```

```
foo()
```

Then if you make your file executable:

```
chmod +x foo.ex
```

You can just type:

```
foo.ex
```

to run your program. You could even shorten the name to simply "foo". EUPHORIA ignores the first line when it starts with `#!`. Be careful though that your first line ends with the unix-style `\n`, and not the *Windows*-style `\r\n`, or the unix shell might get confused. If your file is shrouded, you must give the path to `eub`, not `eui`.

You can also run `bind` to combine your EUPHORIA program with `eui` interpreter, to make a stand-alone executable file. With a stand-alone executable, you *can* redirect standard input and output. Binding is discussed further in [Distributing a Program](#).

Using the [EUPHORIA To C Translator](#), you can also make a stand-alone executable file, and it will normally run much faster than a bound program.

3.1.5 Running under Windows

You can run EUPHORIA programs directly from the *Windows* environment, or from a console shell that you have opened from *Windows*. By "associating" `.ex` files with `eui.exe` and `.exw` files with `euiw.exe`. You will then be able to double click a EUPHORIA source file to run it. The installer will perform this operation for you, if you wish.

3.1.6 Editing a Program

You can use any text editor to edit a EUPHORIA program. However, EUPHORIA comes with its own special editor that is written entirely in EUPHORIA. Type: `ed` followed by the complete name of the file you wish to edit. You can use this editor to edit any kind of text file. When you edit a EUPHORIA file, some extra features such as color syntax highlighting and auto-completion of certain statements, are available to make your job easier.

Whenever you run a EUPHORIA program and get an error message, during compilation or execution, you can simply type `ed` with no file name and you will be automatically positioned in the file containing the error, at the correct line and column, and with the error message displayed at the top of the screen.

Under *Windows* you can associate `ed.bat` with various kinds of text files that you want to edit. Color syntax highlighting is provided for `.ex`, `.exw`, `.exd`, `.e` and `.pro` ([profile files](#)).

Most keys that you type are inserted into the file at the cursor position. Hit the `Esc` key once to get a menu bar of special commands. The arrow keys, and the `Insert/Delete/Home/End/PageUp/PageDown` keys are also active. Under *Linux/FreeBSD* some keys may not be available, and alternate keys are provided. See [Ed - EUPHORIA Editor](#) for a complete description of the editing commands.

If you need to understand or modify any detail of the editor's operation, you can edit the file `ed.ex` in `euphoria/bin` (be sure to make a backup copy so you don't lose your ability to edit). If the name `ed` conflicts with some other command on your system, simply rename the file `euphoria/bin/ed.bat` to something else. Because this editor is written in EUPHORIA, it is remarkably concise and easy to understand. The same functionality implemented in a language like C, would take far more lines of code.

`ed` is a simple text-mode editor that runs on all platforms and is distributed with EUPHORIA. There are other editors in [The Archive](#). In addition to that many popular editors include syntax highlighting for EUPHORIA or can be made to.

3.1.7 Distributing a Program

EUPHORIA provides you with 4 distinct ways of distributing a program.

- "source-code", with the EUPHORIA "interpreter"
- "shroud" into `.il` code, with EUPHORIA "backend"
- "bind" into a EUPHORIA executable
- "translate" into a C-compiled executable

In the first method you simply ship your users the interpreter along with your EUPHORIA source files including any EUPHORIA includes that may be necessary from the `euphoria/include` directory. If the EUPHORIA source files and the interpreter are placed together in one directory then your user can run your program by typing `eui` followed by the path of your main executable sort file. You might also provide a small `.bat` file so people won't actually have to type the interpreter name. This method assumes that you are willing to share your EUPHORIA source code with your users.

The Binder gives you two more methods of distribution. You can **shroud** your program, or you can **bind** your program. **Shrouding** combines all of the EUPHORIA source code that your program needs to create a single `.il` file. **Binding** combines your shrouded program with the EUPHORIA backend (`eub` or `eubw`) to create a single, stand-alone executable file. For example, if your program is called `"myprog.ex"` you can create `"myprog.exe"` which will run identically. For more information about shrouding and binding, see [Binding and Shrouding](#).

Finally, with the [EUPHORIA To C Translator](#), you can translate your EUPHORIA program into C and then compile it with a C compiler to get an executable program.

3.2 Using EUPHORIA

3.2.1 Command line switches

You can launch EUPHORIA with some extra command line switches, in order to add or change configuration elements. When running a GUI, there is always some way to open a prompt and enter any text with options, arguments and whatever the program being launched may need for proper, expected operation. Under *Windows*, this is achieved by clicking the `Run . . .` start menu entry, or hitting `Windows-R`.

Command line switches may be changed or added, one at a time.

`batch_command_line`

`-BATCH (all)`

Executes the program but if any error occurs, the "Press Enter" prompt is not presented. The exit code will be set to 1 on error, 0 on success. This option can also be set via the [with batch](#) directive.

`-COPYRIGHT (all)`

Displays the copyright banner for euphoria.

`eu.cfg`

`-C config_file (all)`

Specifies either a file name or the path for where the default file called `eu.cfg` exists. The configuration file which holds a set of additional command line switches.

`-CON (translator)`

Windows only. Specifies that the translated program should be a console application. The default is to build a windowed application.

`-D word (all)`

Defines a word as being set. Words are processed by the [ifdef statement](#). Words can also be defined via the [with / without define](#) directive.

`-DEBUG (translator)`

The translated euphoria code will include

`-DLL, -SO (translator)`

Compiles and links the translated euphoria code into a DLL, SO or DYLIB (depending on the platform).

`-EUDIR dir (all)`

This overrides the environment variable EUDIR.

`-H, (all)`

Displays the list of available command line options.

- `-I include_path (all)`
Specifies an extra include path.
- `-LIB file (translator)`
Specifies the run-time library to use when translating euphoria programs.
- `-PLAT word (translator)`
Specify the target platform for translation. This allows euphoria code to be translated for any supported platform from any other supported platform. Supported platforms: FREEBSD, LINUX, NETBSD, OPENBSD, OSX, SUNOS, WIN
- `-STRICT (all)`
This turns on all warnings, overriding any with/without warning statement found in the source. This option can also be set via the [with/without warning](#) directive.
- `-TEST (all)`
Parses the code only and issues any warnings or errors to `STDOUT`. On error the exit code will be 1, otherwise 0. If an error was found, the normal "Press Enter" prompt will not be presented when using the `-TEST` parameter which enables many editor/IDE programs to test the syntax of your EUPHORIA source in real time.
- `-VERSION (all)`
Displays the version of euphoria that is running.
- `-W warning_name (all)`
Resets, or adds to, the current list of warnings that may be emitted. The list of known names is to be found in the subsection [with/without warning](#). A name should appear without quotes. If the `warning_name` begins with a plus symbol '+', this warning is added to the current set of warnings checked for, otherwise the first usage resets the list to the warning being introduced, and each subsequent `-W warning_name` adds to the list.
- `-WF file_name (all)`
Sets the file where the warnings should go instead of the standard error. Warnings are written to that file regardless of whether or not there are errors in the source. If there are no warnings, the `-wf` file is not created. If the `-wf` file cannot be created, a suitable message is displayed on `STDERR` and written to `ex.err`.
- `-X;`
Resets, or adds to, the list of warnings that will not be issued. This is opposite of the `-W` switch.

The case of the switches is ignored, so `-I` and `-i` are equivalent.

3.2.2 Further notes

- Included files are searched for in all included paths, in the following order:
 1. The current path
 2. Paths specified in a `-I` command line switch, which can also come from any configuration files found.

3. Paths listed in the `EUINC` environment variable, in the order in which they appear
4. Paths listed in the `EUDIR` environment variable, in the order in which they appear
5. The interpreter's path

3.2.3 eu.cfg

EUPHORIA supports reading command line switches from configuration files. The default name for the configuration file is `eu.cfg`. However you can specify different ones by using the `-C` switch.

3.2.3.1 Configuration file format

The file is a text file. Each line in the file is either a command line switch, a section header, an include path or a comment.

- Comments are lines that begin with a double dash `--`. Everything on the line is ignored.
- A section header is a *name* enclosed in square brackets. eg. `[interpret]`.
 - ◆ There are a number of predefined sections.
 - ◆ The lines in a section are only added to the command line switches if they apply to the mode that EUPHORIA is running in.
 - ◇ `[win32]` Applies to Windows platform only.
 - ◇ `[unix]` Applies to any Unix platform only.
 - ◇ `[interpret]` Applies to the interpreter running in any platform.
 - ◇ `[translate]` Applies to the translator running in any platform.
 - ◇ `[bind]` Applies to the binder running in any platform.
 - ◇ `[interpret:win32]` Applies to the interpreter when running under Windows only.
 - ◇ `[interpret:unix]` Applies to the interpreter when running under Unix only.
 - ◇ `[translate:win32]` Applies to the translator when running under Windows only.
 - ◇ `[translate:unix]` Applies to the translator when running under Unix only.
 - ◇ `[bind:win32]` Applies to the binder when running under Windows only.
 - ◇ `[bind:unix]` Applies to the binder when running under Unix only.
 - ◇ `[all]` Applies to all running modes.
 - ◆ All configuration lines before the first section header are assumed to be the `[all]` section.
 - ◆ You can have any number of section headers, but only the predefined ones are used. All lines in other sections are treated as comments.
- A command line switch is a line that begins with a single dash. The entire line is added to the actual command line as if it was originally there.
- An include path is any other line that is not one of the above. The string `-I` is prepended to the line and then it is added to the command line.

When EUPHORIA starts up, it looks for configuration files in the following order ...

- for UNIX systems
 1. `/etc/euphoria/eu.cfg`
 2. `${EUDIR}/eu.cfg`
 3. `${HOME}/eu.cfg`
 4. From where ever the executable is run from "`<exepath>/eu.cfg`"

5. Current working directory - `"/eu.cfg"`
 6. Command line `-C` switches
- for WINDOWS systems
 1. `%ALLUSERSPROFILE%\euphoria\eu.cfg`
 2. `%APPDATA%\euphoria\eu.cfg`
 3. `%EUDIR%\eu.cfg`
 4. `%HOMEDRIVE%\%HOMEPATH%\eu.cfg`
 5. From where ever the executable is run from `"<exepath>/eu.cfg"`
 6. Current working directory - `"/eu.cfg"`
 7. Command line `-C` switches
- EUPHORIA processes every configuration file found, and in the order described above. This means that setting specified in earlier configuration files may be overridden by subsequent configuration files. For example, a configuration file in the current directory will override the same settings in a configuration file in the executable's directory.
 - If a configuration file contains a `"-C"` switch, the new configuration file specified on that switch is processed before subsequent lines in the old file.
 - A configuration file is only ever processed once. Additional references to the same file are ignored.

4 Language Reference

Definition

- Objects
- Identifiers
- Comments
- Expressions
- Precedence Chart

Declarations

- Identifiers
- Specifying the type of a variable
- Scope

Assignment statement

- Assignment with Operator

Branching Statements

- if statement
- switch statement
- ifdef statement

Loop statements

- while statement
- loop until statement
- for statement

Flow control statements

- exit statement
- break statement
- continue statement
- retry statement
- with entry statement
- goto statement
- Header Labels

Short-Circuit Evaluation

Special Top-Level Statements

- include
- with / without

4.1 Definition

4.1.1 Objects

4.1.1.1 Atoms and Sequences

All data **objects** in EUPHORIA are either **atoms** or **sequences**. An **atom** is a single numeric value. A **sequence** is a collection of objects, either atoms or sequences themselves. A sequence can contain any

mixture of atom and sequences; a sequence does not have to contain all the same data type. Because the **objects** contained in a sequence can be an arbitrary mix of atoms or sequences, it is an extremely versatile data structure, capable of representing any sort of data.

A sequence is represented by a list of objects in brace brackets { }, separated by commas. Atoms can have any integer or double-precision floating point value. They can range from approximately -1e300 (minus one times 10 to the power 300) to +1e300 with 15 decimal digits of accuracy. Here are some EUPHORIA objects:

```
-- examples of atoms:
0
1000
98.6
-1e6
23_100_000

-- examples of sequences:
{2, 3, 5, 7, 11, 13, 17, 19}
{1, 2, {3, 3, 3}, 4, {5, {6}}}
{"jon", "smith", 52389, 97.25}
{} -- the 0-element sequence
```

By default, number literals use *base 10*, but you can have integer literals written in other bases, namely binary (*base 2*), octal (*base 8*), and hexadecimal (*base 16*). To do this, the number is prefixed by a 2-character code that lets EUPHORIA know which base to use.

Code	Base
0b	2 = Binary
0t	8 = Octal
0d	10 = Decimal
0x	16 = Hexadecimal

For example:

```
0b101 --> decimal 5
0t101 --> decimal 65
0d101 --> decimal 101
0x101 --> decimal 257
```

Additionally, hexadecimal integers can also be written by prefixing the number with the '#' character.

For example:

```
#FE -- 254
#A000 -- 40960
#FFFF00008 -- 68718428168
-#10 -- -16
```

Only digits and the letters A, B, C, D, E, F, in either uppercase or lowercase, are allowed in hexadecimal numbers. Hexadecimal numbers are always positive, unless you add a minus sign in front of the # character. So for instance #FFFFFFFF is a huge positive number (4294967295), **not** -1, as some machine-language programmers might expect.

Sometimes, and especially with large numbers, it can make reading numeric literals easier when they have embedded grouping characters. We are familiar with using commas (periods in Europe) to group large numbers by 3-digit subgroups. In EUPHORIA we use the underscore character to achieve the same thing, and we can group them anyway that is useful to us.

```
atom big = 32_873_787  -- Set 'big' to the value 32873787

atom salary = 56_110.66 -- Set salary to the value 56110.66

integer defflags = #0323_F3CD

object phone = 61_3_5536_7733

integer bits = 0b11_00010_1
```

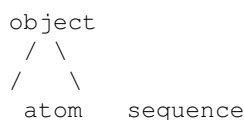
Sequences can be nested to any depth, i.e. you can have sequences within sequences within sequences and so on to any depth (until you run out of memory). Brace brackets are used to construct sequences out of a list of expressions. These expressions can be constant or evaluated at run-time. e.g.

```
{ x+6, 9, y*w+2, sin(0.5) }
```

The "**Hierarchical Objects**" part of the EUPHORIA acronym comes from the hierarchical nature of nested sequences. This should not be confused with the class hierarchies of certain object-oriented languages.

Why do we call them atoms? Why not just "numbers"? Well, an `atom` is just a number, but we wanted to have a distinctive term that emphasizes that they are indivisible (that's what "atom" means in Greek). In the world of physics you can 'split' an atom into smaller parts, but you no longer have an atom--only various particles. You can 'split' a number into smaller parts, but you no longer have a number--only various digits.

Atoms are the basic building blocks of all the data that a EUPHORIA program can manipulate. With this analogy, **sequences** might be thought of as "molecules", made from atoms and other molecules. A better analogy would be that sequences are like directories, and atoms are like files. Just as a directory on your computer can contain both files and other directories, a sequence can contain both atoms and other sequences (and *those* sequences can contain atoms and sequences and so on).



As you will soon discover, sequences make EUPHORIA very simple *and* very powerful. **Understanding atoms and sequences is the key to understanding EUPHORIA.**

Performance Note:

Does this mean that all atoms are stored in memory as 8-byte floating-point numbers? No. The EUPHORIA interpreter usually stores integer-valued atoms as machine integers (4 bytes) to save space and improve execution speed. When fractional results occur or integers get too big, conversion to IEEE 8-byte floating-point format happens automatically.

4.1.1.2 Character Strings and Individual Characters

A **character string** is just a sequence of characters. It may be entered in a number of ways ...

- using double-quotes e.g.

```
"ABCDEFGH"
```

- using raw string notation e.g.

```
-- Using back-quotes  
`ABCDEFGH`
```

or

```
-- Using three double-quotes  
"""ABCDEFGH"""
```

- using hexadecimal strings e.g.

```
x"65 66 67 AE"
```

- using unicode strings e.g.

```
u"65 66 67 AE"  -- utf-16 ==> {25958, 26542}  
U"65 66 67 AE"  -- utf-32 ==> {1701210030}
```

The rules for double-quote strings are...

1. they begin and end with a double-quote character
2. they cannot contain a double-quote
3. they must be only on a single line
4. they cannot contain the TAB character
5. if they contain the back-slash `'\'` character, that character must immediately be followed by one of the special *escape* codes. The back-slash and escape code will be replaced by the appropriate single character equivalent. If you need to include double-quote, end-of-line, back-slash, or TAB characters inside a double-quoted string, you need to enter them in a special manner.

e.g.

```
"Bill said\n\t\"This is a back-slash \\ character\".\n"
```

Which, when displayed should look like ...

```
Bill said  
  "This is a back-slash \ character".
```

The rules for raw strings are...

1. enclose with three double-quotes `"""..."""` or back-quote. ``...``

2. The resulting string will never have any carriage-return characters in it.
3. If the resulting string begins with a new-line, the initial new-line is removed and any trailing new-line is also removed.
4. A special form is used to automatically remove leading whitespace from the source code text. You might code this form to align the source text for ease of reading. If the first line after the raw string start token begins with one or more underscore characters, the number of consecutive underscores signifies the maximum number of whitespace characters that will be removed from each line of the raw string text. The underscores represent an assumed left margin width. **Note**, these leading underscores do not form part of the raw string text.

e.g.

```
-- No leading underscores and no leading whitespace
\
Bill said
    "This is a back-slash \ character".
\
```

Which, when displayed should look like ...

```
Bill said
    "This is a back-slash \ character".

-- No leading underscores and but leading whitespace
\
    Bill said
        "This is a back-slash \ character".
\
```

Which, when displayed should look like ...

```
Bill said
    "This is a back-slash \ character".

-- Leading underscores and leading whitespace
\
____Bill said
    "This is a back-slash \ character".
\
```

Which, when displayed should look like ...

```
Bill said
    "This is a back-slash \ character".
```

Extended string literals are useful when the string contains new-lines, tabs, or back-slash characters because they do not have to be entered in the special manner. The back-quote form can be used when the string literal contains a set of three double-quote characters, and the triple quote form can be used when the text literal contains back-quote characters. If a literal contains both a back quote and a set of three double-quotes, you will need to concatenate two literals.

```
constant TQ = `This text contains "" for some reason.`
constant BQ = """"This text contains a back quote ` for some reason."""
```

```
constant QQ = """This text contains a back quote `""" & `and """ for some reason.`
```

The rules for hexadecimal strings are...

1. they begin with the pair x" and end with a double-quote (") character
2. they can only contain hexadecimal digits (0-9 A-F a-f), and space, underscore, tab, newline, carriage-return. Anything else is invalid.
3. each pair of contiguous hex digits represent a single sequence element with a value from 0 to 255
4. they can span multiple lines
5. The non-hex digits are treated as punctuation and used to delimit individual values.
6. Hexadecimal strings can be used to encode UTF-8 strings, even though the resulting string does not have to be a valid UTF-8 string.

```
x"1 2 34 5678_AbC" == {0x01, 0x02, 0x34, 0x56, 0x78, 0xAB, 0x0C}
```

The rules for unicode strings are...

1. they begin with the pair u" for UTF-16 and U" for UTF-32 strings, and end with a double-quote (") character
2. they can only contain hexadecimal digits (0-9 A-F a-f), and space, underscore, tab, newline, carriage-return. Anything else is invalid.
3. For UTF-16 strings, each set of four contiguous hex digits represent a single sequence element with a value from 0x0000 to 0xFFFF
4. For UTF-32 strings, each set of eight contiguous hex digits represent a single sequence element with a value from 0x0000 to 0xFFFFFFFF
5. they can span multiple lines
6. The non-hex digits are treated as punctuation and used to delimit individual values.
7. The resulting string does not have to be a valid UTF-16/UTF-32 string.

```
u"1 2 34 5678AbC" == {0x0001, 0x0002, 0x0034, 0x5678, 0x0ABC}
u"1 2 34 5678AbC" == {0x0000_0001, 0x0000_0002, 0x0000_0034, 0x05678ABC}
```

Character strings may be manipulated and operated upon just like any other sequences. For example the above string is entirely equivalent to the sequence:

```
{65, 66, 67, 68, 69, 70, 71}
```

which contains the corresponding ASCII codes. The EUPHORIA compiler will immediately convert "ABCDEFGH" to the above sequence of numbers. In a sense, there are no "strings" in EUPHORIA, only sequences of numbers. A quoted string is really just a convenient notation that saves you from having to type in all the ASCII codes. emptyseq It follows that "" is equivalent to {}. Both represent the sequence of length-0, also known as the **empty sequence**. As a matter of programming style, it is natural to use "" to suggest a length-0 sequence of characters, and {} to suggest some other kind of sequence.

An **individual character** is an **atom**. It must be entered using single quotes. There is a difference between an individual character (which is an atom), and a character string of length-1 (which is a sequence). e.g.

```
'B'  -- equivalent to the atom 66 - the ASCII code for B
"B"  -- equivalent to the sequence {66}
```

Again, 'B' is just a notation that is equivalent to typing 66. There aren't really any "characters" in EUPHORIA, just numbers (atoms).

Keep in mind that an atom is *not* equivalent to a one-element sequence containing the same value, although there are a few built-in routines that choose to treat them similarly.

4.1.1.3 Escaped Characters

Special characters may be entered using a back-slash:

Code	Meaning
<code>\n</code>	newline
<code>\r</code>	carriage return
<code>\t</code>	tab
<code>\\</code>	backslash
<code>\"</code>	double quote
<code>\'</code>	single quote
<code>\0</code>	null
<code>\e</code>	escape
<code>\E</code>	escape
<code>\x/hh/</code>	A 2-hex-digit value, e.g. <code>"\x5F" ==> {95}</code>
<code>\u/hhhh/</code>	A 4-hex-digit value, e.g. <code>"\u2A7C" ==> {10876}</code>
<code>\U/hhhhhhhh/</code>	An 8-hex-digit value, e.g. <code>"\U8123FEDC" ==> {2166619868}</code>

For example, `"Hello, World!\n"`, or `'\\'`. The EUPHORIA editor displays character strings in green.

4.1.2 Identifiers

An identifier is just the name you give something in your program. This can be a variable, constant, function, procedure, parameter, or namespace. An identifier must begin with either a letter or an underscore, then followed by zero or more letters, digits or underscore characters. There is no theoretical limit to how large an identifier can be but in practice it should be no more than about 30 characters.

Identifiers are case-sensitive. This means that `"Name"` is a different identifier from `"name"`, or `"NAME"`, etc...

Examples of valid identifiers...

```
n
color26
ShellSort
quick_sort
a_very_long_identifier_that_is_really_too_long_for_its_own_good
_alpha
```

Examples of invalid identifiers...

```
0n      -- must not start with a digit
^color26 -- must not start with a punctuation character
Shell Sort -- Cannot have spaces in identifiers.
quick-sort -- must only consist of letters, digits or underscore.
```

4.1.3 Comments

Comments are ignored by EUPHORIA and have no effect on execution speed. The editor displays comments in red.

There are three forms of comment text.

- The *line* format comment is started by two dashes and extends to the end of the current line.

e.g.

```
-- This is a comment which extends to the end of this line only.
```

- The *multi-line* format comment is started by `/*` and extends to the next occurrence of `*/`, even if that occurs on a different line.

e.g.

```
/* This is a comment which
   extends over a number
   of text lines.
*/
```

- On the first line only of your program, you can use a special comment beginning with the two character sequence `#!`. This is mainly used to tell Linux shells which program to execute the 'script' program with.

e.g.

```
#!/home/rob/euphoria/bin/eui
```

This informs the Linux shell that your file should be executed by the EUPHORIA interpreter, and gives the full path to the interpreter. If you make your file executable, you can run it, just by typing its name, and without the need to type `"exi"`. On *Windows* this line is just treated as a comment (though Apache Web server on *Windows* does recognize it.). If your file is a shrouded `.il` file, use `eub.exe` instead of `eui`.

Line comments are typically used to annotate a single (or small section) of code, whereas multi-line comments are typically used to give larger pieces of documentation inside the source text.

4.1.4 Expressions

Like other programming languages, EUPHORIA lets you calculate results by forming expressions. However, in EUPHORIA you can perform calculations on entire sequences of data with one expression, where in most

other languages you would have to construct a loop. In EUPHORIA you can handle a sequence much as you would a single number. It can be copied, passed to a subroutine, or calculated upon as a unit. For example,

```
{1, 2, 3} + 5
```

is an expression that adds the sequence {1, 2, 3} and the atom 5 to get the resulting sequence {6, 7, 8}.

We will see more examples later.

4.1.4.1 Relational Operators

The relational operators `<` `>` `<=` `>=` `=` `!=` each produce a 1 (true) or a 0 (false) result.

```
8.8 < 8.7    -- 8.8 less than 8.7 (false)
-4.4 > -4.3   -- -4.4 greater than -4.3 (false)
8 <= 7        -- 8 less than or equal to 7 (false)
4 >= 4        -- 4 greater than or equal to 4 (true)
1 = 10       -- 1 equal to 10 (false)
8.7 != 8.8    -- 8.7 not equal to 8.8 (true)
```

As we will soon see you can also apply these operators to sequences.

4.1.4.2 Logical Operators

The logical operators `and`, `or`, `xor`, and `not` are used to determine the "truth" of an expression. e.g.

```
1 and 1      -- 1 (true)
1 and 0      -- 0 (false)
0 and 1      -- 0 (false)
0 and 0      -- 0 (false)

1 or 1       -- 1 (true)
1 or 0       -- 1 (true)
0 or 1       -- 1 (true)
0 or 0       -- 0 (false)

1 xor 1      -- 0 (false)
1 xor 0      -- 1 (true)
0 xor 1      -- 1 (true)
0 xor 0      -- 0 (false)

not 1        -- 0 (false)
not 0        -- 1 (true)
```

You can also apply these operators to numbers other than 1 or 0. The rule is: zero means false and non-zero means true. So for instance:

```
5 and -4     -- 1 (true)
not 6        -- 0 (false)
```

These operators can also be applied to sequences. See below.

In some cases [short-circuit](#) evaluation will be used for expressions containing `and` or `or`. Specifically, short circuiting applies inside decision taking statements, like the [if statement](#), [while statement](#) or [loop statement](#). More on this later.

4.1.4.3 Arithmetic Operators

The usual arithmetic operators are available: add, subtract, multiply, divide, unary minus, unary plus.

```
3.5 + 3  -- 6.5
3 - 5    -- -2
6 * 2    -- 12
7 / 2    -- 3.5
-8.1     -- -8.1
+8       -- +8
```

Computing a result that is too big (i.e. outside of -1e300 to +1e300) will result in one of the special atoms **+infinity** or **-infinity**. These appear as `inf` or `-inf` when you print them out. It is also possible to generate `nan` or `-nan`. "nan" means "not a number", i.e. an undefined value (such as `inf` divided by `inf`). These values are defined in the IEEE floating-point standard. If you see one of these special values in your output, it usually indicates an error in your program logic, although generating `inf` as an intermediate result may be acceptable in some cases. For instance, `1/inf` is 0, which may be the "right" answer for your algorithm.

Division by zero, as well as bad arguments to math library routines, e.g. square root of a negative number, log of a non-positive number etc. cause an immediate error message and your program is aborted.

The only reason that you might use unary plus is to emphasize to the reader of your program that a number is positive. The interpreter does not actually calculate anything for this.

4.1.4.4 Operations on Sequences

All of the relational, logical and arithmetic operators described above, as well as the math routines described in [Language Reference](#), can be applied to sequences as well as to single numbers (atoms).

When applied to a sequence, a unary (one operand) operator is actually applied to each element in the sequence to yield a sequence of results of the same length. If one of these elements is itself a sequence then the same rule is applied again recursively. e.g.

```
x = -{1, 2, 3, {4, 5}}  -- x is {-1, -2, -3, {-4, -5}}
```

If a binary (two-operand) operator has operands which are both sequences then the two sequences must be of the same length. The binary operation is then applied to corresponding elements taken from the two sequences to get a sequence of results. e.g.

```
x = {5, 6, 7, 8} + {10, 10, 20, 100}
-- x is {15, 16, 27, 108}
x = {{1, 2, 3}, {4, 5, 6}} + {-1, 0, 1} -- ERROR: 2 != 3
-- but
x = {{1, 2, 3} + {-1, 0, 1}, {4, 5, 6} + {-1, 0, 1}} -- CORRECT
-- x is {{0, 2, 4}, {3, 5, 7}}
```

If a binary operator has one operand which is a sequence while the other is a single number (atom) then the single number is effectively repeated to form a sequence of equal length to the sequence operand. The rules for operating on two sequences then apply. Some examples:

```
y = {4, 5, 6}
w = 5 * y          -- w is {20, 25, 30}

x = {1, 2, 3}
z = x + y          -- z is {5, 7, 9}
z = x < y          -- z is {1, 1, 1}

w = {{1, 2}, {3, 4}, {5}}
w = w * y          -- w is {{4, 8}, {15, 20}, {30}}

w = {1, 0, 0, 1} and {1, 1, 1, 0} -- {1, 0, 0, 0}
w = not {1, 5, -2, 0, 0}         -- w is {0, 0, 0, 1, 1}

w = {1, 2, 3} = {1, 2, 4}        -- w is {1, 1, 0}

-- note that the first '=' is assignment, and the
-- second '=' is a relational operator that tests
-- equality
```

Note: When you wish to compare two strings (or other sequences), you should **not** (as in some other languages) use the '=' operator:

```
if "APPLE" = "ORANGE" then -- ERROR!
```

'=' is treated as an operator, just like '+', '*' etc., so it is applied to corresponding sequence elements, and the sequences must be the same length. When they are equal length, the result is a sequence of ones and zeros. When they are not equal length, the result is an error. Either way you'll get an error, since an if-condition must be an atom, not a sequence. Instead you should use the `equal()` built-in routine:

```
if equal("APPLE", "ORANGE") then -- CORRECT
```

In general, you can do relational comparisons using the `compare()` built-in routine:

```
if compare("APPLE", "ORANGE") = 0 then -- CORRECT
```

You can use `compare()` for other comparisons as well:

```
if compare("APPLE", "ORANGE") < 0 then -- CORRECT
-- enter here if "APPLE" is less than "ORANGE" (TRUE)
```

Especially useful is the idiom `compare(x, "") = 1` to determine whether `x` is a non empty sequence. `compare(x, "") = -1` would test for `x` being an atom, but `atom(x) = 1` does the same faster and is clearer to read.

4.1.4.5 Subscripting of Sequences

A single element of a sequence may be selected by giving the element number in square brackets. Element numbers start at 1. Non-integer subscripts are rounded down to an integer.

For example, if `x` contains `{5, 7.2, 9, 0.5, 13}` then `x[2]` is `7.2`. Suppose we assign something different to `x[2]`:

```
x[2] = {11, 22, 33}
```

Then `x` becomes: `{5, {11, 22, 33}, 9, 0.5, 13}`. Now if we ask for `x[2]` we get `{11, 22, 33}` and if we ask for `x[2][3]` we get the atom `33`. If you try to subscript with a number that is outside of the range 1 to the number of elements, you will get a subscript error. For example `x[0]`, `x[-99]` or `x[6]` will cause errors. So will `x[1][3]` since `x[1]` is not a sequence. There is no limit to the number of subscripts that may follow a variable, but the variable must contain sequences that are nested deeply enough. The two dimensional array, common in other languages, can be easily represented with a sequence of sequences:

```
x = {
    {5, 6, 7, 8, 9},      -- x[1]
    {1, 2, 3, 4, 5},      -- x[2]
    {0, 1, 0, 1, 0}       -- x[3]
}
```

where we have written the numbers in a way that makes the structure clearer. An expression of the form `x[i][j]` can be used to access any element.

The two dimensions are not symmetric however, since an entire "row" can be selected with `x[i]`, but you need to use `vslice()` in the Standard Library to select an entire column. Other logical structures, such as n-dimensional arrays, arrays of strings, structures, arrays of structures etc. can also be handled easily and flexibly:

3-D array:

```
y = {
    {{1,1}, {3,3}, {5,5}},
    {{0,0}, {0,1}, {9,1}},
    {{-1,9}, {1,1}, {2,2}}
}
```

```
-- y[2][3][1] is 9
```

Array of strings:

```
s = {"Hello", "World", "EUPHORIA", "", "Last One"}

-- s[3] is "EUPHORIA"
-- s[3][1] is 'E'
```

A Structure:

```
employee = {
    {"John", "Smith"},
    45000,
    27,
    185.5
}
```


To access "fields" or elements within a structure it is good programming style to make up an enum that names the various fields. This will make your program easier to read. For the example above you might have:

```
enum NAME, FIRST_NAME, LAST_NAME, SALARY, AGE, WEIGHT

employees = {
    {"John", "Smith"}, 45000, 27, 185.5}, -- a[1]
    {"Bill", "Jones"}, 57000, 48, 177.2}, -- a[2]
    -- .... etc.
}

-- employees[2][SALARY] would be 57000.
```

The `length()` built-in function will tell you the length of a sequence. So the last element of a sequence `s`, is:

```
s[length(s)]
```

A short-hand for this is:

```
s[$]
```

Similarly,

```
s[length(s)-1]
```

can be simplified to:

```
s[$-1]
```

The `$` symbol equals the length of the sequence. `$` may only appear between square braces. Where there's nesting, e.g.:

```
s[$ - t[$-1] + 1]
```

The first `$` above refers to the length of `s`, while the second `$` refers to the length of `t` (as you'd probably expect). An example where `$` can save a lot of typing, make your code clearer, and probably even faster is:

```
longname[$][$] -- last element of the last element
```

Compare that with the equivalent:

```
longname[length(longname)][length(longname[length(longname)])]
```

Subscripting and function side-effects:

In an assignment statement, with left-hand-side subscripts:

```
lhs_var[lhs_expr1][lhs_expr2]... = rhs_expr
```

The expressions are evaluated, and any subscripting is performed, from left to right. It is possible to have function calls in the right-hand-side expression, or in any of the left-hand-side expressions. If a function call

has the side-effect of modifying the `lhs_var`, it is not defined whether those changes will appear in the final value of the `lhs_var`, once the assignment has been completed. To be sure about what is going to happen, perform the function call in a separate statement, i.e. do not try to modify the `lhs_var` in two different ways in the same statement. Where there are no left-hand-side subscripts, you can always assume that the final value of the `lhs_var` will be the value of `rhs_expr`, regardless of any side-effects that may have changed `lhs_var`.

EUPHORIA data structures are almost infinitely flexible.

Arrays in many languages are constrained to have a fixed number of elements, and those elements must all be of the same type. EUPHORIA eliminates both of those restrictions by defining all arrays (sequences) as a list of zero or more EUPHORIA objects whose element count can be changed at any time. You can easily add a new structure to the employee sequence above, or store an unusually long name in the NAME field and EUPHORIA will take care of it for you. If you wish, you can store a variety of different employee "structures", with different sizes, all in one sequence. However, when you retrieve a sequence element, it is not guaranteed to be of any type. You, as a programmer, need to check that the retrieved data is of the type you'd expect, EUPHORIA will not. The only thing it will check is whether an assignment is legal. For example, if you try to assign a sequence to an integer variable, EUPHORIA will complain at the time your code does the assignment.

Not only can a EUPHORIA program represent all conventional data structures but you can create very useful, flexible structures that would be hard to declare in many other languages.

Note that expressions in general may not be subscripted, just variables. For example:

`{5+2, 6-1, 7*8, 8+1}[3]` is *not* supported, nor is something like: `date()[MONTH]`. You have to assign the sequence returned by `date()` to a variable, then subscript the variable to get the month.

4.1.4.6 Slicing of Sequences

A sequence of consecutive elements may be selected by giving the starting and ending element numbers. For example if `x` is `{1, 1, 2, 2, 2, 1, 1, 1}` then `x[3..5]` is the sequence `{2, 2, 2}`. `x[3..3]` is the sequence `{2}`. `x[3..2]` is also allowed. It evaluates to the length-0 sequence `{}`. If `y` has the value: `{"fred", "george", "mary"}` then `y[1..2]` is `{"fred", "george"}`.

We can also use slices for overwriting portions of variables. After `x[3..5] = {9, 9, 9}` `x` would be `{1, 1, 9, 9, 9, 1, 1, 1}`. We could also have said `x[3..5] = 9` with the same effect. Suppose `y` is `{0, "EUPHORIA", 1, 1}`. Then `y[2][1..4]` is "Euph". If we say `y[2][1..4] = "ABCD"` then `y` will become `{0, "ABCDoria", 1, 1}`.

In general, a variable name can be followed by 0 or more subscripts, followed in turn by 0 or 1 slices. Only variables may be subscripted or sliced, not expressions.

We need to be a bit more precise in defining the rules for **empty slices**. Consider a slice `s[i..j]` where `s` is of length `n`. A slice from `i` to `j`, where `j = i - 1` and `i >= 1` produces the **empty sequence**, even if `i = n + 1`. Thus `1..0` and `n + 1..n` and everything in between are legal (**empty**) slices. Empty slices are quite useful in many algorithms. A slice from `i` to `j` where `j < i - 1` is illegal, i.e. "reverse" slices such as `s[5..3]` are not allowed.

We can also use the `$` shorthand with slices, e.g.

```
s[2..$]
s[5..$-2]
s[$-5..$]
s[$][1..floor($/2)] -- first half of the last element of s
```

4.1.4.7 Concatenation of Sequences and Atoms - The '&' Operator

Any two objects may be concatenated using the **&** operator. The result is a sequence with a length equal to the sum of the lengths of the concatenated objects (where atoms are considered here to have length 1). e.g.

```
{1, 2, 3} & 4          -- {1, 2, 3, 4}
4 & 5                  -- {4, 5}
{{1, 1}, 2, 3} & {4, 5} -- {{1, 1}, 2, 3, 4, 5}
x = {}
y = {1, 2}
y = y & x              -- y is still {1, 2}
```

You can delete element *i* of any sequence *s* by concatenating the parts of the sequence before and after *i*:

```
s = s[1..i-1] & s[i+1..length(s)]
```

This works even when *i* is 1 or `length(s)`, since `s[1..0]` is a legal empty slice, and so is `s[length(s)+1..length(s)]`.

4.1.4.8 Sequence-Formation

Finally, sequence-formation, using braces and commas:

```
{a, b, c, ... }
```

is also an operator. It takes *n* operands, where *n* is 0 or more, and makes an *n*-element sequence from their values. e.g.

```
x = {apple, orange*2, {1,2,3}, 99/4+foobar}
```

The sequence-formation operator is listed at the bottom of the a [precedence chart](#).

4.1.4.9 Other Operations on Sequences

Some other important operations that you can perform on sequences have English names, rather than special characters. These operations are built-in to **eui.exe/euiw**, so they'll always be there, and so they'll be fast. They are described in detail in the [Language Reference](#), but are important enough to EUPHORIA programming that we should mention them here before proceeding. You call these operations as if they were subroutines, although they are actually implemented much more efficiently than that.

4.1.4.9.1 length(sequence s)

Returns the length of a sequence s1.

This is the number of elements in s1. Some of these elements may be sequences that contain elements of their own, but `length` just gives you the "top-level" count. You'll get an error if you ask for the length of an atom. e.g.

```
length({5,6,7})           -- 3
length({1, {5,5,5}, 2, 3}) -- 4 (not 6!)
length({})                -- 0
length(5)                 -- error!
```

4.1.4.9.2 repeat(object o1, integer times)

Returns a sequence that consists of an item repeated count times. e.g.

```
repeat(0, 100)           -- {0,0,0,...,0}   i.e. 100 zeros
repeat("Hello", 3)       -- {"Hello", "Hello", "Hello"}
repeat(99, 0)            -- {}
```

The item to be repeated can be any atom or sequence.

4.1.4.9.3 append(sequence s1, object o1)

Returns a sequence by adding an object o1 to the end of a sequence s1.

```
append({1,2,3}, 4)       -- {1,2,3,4}
append({1,2,3}, {5,5,5}) -- {1,2,3,{5,5,5}}
append({}, 9)            -- {9}
```

The length of the new sequence is always 1 greater than the length of the original sequence. The item to be added to the sequence can be any atom or sequence.

4.1.4.9.4 prepend(sequence s1, object o1)

Returns a new sequence by adding an element to the beginning of a sequence s. e.g.

```
append({1,2,3}, 4)       -- {1,2,3,4}
prepend({1,2,3}, 4)       -- {4,1,2,3}

append({1,2,3}, {5,5,5}) -- {1,2,3,{5,5,5}}
prepend({}, 9)           -- {9}
append({}, 9)            -- {9}
```

The length of the new sequence is always 1 greater than the length of the original sequence. The item to be added to the sequence can be any atom or sequence.

These two built-in functions, `append()` and `prepend()`, have some similarities to the concatenate operator, `&`, but there are clear differences. e.g.

```
-- appending a sequence is different
append({1,2,3}, {5,5,5})  -- {1,2,3,{5,5,5}}
{1,2,3} & {5,5,5}         -- {1,2,3,5,5,5}

-- appending an atom is the same
append({1,2,3}, 5)        -- {1,2,3,5}
{1,2,3} & 5                -- {1,2,3,5}
```

4.1.4.9.5 insert(sequence in_what, object what, atom position)

This function takes a target sequence, `in_what`, shifts its tail one notch and plugs the object `what` in the hole just created. The modified sequence is returned. For instance:

```
s = insert("Joe", 'h', 3) -   - s id "Johe", another string
s = insert("Joe", "h", 3)   -- s is {'J','o',{'h'},'e'}, not a string
s = insert({1,2,3}, 4, -0.5) -- s is {4,1,2,3}, like prepend()
s = insert({1,2,3}, 4, 8.5)  -- s is {1,2,3,4}, like append()
```

The length of the returned sequence is 1 more than the one of `in_what`. This is the same rule as for `append()` and `prepend()` above, which are actually special cases of `insert()`.

4.1.4.9.6 splice(sequence in_what, object what, atom position)

If `what` is an atom, this is the same as `insert()`. But if `what` is a sequence, that sequence is inserted as successive elements into `in_what` at `position`. Example:

```
s = splice("Joe", 'h', 3) -- s is "Johe", like insert()
s = splice("Joe", "hn Do", 3) -- s is "John Doe", another string
s = splice("Joh", "n Doe", 9.3) -- s is "John Doe", like with the & operator
s = splice({1,2,3}, 4, -2) -- s is {4,1,2,3}, like with the & operator in reversed order
```

The length of `splice(in_what, what, position)` always is `length(in_what) + length(what)`, like for concatenation using `&`. For this purpose, atoms behave as if they were of length 1.

4.1.5 Precedence Chart

When two or more operators follow one another in an expression, there must be rules to tell in which order they should be evaluated, as different orders usually lead to different results. It is common and convenient to use a **precedence order** on operators. Operators with the highest degree of precedence are evaluated first, then those with highest precedence among what remains, and so on.

The precedence of operators in expressions is as follows:

highest precedence

```
**highest precedence**
```

4.1.4 Expressions

```
function/type calls
unary-  unary+  not
*  /
+  -
&
<  >  <=  >=  =  !=
and  or  xor
```

lowest precedence

```
{ , , , }
```

Thus $2+6*3$ means $2+(6*3)$ rather than $(2+6)*3$. Operators on the same line above have equal precedence and are evaluated left to right. You can force any order of operations by placing round brackets () around an expression. For instance, $6/3*5$ is $2*5$, not $6/15$.

Different languages or contexts may have slightly different precedence rules. You should be careful when translating a formula from a language to another; EUPHORIA is no exception. Adding superfluous parentheses to explicitly denote the exact order of evaluation doesn't cost much, and may help either readers used to some other precedence chart or translating to or from another context with slightly different rules. Watch out for `and` and `or`, or `*` and `/`.

The equals symbol '=' used in an [assignment statement](#) is not an operator, it's just part of the syntax of the language.

4.2 Declarations

4.2.1 Identifiers

Identifiers, which encompass all explicitly declared variable, constant or routine names, may be of any length. Upper and lower case are distinct. Identifiers must start with a letter or underscore and then be followed by any combination of letters, digits and underscores. The following **reserved words** have special meaning in EUPHORIA and cannot be used as identifiers:

<code>and</code>	<code>export</code>	<code>public</code>
<code>as</code>	<code>fallthru</code>	<code>retry</code>
<code>break</code>	<code>for</code>	<code>return</code>
<code>by</code>	<code>function</code>	<code>routine</code>
<code>case</code>	<code>global</code>	<code>switch</code>
<code>constant</code>	<code>goto</code>	<code>then</code>
<code>continue</code>	<code>if</code>	<code>to</code>
<code>do</code>	<code>ifdef</code>	<code>type</code>
<code>else</code>	<code>include</code>	<code>until</code>
<code>elseif</code>	<code>label</code>	<code>while</code>
<code>elsif</code>	<code>loop</code>	<code>with</code>
<code>elseifdef</code>	<code>namespace</code>	<code>without</code>
<code>end</code>	<code>not</code>	<code>xor</code>
<code>entry</code>	<code>or</code>	
<code>enum</code>	<code>override</code>	

`exit` `procedure`

The EUPHORIA editor displays these words in blue

The following are EUPHORIA built-in routines. It is best if you do not use these for your own identifiers:

<code>?</code>	<code>peek</code>	<code>peek2u</code>	<code>sprintf</code>
<code>abort</code>	<code>peek2s</code>	<code>peek4s</code>	<code>sqrt</code>
<code>and_bits</code>	<code>log</code>	<code>peek4u</code>	<code>system</code>
<code>append</code>	<code>machine_func</code>	<code>peek_string</code>	<code>system_exec</code>
<code>arctan</code>	<code>machine_proc</code>	<code>peeks</code>	<code>tail</code>
<code>atom</code>	<code>match</code>	<code>pixel</code>	<code>tan</code>
<code>c_func</code>	<code>match_from</code>	<code>platform</code>	<code>task_clock_start</code>
<code>c_proc</code>	<code>get_key</code>	<code>poke</code>	<code>task_clock_stop</code>
<code>call</code>	<code>get_pixel</code>	<code>poke2</code>	<code>task_create</code>
<code>call_func</code>	<code>getc</code>	<code>poke4</code>	<code>task_list</code>
<code>call_proc</code>	<code>getenv</code>	<code>position</code>	<code>task_schedule</code>
<code>clear_screen</code>	<code>gets</code>	<code>power</code>	<code>task_self</code>
<code>close</code>	<code>hash</code>	<code>prepend</code>	<code>task_status</code>
<code>command_line</code>	<code>head</code>	<code>print</code>	<code>task_suspend</code>
<code>compare</code>	<code>include_paths</code>	<code>printf</code>	<code>task_yield</code>
<code>cos</code>	<code>insert</code>	<code>profile</code>	<code>time</code>
<code>date</code>	<code>mem_copy</code>	<code>profile_time</code>	<code>trace</code>
<code>delete</code>	<code>mem_set</code>	<code>puts</code>	<code>warning</code>
<code>delete_routine</code>	<code>not_bits</code>	<code>rand</code>	<code>xor_bits</code>
<code>equal</code>	<code>object</code>	<code>remainder</code>	
<code>find</code>	<code>open</code>	<code>remove</code>	
<code>find_from</code>	<code>option_switches</code>	<code>routine_id</code>	
<code>floor</code>	<code>or_bits</code>	<code>sequence</code>	
<code>integer</code>	<code>repeat</code>	<code>sin</code>	
<code>length</code>	<code>replace</code>	<code>splice</code>	

Identifiers can be used in naming the following:

- procedures
- functions
- types
- variables
- constants
- enums

4.2.1.1 procedures

These perform some computation and may have a list of parameters, e.g.

```
procedure empty()  
end procedure  
  
procedure plot(integer x, integer y)  
    position(x, y)  
    puts(1, '**')  
end procedure
```

There are a fixed number of named parameters, but this is not restrictive since any parameter could be a variable-length sequence of arbitrary objects. In many languages variable-length parameter lists are impossible. In C, you must set up strange mechanisms that are complex enough that the average programmer cannot do it without consulting a manual or a local guru.

A copy of the value of each argument is passed in. The formal parameter variables may be modified inside the procedure but this does not affect the value of the arguments. Pass by reference can be achieved using indexes into some fixed sequence.

****Performance Note**

****** :The interpreter does not actually copy sequences or floating-point numbers unless it becomes necessary. For example,

```
y = {1,2,3,4,5,6,7,8.5,"ABC"}
x = y
```

The statement `x = y` does not actually cause a new copy of `y` to be created. Both `x` and `y` will simply "point" to the same sequence. If we later perform `x[3] = 9`, then a separate sequence will be created for `x` in memory (although there will still be just one shared copy of `8.5` and `"ABC"`). The same thing applies to "copies" of arguments passed in to subroutines.

For a number of procedures or functions--see below--some parameters may have the same value in many cases. The most expected value for any parameter may be given a default value. Omitting the value of such a parameter on a given call will cause its default value to be passed.

```
procedure foo(sequence s, integer n=1)
  ? n + length(s)
end procedure

foo("abc")      -- prints out 4 = 3 + 1. n was not specified, so was set to 1.
foo("abc", 3)   -- prints out 6 = 3 + 3
```

This is not limited to the last parameter(s):

```
procedure bar(sequence s="abc", integer n, integer p=1)
  ? length(s)+n+p
end procedure

bar(, 2)        -- prints out 6 = 3 + 2 + 1
bar(2)          -- errors out, as 2 is not a sequence
bar(, 2,)       -- same as bar(,2)
bar(, 2, 3)     -- prints out 8 = 3 + 2 + 3
bar({}, 2, )    -- prints out 3 = 0 + 2 + 1
bar()           -- errors out, second parameter is omitted, but doesn't have a default value
```

Any expression may be used in a default value. Parameters that have been already mentioned may even be part of the expression:

```
procedure baz(sequence s, integer n=length(s))
  ? n
end procedure

baz("abcd") -- prints out 4
```


4.2.1.2 functions

These are just like procedures, but they return a value, and can be used in an expression, e.g.

```
function max(atom a, atom b)
  if a >= b then
    return a
  else
    return b
  end if
end function
```

Any EUPHORIA object can be returned. You can, in effect, have multiple return values, by returning a sequence of objects. e.g.

```
return {x_pos, y_pos}
```

However, EUPHORIA does not have variable lists. When you return a sequence, you still have to dispatch its contents to variables as needed. And you cannot pass a sequence of parameters to a routine, unless using [call_func](#) or [call_proc](#), which carries a performance penalty.

We will use the general term "subroutine", or simply "routine" when a remark is applicable to both procedures and functions.

Defaulted parameters can be used in functions exactly as they are in procedures. See the section above for a few examples.

4.2.1.3 types

These are special functions that may be used in declaring the allowed values for a variable. A type must have exactly one parameter and should return an atom that is either true (non-zero) or false (zero). Types can also be called just like other functions. See [Specifying the Type of a variable](#).

Although there are no restrictions to using defaulted parameters with types, their use is so much constrained by a type having exactly one parameter that they are of little practical help there.

You cannot use a type to perform any adjustment to the value being checked, if only because this value may be the temporary result of an expression, not an actual variable.

4.2.1.4 variables

These may be assigned values during execution e.g.

```
-- x may only be assigned integer values
integer x
x = 25

-- a, b and c may be assigned *any* value
object a, b, c
```

```
a = {}  
b = a  
c = 0
```

When you declare a variable you name the variable (which protects you against making spelling mistakes later on) and you define which sort of values may legally be assigned to the variable during execution of your program.

The simple act of declaring a variable does not assign any value to it. If you attempt to read it before assigning any value to it, EUPHORIA will issue a run-time error as "variable xyz has never been assigned a value".

To guard against forgetting to initialise a variable, and also because it may make the code clearer to read, you can combine declaration and assignment:

```
integer n = 5
```

This is equivalent to

```
integer n  
n = 5
```

It is not infrequent that one defines a private variable that bears the same name as one already in scope. You can reuse the value of that variable when performing an initialisation on declare by using a default namespace for the current [file](#):

```
namespace app  
  
integer n  
n=5  
  
procedure foo()  
    integer n = app:n + 2  
    ? n  
end procedure  
  
foo() -- prints out 7
```

4.2.1.5 constants

These are variables that are assigned an initial value that can never change e.g.

```
constant MAX = 100  
constant Upper = MAX - 10, Lower = 5  
constant name_list = {"Fred", "George", "Larry"}
```

The result of any expression can be assigned to a constant, even one involving calls to previously defined functions, but once the assignment is made, the value of the constant variable is "locked in".

Constants may not be declared inside a subroutine.

4.2.1.6 enum

An enumerated value is a special type of constant where the first value defaults to the number 1 and each item after that is incremented by 1.

```
enum ONE, TWO, THREE, FOUR

-- ONE is 1, TWO is 2, THREE is 3, FOUR is 4
```

You can change the value of any one item by assigning it a numeric value. Enums can only take numeric values. You cannot set the starting value to an expression or other variable. Subsequent values are always the previous value plus one, unless they too are assigned a default value.

```
enum ONE, TWO, THREE, ABC=10, DEF, XYZ

-- ONE is 1, TWO is 2, THREE is 3
-- ABC is 10, DEF is 11, XYZ is 12
```

EUPHORIA sequences use integer indexes, but with `enum` you may write code like this:

```
enum X, Y
sequence point = { 0,0 }
point[X] = 3
point[Y] = 4
```

4.2.2 Specifying the type of a variable

So far you've already seen some examples of variable types but now we will define types more precisely.

Variable declarations have a type name followed by a list of the variables being declared. For example,

```
object a

global integer x, y, z

procedure fred(sequence q, sequence r)
```

The types: **object**, **sequence**, **atom** and **integer** are **predefined**. Variables of type **object** may take on *any* value. Those declared with type **sequence** must always be sequences. Those declared with type **atom** must always be atoms.

Variables declared with type **integer** must be atoms with integer values from -1073741824 to +1073741823 inclusive. You can perform exact calculations on larger integer values, up to about 15 decimal digits, but declare them as **atom**, rather than integer.

Note:

In a procedure or function parameter list like the one for `fred()` above, a type name may only be followed by a single parameter name.

Performance Note:

Calculations using variables declared as integer will usually be somewhat faster than calculations involving variables declared as atom. If your machine has floating-point hardware, EUPHORIA will use it to manipulate atoms that are not integers. If your machine doesn't have floating-point hardware (this may happen on old 386 or 486 PCs), EUPHORIA will call software floating-point arithmetic routines contained in **euid.exe** (or in Windows). You can force eui.exe to bypass any floating-point hardware, by setting an environment variable:

```
SET NO87=1
```

The slower software routines will be used, but this could be of some advantage if you are worried about the floating-point bug in some early Pentium chips.

:udt To augment the [predefined types](#), you can create **user-defined types**. All you have to do is define a single-parameter function, but declare it with **type ... end type** instead of **function ... end function**. For example,

```
type hour(integer x)
    return x >= 0 and x <= 23
end type

hour h1, h2

h1 = 10      -- ok
h2 = 25      -- error! program aborts with a message
```

Variables h1 and h2 can only be assigned integer values in the range 0 to 23 inclusive. After each assignment to h1 or h2 the interpreter will call hour(), passing the new value. The value will first be checked to see if it is an integer (because of "integer x"). If it is, the return statement will be executed to test the value of x (i.e. the new value of h1 or h2). If hour() returns true, execution continues normally. If hour() returns false then the program is aborted with a suitable diagnostic message.

"hour" can be used to declare subroutine parameters as well:

```
procedure set_time(hour h)
```

set_time() can only be called with a reasonable value for parameter h, otherwise the program will abort with a message.

A variable's type will be checked after each assignment to the variable (except where the compiler can predetermine that a check will not be necessary), and the program will terminate immediately if the type function returns false. Subroutine parameter types are checked each time that the subroutine is called. This checking guarantees that a variable can never have a value that does not belong to the type of that variable.

Unlike other languages, the type of a variable does not affect any calculations on the variable, nor the way its contents are displayed. Only the value of the variable matters in an expression. The type just serves as an error check to prevent any "corruption" of the variable. User-defined types can catch unexpected logical errors in your program. They are not designed to catch or correct user input errors. In particular, they cannot adjust a wrong value to some other, presumably legal, one.

[[:type_check]] Type checking can be turned off or on between subroutines using the with `type_check` or

`without type_check` (see [[:special statements]]. It is initially on by default.

Note to Bench markers:

When comparing the speed of EUPHORIA programs against programs written in other languages, you should specify **without type_check** at the top of the file. This gives EUPHORIA permission to skip run-time type checks, thereby saving some execution time. All other checks are still performed, e.g. subscript checking, uninitialized variable checking etc. Even when you turn off type checking, EUPHORIA reserves the right to make checks at strategic places, since this can actually allow it to run your program *faster* in many cases. So you may still get a type check failure even when you have turned off type checking. Whether type checking is on or off, you will never get a *machine-level* exception. **You will always get a meaningful message from EUPHORIA when something goes wrong.** (*This might not be the case when you [poke](#) directly into memory, or call routines written in C or machine code.*)

EUPHORIA's method of defining types is simpler than what you will find in other languages, yet EUPHORIA provides the programmer with *greater* flexibility in defining the legal values for a type of data. Any algorithm can be used to include or exclude values. You can even declare a variable to be of type object which will allow it to take on *any* value. Routines can be written to work with very specific types, or very general types.

For many programs, there is little advantage in defining new types, and you may wish to stick with the four [predefined types](#). Unlike other languages, EUPHORIA's type mechanism is optional. You don't need it to create a program.

However, for larger programs, strict type definitions can aid the process of debugging. Logic errors are caught close to their source and are not allowed to propagate in subtle ways through the rest of the program.

Furthermore, it is easier to reason about the misbehavior of a section of code when you are guaranteed that the variables involved always had a legal value, if not the desired value.

Types also provide meaningful, machine-checkable documentation about your program, making it easier for you or others to understand your code at a later date. Combined with the subscript checking, uninitialized variable checking, and other checking that is always present, strict run-time type checking makes debugging much easier in EUPHORIA than in most other languages. It also increases the reliability of the final program since many latent bugs that would have survived the testing phase in other languages will have been caught by EUPHORIA.

Anecdote 1:

In porting a large C program to EUPHORIA, a number of latent bugs were discovered. Although this C program was believed to be totally "correct", we found: a situation where an uninitialized variable was being read; a place where element number "-1" of an array was routinely written and read; and a situation where something was written just off the screen. These problems resulted in errors that weren't easily visible to a casual observer, so they had survived testing of the C code.

Anecdote 2:

The Quick Sort algorithm presented on page 117 of *Writing Efficient Programs* by Jon Bentley has a subscript error! The algorithm will sometimes read the element just *before* the beginning of the array to be sorted, and will sometimes read the element just *after* the end of the array. Whatever garbage is read, the algorithm will still work - this is probably why the bug was never caught. But what if there isn't any (virtual) memory just before or just after the array? Bentley later modifies the algorithm such that this bug goes away--but he presented this version as being correct. *Even the experts need*

subscript checking!

Performance Note:

When typical user-defined types are used extensively, type checking adds only 20 to 40 percent to execution time. Leave it on unless you really need the extra speed. You might also consider turning it off for just a few heavily-executed routines. [Profiling](#) can help with this decision.

4.2.2.1 integer

An EUPHORIA `integer` is a mathematical integer restricted to the range -1,073,741,824 to +1,073,741,823. As a result, a variable of the integer type, while allowing computations as fast as possible, cannot hold 32-bit machine addresses, even though the latter are mathematical integers. You must use the `atom` type for this purpose. Also, even though the product of two integers is a mathematical integer, it may not fit into an integer, and should be kept in an atom instead.

4.2.2.2 atom

An `atom` can hold three kinds of data:

- Mathematical integers in the range `-power(2, 53)` to `+power(2, 53)`
- Floating point numbers, in the range `-power(2, 1024)+1` to `+power(2, 1024)-1`
- Large mathematical integers in the same range, but with a fuzz that grows with the magnitude of the integer.

`power(2, 53)` is slightly above 9.10^{15} , `power(2, 1024)` is in the 10^{308} range.

Because of these constraints, which arise in part from common hardware limitations, some care is needed for specific purposes:

- The sum or product of two integers is an `atom`, but may not be an `integer`.
- Memory addresses, or handles acquired from anything non EUPHORIA, including the operating system, **must** be stored as an `atom`.
- For large numbers, usual operations may yield strange results:

```
integer n = power(2, 27) -- ok
integer n_plus = n + 1, n_minus = n - 1 -- ok
atom a = n * n -- ok
atom a1 = n_plus * n_minus -- still ok
? a - a1 -- prints 0, should be 1 mathematically
```

This is not an EUPHORIA bug. The IEEE 754 standard for floating point numbers provides for 53 bits of precision for any real number, and an accurate computation of `a-a1` would require 54 of them. Intel FPU chips do have 64 bit precision registers, but the low order 16 bits are only used internally, and Intel recommends against using them for high precision arithmetic. Their SIMD machine instruction set only uses the IEEE 754 defined format.

4.2.2.3 sequence

A sequence is a type that is a *container*. A sequence has *elements* which can be accessed through their *index*, like in `my_sequence[3]`. sequences are so generic as being able to store all sorts of data structures: strings, trees, lists, anything. Accesses to sequences are always bound checked, so that you cannot read or write an element that doesn't exist, ever. A large amount of extraction and shape change operations on sequences is available, both as built-in operations and library routines. The elements of a sequence can have any type.

sequences are implemented very efficiently. Programmers used to pointers will soon notice that they can get most usual pointer operations done using sequence indexes. The loss in efficiency is usually hard to notice, and the gain in code safety and bug prevention far outweighs it.

4.2.2.4 object

This type can hold any data EUPHORIA can handle, both atoms and sequences.

The `object()` type returns 0 if a variable is not initialised, else 1.

4.2.3 Scope

4.2.3.1 Why scopes, and what are they?

The *scope* of an identifier is the portion of the program where that its declaration is in effect, i.e. where that identifier is *visible*.

EUPHORIA has many pre-defined procedures, functions and types. These are defined automatically at the start of any program. The EUPHORIA editor shows them in magenta. These pre-defined names are not reserved. You can override them with your own variables or routines.

Every user-defined identifier must be declared before it is used. You can read a EUPHORIA program from beginning to end without encountering any user-defined variables or routines that haven't been defined yet.

It is possible to call a routine that comes later in the source, but you must use the special functions, `routine_id()`, and either `call_func()` or `call_proc()` to do it. For example, procedures, functions and types can call themselves or one another *recursively*. Mutual recursion, where routine A calls routine B which directly or indirectly calls routine A, implies one of A or B being called before it is defined. This is perhaps the most frequent situation which requires using the `routine_id()` mechanism. See [Indirect Routine Calling](#) for more details.

4.2.3.2 Defining the scope of an identifier

The scope of an identifier is a description of what code can 'access' it. Code in the same scope of an identifier can access that identifier and code not in the same scope cannot access it.

The scope of a **variable** depends upon where and how it is declared.

- If it is declared within a **for**, **while**, **loop** or **switch**, its scope starts at the declaration and ends at the respective **end** statement.
- In an **if** statement, the scope starts at the declaration and ends either at the next **else**, **elsif** or **end if** statement.
- If a variable is declared within a routine (known as a private variable) and outside one of the structures listed above, the scope of the variable starts at the declaration and ends at the routine's **end** statement.
- If a variable is declared outside of a routine (known as a module variable), and does not have a scope modifier, its scope starts at the declaration and ends at the end of the file it is declared in.

The scope of a **constant** that does not have a scope modifier, starts at the declaration and ends at the end of the file it is declared in.

The scope of a **enum** that does not have a scope modifier, starts at the declaration and ends at the end of the file it is declared in.

The scope of all **procedures**, **functions** and **types**, which do not have a scope modifier, starts at the beginning of the source file and ends at the end of the source file in which they are declared. In other words, these can be accessed by any code in the same file.

Constants, enums, module variables, procedures, functions and types, which do not have a scope modifier are referred to as **local**. However, these identifiers can have a scope modifier preceding their declaration, which causes their scope to extend beyond the file they are declared in.

- If the keyword **global** precedes the declaration, the scope of these identifiers extends to the whole application. They can be accessed by code anywhere in the application files.
- If the keyword **public** precedes the declaration, the scope extends to any file that explicitly includes the file in which the identifier is declared, or to any file that includes a file that in turn `public includes` the file containing the `public` declaration.
- If the keyword **export** precedes the declaration, the scope only extends to any file that directly includes the file in which the identifier is declared.

When you **include** a EUPHORIA file in another file, only the identifiers declared using a scope modifier are accessible to the file doing the include. The other declarations in the included file are invisible to the file doing the include, and you will get an error message, "Errors resolving the following references", if you try to use them.

There is a variant of the **include** statement, called **public include**, which will be discussed later and behaves differently on **public** symbols.

Note that **constant** and **enum** declarations must be outside of any subroutine.

EUPHORIA encourages you to restrict the scope of identifiers. If all identifiers were automatically global to the whole program, you might have a lot of naming conflicts, especially in a large program consisting of files written by many different programmers. A naming conflict might cause a compiler error message, or it could lead to a very subtle bug, where different parts of a program accidentally modify the same variable without being aware of it. Try to use the most restrictive scope that you can. Make variables **private** to one routine

where possible, and where that isn't possible, make them **local** to a file, rather than **global** to the whole program. And whenever an identifier needs to be known from a few files only, make it **public** or **export** so as to hide it from whoever doesn't need to see it - and might some day define the same identifier.

For example:

```
-- sublib.e
export procedure bar()
?0
end procedure

-- some_lib.e
include sublib.e
export procedure foo()
?1
end procedure
bar() -- ok, declared in sublib.e

-- my_app.exw
include some_lib.e
foo() -- ok, declared in some_lib.e
bar() -- error! bar() is not declared here
```

Why not declare `foo()` as **global**, as it is meant to be used anywhere? Well, one could, but this will increase the risks of name conflicts. This is why, for instance, all public identifiers from the standard library have **public** scope. **global** should be used rarely, if ever. Because earlier versions of EUPHORIA didn't have **public** or **export**, it has to remain there for a while. One should be very sure of not polluting any foreign file's symbol table before using **global** scope--which is what a large part of the EUPHORIA source files do. Built-in identifiers act as if declared as **global**--but they are not declared in any EUPHORIA coded file.

4.2.3.3 Using namespaces

Identifiers marked as **global**, **public** or **export** are known as *exposed* variables because they can be used in files other than the one they were declared in. All other identifiers can only be used within their own file. This information is helpful when maintaining or enhancing the file, or when learning how to use the file. You can make changes to the internal routines and variables, without having to examine other files, or notify other users of the include file.

Sometimes, when using include files developed by others, you will encounter a naming conflict. One of the include file authors has used the same name for a exposed identifier as one of the other authors. One of fixing this, if you have the source, you could simply edit one of the include files to correct the problem, however then you'd have repeat this process whenever a new version of the include file was released.

EUPHORIA has a simpler way to solve this. Using an extension to the include statement, you can say for example:

```
include johns_file.e as john
include bills_file.e as bill

john:x += 1
bill:x += 2
```

In this case, the variable `x` was declared in two different files, and you want to refer to both variables in your file. Using the *namespace identifier* of either `john` or `bill`, you can attach a prefix to `x` to indicate which `x` you are referring to. We sometimes say that `john` refers to one *namespace*, while `bill` refers to another distinct *namespace*. You can attach a namespace identifier to any user-defined variable, constant, procedure or function. You can do it to solve a conflict, or simply to make things clearer. A namespace identifier has local scope. It is known only within the file that declares it, i.e. the file that contains the include statement. Different files might define different namespace identifiers to refer to the same included file.

There is a special, reserved namespace, **eu** for referring to built-in euphoria routines. This can be useful when a built-in routine has been overridden:

```
procedure puts( integer fn, object text )
    eu:puts(fn, "Overloaded puts says: "& text )
end procedure

puts(1, "Hello, world!\n")
eu:puts(1, "Hello, world!\n")
```

Files can also declare a default namespace to be used with the file. When a file with a default namespace is included, if the include statement did not specify a namespace, then the default namespace will be automatically declared in that file. If the include statement declares a namespace for the newly included file, then the specified namespace will be available instead of the default. No two files can use the same namespace identifier. If two files with the same default namespaces are included, at least one will be required to have a different namespace to be specified.

To declare a default namespace in a file, the first token (whitespace and comments are ignored) should be 'namespace' followed by the desired name:

```
-- foo.e : this file does some stuff
namespace foo
```

A namespace that is declared as part of an `include` statement is local to the file where the `include` statement is. A default namespace declared in a file is considered a public symbol in that file. Namespaces and other symbols (e.g., variables, functions, procedures and types) can have the same name without conflict. A namespace declared through an `include` statement will mask a default namespace declared in another file, just like a normal local variable will mask a public variable in another file. In this case, rather than using the default namespace, declare a new namespace through the `include` statement.

Note that declaring a namespace, either through the include statement or as a default namespace does not **require** that every symbol reference must be qualified with that namespace. The namespace simply **allows** the user to deconflict symbols in different files with the same name, or to allow the programmer to be explicit about where symbols are coming from for the purposes of clarity, or to avoid possible future conflicts.

A qualified reference does not absolutely restrict the reference to symbols that actually reside within the specified file. It can also apply to symbols included by that file. This is especially useful for multi-file libraries. Programmers can use a single namespace for the library, even though some of the visible symbols in that library are not declared in the main [file](#):

```
-- lib.e
namespace lib
```

```
public include sublib.e

public procedure main()
...

-- sublib.e
public procedure sub()
...

-- app.ex
include lib.e

lib:main()
lib:sub()
```

4.2.3.4 The visibility of public and export identifiers

When a file needs to see the public or exported identifiers in another file that includes the first file, the first file must include that other (including) file.

For example,

```
-- Parent file: foo.e --
public integer Foo = 1
include bar.e -- bar.e needs to see Foo
showit() -- execute a routine in bar.e

-- Included file: bar.e --
include foo.e -- included so I can see Foo
constant xyz = Foo + 1

public procedure showit()
? xyz
end procedure
```

Public symbols can only be seen by the file that explicitly includes the file where those public symbols are declared.

For example,

```
-- Parent file: foo.e --
include bar.e
showit() -- execute a public routine in bar.e
```

If however, a file wants a third file to also see the symbols that it can, it needs to do a `public include`.

For example,

```
-- Parent file: foo.e --
public include bar.e
showit() -- execute a public routine in bar.e
```

```
public procedure fooer()
. . .
end procedure

-- Appl file: runner.ex --
include foo.e
showit() -- execute a public routine that foo.e can see in bar.e
fooer()  -- execute a public routine in foo.e
```

The `public include` facility is designed to make having a library composed of multiple files easy for an application to use. It allows the main library file to expose symbols in files that *it* includes as if the application had actually included them. That way, symbols meant for the end user can be declared in files other than the main file, and the library can still be organized however the author prefers without affecting the end user.

Another example

Given that we have two files LIBA.e and LIBB.e ...

```
-- LIBA.e --
public constant
    foo1 = 1,
    foo2 = 2

export function foobarr1()
    return 0
end function

export function foobarr2()
    return 0
end function
```

and

```
-- LIBB.e --
-- I want to pass on just the constants not
-- the functions from LIBA.e.
public include LIBA.e
```

The `export` scope modifier is used to limit the extent that symbols can be accessed. It works just like `public` except that `export` symbols are only ever passed up one level only. In other words, if a file wants to use an `export` symbol, that file must include it explicitly.

In this example above, code in LIBB.e can see both the `public` and `export` symbols declared in LIBA.e (`foo1`, `foo2`, `foobarr1` and `foobarr2`) because it explicitly includes LIBA.e. And by using the `public` prefix on the `include` of LIBA.e, it also allows any file that includes LIBB.e to see the `public` symbols from LIBA.e but they will not see any `export` symbols declared in LIBA.e.

In short, a `public include` is used to expose `public` symbols that are included, up one level but not any `export` symbols that were included.

4.2.3.5 The complete set of resolution rules

Resolution is the process by which the interpreter determines which specific symbol will actually be used at any given point in the code. This is usually quite easy as most symbol names in a given scope are unique and so EUPHORIA doesn't have to choose between them. However, when the same symbol name is used in different but enclosing scopes, EUPHORIA has to make a decision about which symbol the coder is referring to.

When EUPHORIA sees an identifier name being used, it looks for the name's declaration starting from the current scope and moving outwards through the enclosing scopes until the name's declaration is found.

The hierarchy of scopes can be viewed like this ...

```
global/public/export
  file
    routine
      block 1
        block 2
          ...
        block n
```

So, if a name is used at a `block` level, EUPHORIA will first check for its declaration in the same block, and if not found will check the enclosing blocks until it reaches the routine level, in which case it checks the routine (including parameter names), and then check the file that the block is declared in and finally check the `global/public/export` symbols.

By the way, EUPHORIA will not allow a name to be declared if it is already declared in the same scope, or enclosing `block` or enclosing `routine`. Thus the following examples are illegal...

```
integer a
if x then
  integer a -- redefinition not allowed.
end if

if x then
  integer a
  if y then
    integer a -- redefinition not allowed.
  end if
end if

procedure foo(integer a)
if x then
  integer a -- redefinition not allowed.
end if
end procedure
```

But note that this below is valid ...

```
integer a = 1
procedure foo()
  integer a = 2
  ? a
```

```
end procedure
? a
```

In this situation, the second declaration of 'a' is said to *shadow* the first one. The output from this example will be ...

```
2
1
```

Symbols all declared in the same file (be they in blocks, routines or at the file level) are easy to check by EUPHORIA for scope clashes. However, a problem can arise when symbol names declared as global/public/export in different files are placed in the same scope during `include` processing. As it is quite possible for these files to come from independent developers that are not aware of each other's symbol names, the potential for name clashes is high. A name clash is just when the same name is declared in the same scope but in different files. EUPHORIA cannot generally decide which name you were referring to when this happens, so it needs you help to resolve it. This is where the `namespace` concept is used.

A namespace is just a name that you assign to an include file so that your code can exactly specify where an identifier that your code is using actually comes from. Using a namespace with an identifier, for example ...

```
include somefile.e as my_lib
include another.e
my_lib:foo()
```

enables EUPHORIA to resolve the identifier (`foo`) as explicitly coming from the file associated with the namespace "my_lib". This means that if `foo` was also declared as global/public/export in *another.e* then that `foo` would be ignored and the `foo` in *somefile.e* would be used instead. Without that namespace, EUPHORIA would have complained (Errors resolving the following references:)

If you need to use both `foo` symbols you can still do that by using two different namespaces. For example ...

```
include somefile.e as my_lib
include another.e as her_ns
my_lib:foo() -- Calls the one in somefile.e
her_ns:foo() -- Calls the one in another.e
```

Note that there is a reserved namespace name that is always in use. The special namespace **eu** is used to let EUPHORIA know that you are accessing a built-in symbol rather than one of the same name declared in someone's file.

For example...

```
include somefile.e as my_lib
result = my_lib:find(something) -- Calls the 'find' in somefile.e
xy = eu:find(X, Y) -- Calls EUPHORIA's built-in 'find'
```

The controlling variable used in a **for statement** is special. It is automatically declared at the beginning of the loop block, and its scope ends at the end of the for-loop. If the loop is inside a function or procedure, the loop variable cannot have the same name as any other variable declared in the routine or enclosing block. When the loop is at the top level, outside of any routine, the loop variable cannot have the same name as any other file-scoped variable. You can use the same name in many different for-loops as long as the loops aren't nested.

You do not declare loop variables as you would other variables because they are automatically declared as atoms. The range of values specified in the for statement defines the legal values of the loop variable.

Variables declared inside other types of blocks, such as a **loop**, **while**, **if** or **switch** statement use the same scoping rules as a for-loop index.

4.2.3.6 The override qualifier

There are times when it is necessary to replace a global, public or export identifier. Typically, one would do this to extend the capabilities of a routine. Or perhaps to supersede the user defined type of some public, export or global variable, since the type itself may not be global.

This can be achieved by declaring the identifier as **override**:

```
override procedure puts(integer channel, sequence text)
    eu:puts(log_file, text)
    eu:puts(channel, text)
end procedure
```

A warning will be issued when you do this, because it can be very confusing, and would probably break code, for the new routine to change the behaviour of the former routine. Code that was calling the former routine expects no difference in service, so there shouldn't be any.

If an identifier is declared global, public or export, but not override, and there is a built-in of the same name, EUPHORIA will not assume an override, and will choose the built-in. A warning will be generated whenever this happens.

4.3 Assignment statement

An **assignment statement** assigns the value of an expression to a simple variable, or to a subscript or slice of a variable. e.g.

```
x = a + b
y[i] = y[i] + 1
y[i..j] = {1, 2, 3}
```

The previous value of the variable, or element(s) of the subscripted or sliced variable are discarded. For example, suppose x was a 1000-element sequence that we had initialized with:

```
object x

x = repeat(0, 1000)  -- a sequence of 1000 zeros
```

and then later we assigned an atom to x with:

```
x = 7
```

This is perfectly legal since `x` is declared as an **object**. The previous value of `x`, namely the 1000-element sequence, would simply disappear. Actually, the space consumed by the 1000-element sequence will be automatically recycled due to EUPHORIA's dynamic storage allocation.

Note that the equals symbol '=' is used for both assignment and for a equality testing. There is never any confusion because an assignment in EUPHORIA is a statement only, it can't be used as an expression (as in C).

4.3.1 Assignment with Operator

EUPHORIA also provides some additional forms of the assignment statement.

To save typing, and to make your code a bit neater, you can combine assignment with one of the operators:

`+ - / * &`

For example, instead of saying:

```
mylongvarname = mylongvarname + 1
```

You can say:

```
mylongvarname += 1
```

Instead of saying:

```
galaxy[q_row][q_col][q_size] = galaxy[q_row][q_col][q_size] * 10
```

You can say:

```
galaxy[q_row][q_col][q_size] *= 10
```

and instead of saying:

```
accounts[start..finish] = accounts[start..finish] / 10
```

You can say:

```
accounts[start..finish] /= 10
```

In general, whenever you have an assignment of the form:

```
left-hand-side = left-hand-side op expression
```

You can say:

```
left-hand-side op= expression
```

where **op** is one of:

+ - * / &

When the left-hand-side contains multiple subscripts/slices, the `op=` form will usually execute faster than the longer form. When you get used to it, you may find the `op=` form to be slightly more readable than the long form, since you don't have to visually compare the left-hand-side against the copy of itself on the right side.

You cannot use assignment with operators while declaring a variable, because that variable is not initialised when you perform the assignment.

4.4 Branching Statements

4.4.1 if statement

An **if statement** tests a condition to see whether it is true or false, and then depending on the result of that test, executes the appropriate set of statements.

The syntax of `if` is

```
IFSTMT ==: IFTEST [ ELSIF ... ] [ELSE] ENDIF
IFTEST ==: if ATOMEXPR [ LABEL ] then [ STMTBLOCK ]
ELSIF  ==: elsif ATOMEXPR then [ STMTBLOCK ]
ELSE   ==: else [ STMTBLOCK ]
ENDIF  ==: end if
```

Description of syntax

- An *if statement* consists of the keyword **if**, followed by an *expression* that evaluates to an atom, optionally followed by a *label* clause, followed by the keyword **then**. Next is a set of zero or more statements. This is followed by zero or more *elsif* clauses. Next is an optional *else* clause and finally there is the keyword **end** followed by the keyword **if**.
- An *elsif* clause consists of the key word **elsif**, followed by an *expression* that evaluates to an atom, followed by the keyword **then**. Next is a set of zero or more statements.
- An *else* clause consists of the keyword **else** followed by a set of zero or more statements.

In EUPHORIA, *false* is represented by an atom whose value is zero and *true* is represented by an atom that has any non-zero value.

- When an *expression* being tested is true, EUPHORIA executes the statements immediately following the **then** keyword after the *expression*, up to the corresponding **elsif** or **else**, whichever comes next, then skips down to the corresponding **end if**.
- When an *expression* is false, EUPHORIA skips over any statements until it comes to the next corresponding **elsif** or **else**, whichever comes next. If this is an **elsif** then its *expression* is tested otherwise any statements following the **else** are executed.

For example:

```
if a < b then
    x = 1
```

```
end if

if a = 9 and find(0, s) then
    x = 4
    y = 5
else
    z = 8
end if

if char = 'a' then
    x = 1
elseif char = 'b' or char = 'B' then
    x = 2
elseif char = 'c' then
    x = 3
else
    x = -1
end if
```

Notice that **elseif** is a contraction of *else if*, but it's cleaner because it doesn't require an **end if** to go with it. There is just one **end if** for the entire *if statement*, even when there are many **elseif** clauses contained in it.

The **if** and **elseif** expressions are tested using [short-circuit](#) evaluation.

An *if statement* can have a *label clause* just before the first **then** keyword. See the section on [labelled headers](#). Note that an *elsif clause* can not have a label.

4.4.2 switch statement

The switch statement is used to run a specific set of statements, depending on the value of an expression. It often replaces a set of if-elsif statements due to it's ability to be highly optimized, thus much greater performance. There are some key differences, however. A switch statement operates upon the value of a single expression, and the program flow continues based upon defined cases. The syntax of a switch statement:

```
switch <expr> [with fallthru] [label "<label name>"] do
    case <val>[, <val2>, ...] then
        [code block]
        [[break [label]]|fallthru]
    case <val>[, <val2>, ...] then
        [code block]
        [[break [label]]|fallthru]
    case <val>[, <val2>, ...] then
        [code block]
        [[break [label]]|fallthru]
    ...

    [case else]
        [code block]
        [[break [label]]|fallthru]
end switch
```

The above example could be written with `if` statements like this ..

```
object temp = expression
object breaking = false
if equal(temp, val1) then
    [code block 1]
    [breaking = true]
end if
if not breaking and equal(temp, val2) then
    [code block 2]
    [breaking = true]
end if
if not breaking and equal(temp, val3) then
    [code block 3]
    [breaking = true]
end if
...
if not breaking then
    [code block 4]
    [breaking = true]
end if
```

The `<val>` in a case must be either an atom, literal string, constant or enum. Multiple values for a single case can be specified by separating the values by commas. By default, control flows to the end of the switch block when the next case is encountered. The default behavior can be modified in two ways. The default for a particular switch block can be changed so that control passes to the next executable statement whenever a new case is encountered by using `with fallthru` in the switch statement:

```
switch x with fallthru do
    case 1 then
        ? 1
    case 2 then
        ? 2
        break
    case else
        ? 0
end switch
```

Note that when `with fallthru` is used, the `break` statement can be used to jump out of the switch block. The behavior of individual cases can be changed by using the `fallthru` statement:

```
switch x do
    case 1 then
        ? 1
        fallthru
    case 2 then
        ? 2
    case else
        ? 0
end switch
```

Note that the `break` statement before `case else` was omitted, because the equivalent action is taken automatically by default.

```
switch length(x) do
```



```
case 1 then
    -- do something
    fallthru
case 2 then
    -- do something extra
case 3 then
    -- do something usual

case else
    -- do something else
end switch
```

The label "name" is optional and if used it gives a name to the switch block. This name can be used in nested switch break statements to break out of an enclosing switch rather than just the owning switch.

Example:

```
switch opt label "LBLa" do
    case 1, 5, 8 then
        FuncA()

    case 4, 2, 7 then
        FuncB()
        switch alt label "LBLb" do
            case "X" then
                FuncC()
                break "LBLa"

            case "Y" then
                FuncD()

            case else
                FuncE()
        end switch
        FuncF()

    case 3 then
        FuncG()
        break

    case else
        FuncH()
end switch
FuncM()
```

In the above, if opt is 2 and alt is "X" then it runs...

FuncB() FuncC() FuncM()

But if opt is 2 and alt is "Y" then it runs ...

FuncB() FuncD() FuncF() FuncG() FuncM()

In other words, the break "LBLa" skips to the end of the switch called "LBLa" rather than the switch called "LBLb".

4.4.3 ifdef statement

The `ifdef` statement has a similar syntax to the `if` statement.

```
ifdef SOME_WORD then
  --... zero or more statements
elsifdef SOME_OTHER_WORD then
  --... zero or more statments
elsdef
  --... zero or more statements
end ifdef
```

Of course, the `elsifdef` and `elsdef` clauses are optional, just like `elsif` and `else` are option in an `if` statement.

The major differences between and `if` and `ifdef` statement are that `ifdef` is executed at parse time not runtime, and `ifdef` can only test for the existance of a defined word whereas `if` can test any boolean expression.

Note that since the `ifdef` statement executes at parse time, run-time values cannot be checked, only words defined by the `-D` command line switch, or by the `with define` directive, or one of the special predefined words.

The purpose of `ifdef` is to allow you to change the way your program operates in a very efficient manner. Rather than testing for a specific condition repeatedly during the running of a program, `ifdef` tests for it once during parsing and then generates the precise IL code to handle the condition.

For example, assume you have some debugging code in your application that displays information to the screen. Normally you would not want to see this display so you set a condition so it only displays during a 'debug' session. The first example below shows how would could do this just using the `if` statement, and the second example shows the same thing but using the `ifdef` statement.

```
-- Example 1. --
if find("-DEBUG", command_line()) then
  writefln("Debug x=[], y=[]", {x,y})
end if
```

```
-- Example 1. --
ifdef DEBUG then
  writefln("Debug x=[], y=[]", {x,y})
end ifdef
```

As you can see, they are almost identical. However, in the first example, everytime the program gets to this point in the code, it tests the command line for the `-DEBUG` switch before deciding to display the information or not. But in the second example, the existance of `DEBUG` is tested *once* at parse time, and if it exists then, EUPHORIA generates the IL code to do the display. Thus when the program is running then everytime it gets to this point in the code, it does **not** check that `DEBUG` exists, instead it already knows it does so it just does the display. If however, `DEBUG` did not exist at parse time, then the IL code for the display would simply be omitted, meaning that during the running of the program, when it gets to this point in the code, it does not recheck for `DEBUG`, instead it already knows it doesn't exist and the IL code to do the display also doesn't exist so nothing is displayed. This can be a much needed performance boost for a program.

EUPHORIA predefines some words itself:

4.4.3.1 EUPHORIA Version Definitions

- **EU4** - Major EUPHORIA Version
- **EU400** - Major and Minor EUPHORIA Version
- **EU40000** - Major, Minor and Release EUPHORIA Version

EUPHORIA is release with the common version scheme of Major, Minor and Release version identifiers in the form of major.minor.release. When 4.0.1 is released, **EU40000** will no longer be defined, but **EU40001** will be defined. When 4.1 is release, **EU400** will no longer be defined, but **EU401** will be defined. Finally, when 5.0 is released, **EU4** will no longer be defined, but **EU5** will be defined.

4.4.3.2 Platform Definitions

- **WIN32_CONSOLE** - Platform is Windows and is being executed with the Console version of the interpreter (**eui.exe**)
- **WIN32_GUI** - Platform is Windows and is being executed with the GUI version of the interpreter (**euiw.exe**)
- **WIN32** - Platform is Windows (GUI or Console)
- **WINDOWS** - Platform is any platform that was derived from Windows
- **LINUX** - Platform is Linux
- **OSX** - Platform is Mac OS X
- **SUNOS** - Platform is SunOS (Open Solaris)
- **FREEBSD** - Platform is FreeBSD
- **OPENBSD** - Platform is OpenBSD
- **BSD** - Platform is a BSD variant (FreeBSD, OpenBSD, SunOS and OS X)
- **UNIX** - Platform is any Unix

4.4.3.3 Application Definitions

- **EUI** - Application is being interpreted by **eui**.
- **EUC** - Application is being translated by **euc**.
- **EUC_CON** - Application is being translated by **euc** as a *console* application.
- **EUC_DLL** - Application is being translated by **euc** into a *DLL* file.
- **EUB** - Application is being converted to a bound program by **eub**.
- **EUB_SHROUD** - Application is being converted to a shrouded program by **eub**.
- **EUB_CON** - Application is being converted to a bound *console* program by **eub**.
- **EUB_WIN32** - Application is being converted to a bound *MS-Windows* program by **eub**.

4.4.3.4 Library Definitions

- **DATA_EXECUTE** - Application will always get executable memory from `allocate()` even when the system has Data Execute Protection enabled for the EUPHORIA Interpreter.
- **SAFE** - Enables safe runtime checks for operations for routines found in `machine.e` and `dll.e`

- **UCSTYPE_DEBUG** - Found in include/std/ucstypes.e
- **CRASH** - Found in include/std/unittest.e

More examples

```
-- file: myproj.ex
puts(1, "Hello, I am ")
ifdef EUC then
    puts(1, "a translated")
end ifdef
ifdef EUI then
    puts(1, "an interpreted")
end ifdef
ifdef EUB then
    puts(1, "a bound")
end ifdef
ifdef EUB_SHROUD then
    puts(1, ", shrouded")
end ifdef
puts(1, " program.\n")
```

```
C:\myproj> eui myproj.ex
Hello, I am an interpreted program.
C:\myproj> euc -con myprog.ex
... translating ...
... compiling ...
C:\myproj> myprog.exe
Hello, I am a translated program.
C:\myproj> bind myprog.ex
...
C:\myproj> myprog.exe
Hello, I am a bound program.
C:\myproj> shroud myprog.ex
...
C:\myproj> eub myprog.il
Hello, I am a bound, shrouded program.
```

It is possible for one or more of the above definitions to be true at the same time. For instance, EUC and EUC_DLL will both be true when the source file has been translated to a DLL. If you wish to know if your source file is translated and not a DLL, then you can

```
ifdef EUC and not EUC_DLL then
    -- translated to an application
end ifdef
```

4.4.3.5 Using ifdef

You can define your own words either in source:

```
with define MY_WORD          -- defines
without define OTHER_WORD    -- undefines
```

or by command line:

```
eui -D MY_WORD myprog.ex
```

This can handle many tasks such as change the behavior of your application when running on *Linux* vs. *Windows*, enable or disable debug style code or possibly work differently in demo/shareware applications vs. registered applications.

You should surround code that is not portable with `ifdef` like:

```
ifdef WIN32 then
    -- Windows specific code.
elseif
    include std/error.e
    crash("This program must be run with the Windows interpreter.")
end ifdef
```

When writing **include files** that you cannot run on some platform, issue a crash call in the **include file**. Yet make sure that public constants and procedures are defined for the unsupported platform as well.

```
ifdef UNIX then
    include std/bash.e
end ifdef

-- define exported and public constants and procedures for
-- OSX as well
ifdef WIN32 or OSX then
    -- OSX is not supported but we define public symbols for it anyhow.
```

The reason for doing this is so that the user that includes your include file sees an "OS not supported" message instead of an "undefined reference" message.

Defined words must follow the same character set of an identifier, that is, it must start with either a letter or underscore and contain any mixture of letters, numbers and underscores. It is common for defined words to be in all upper case, however, it is not required.

A few examples:

```
for a = 1 to length(lines) do
    ifdef DEBUG then
        printf(1, "Line %i is %i characters long\n", {a, length(lines[a])})
    end ifdef
end for

sequence os_name
ifdef UNIX then
    include unix_ftp.e
elseifdef WIN32 then
    include win32_ftp.e
else
    crash("Operating system is not supported")
end ifdef

ifdef SHAREWARE then
    if record_count > 100 then
        message("Shareware version can only contain 100 records. Please register")
```



```
        abort(1)
    end if
end ifdef
```

The `ifdef` statement is very efficient in that it makes the decision only once during parse time and only emits the TRUE portions of code to the resulting interpreter. Thus, in loops that are iterated many times there is zero performance hit when making the decision. Example:

```
while 1 do
    ifdef DEBUG then
        puts(1, "Hello, I am a debug message\n")
    end ifdef
    -- more code
end while
```

If `DEBUG` is defined, then the interpreter/translator actually sees the code as being:

```
while 1 do
    puts(1, "Hello, I am a debug message\n")
    -- more code
end while
```

Now, if `DEBUG` is not defined, then the code the interpreter/translator sees is:

```
while 1 do
    -- more code
end while
```

Do be careful to put the numbers after the platform names for Windows:

```
-- This puts() routine will never be called
-- even when run by the Windows interpreter!
ifdef WIN then
    puts(1, "I am on Windows\n")
end ifdef
```

4.5 Loop statements

An iterative code block repeats its own execution zero, one or more times. There are several ways to specify for how long the process should go on, and how to stop or otherwise alter it. An iterative block may be informally called a loop, and each execution of code in a loop is called an iteration of the loop.

EUPHORIA has three flavours of loops. They all may harbor a [labelled header](#), in order to make exiting or resuming them more flexible.

4.5.1 while statement

A **while statement** tests a condition to see if it is non-zero (true), and while it is true a loop is executed.

Syntax Format:

```
while expr [with entry] [label "name" ] do
```

Example 1

```
while x > 0 do  
    a = a * 2  
    x = x - 1  
end while
```

Example 2

```
while sequence(Line) with entry do  
    proc(Line)  
entry  
    Line = gets(handle)  
end while
```

Example 3

```
while true label "main" do  
    res = funcA()  
    if res > 5 then  
        if funcB() > some_value then  
            continue "main" -- go to start of loop  
        end if  
        procC()  
    end if  
    procD(res)  
    for i = 1 to res do  
        if i > some_value then  
            exit "main" -- exit the "main" loop, not just this 'for' loop.  
        end if  
        procF(i,res)  
    end if  
  
    res = funcE(res, some_value)  
end while
```

4.5.2 loop until statement

A **loop** statement tests a condition to see if it is non-zero (true), and until it is true a loop is executed.

Syntax Format:

```
loop [with entry] [label "name" ] do
```

```
statements  
until expr  
end loop
```

```
loop do  
    a = a * 2  
    x = x - 1
```

```

    until x<=0
end loop

loop with entry do
    a = a * 2
    entry
    x = x - 1
    until x<=0
end loop

loop label "GONEXT" do
    a = a * 2
    y += 1
    if y = 7 then continue "GONEXT" end if
    x = x - 1
    until x<=0
end loop

```

This differs from a **while** loop in more than one way:

1. The body of the loop is executed at least once, since testing takes place AFTER it completes. In a **while**, the test is taken BEFORE execution.
2. The header contains only optional **label** and **entry** tags.

4.5.3 for statement

A **for** statement sets up a special loop that has its own **loop variable**. The **loop variable** starts with the specified initial value and increments or decrements it to the specified final value. The **for** statement is used when you need to repeat a set of statements a specific number of times.

Example:

```

-- Display the numbers 1 to 6 on the screen.
puts(1, "1\n")
puts(1, "2\n")
puts(1, "3\n")
puts(1, "4\n")
puts(1, "5\n")
puts(1, "6\n")

```

This block of code simply starts at the first line and runs each in turn. But it could be written more simply and flexibly by using a **for** statement.

```

for i = 1 to 6 do
    printf(1, "%d\n", i)
end for

```

Now it's just three lines of code rather than six. More importantly, if we needed to change the program to print the numbers from 1 to 100, we only have to change one line rather than add 94 new lines.

```

for i = 1 to 100 do -- One line change.
    printf(1, "%d\n", i)
end for

```

Or using another way ...

```
for i = 1 to 10 do
    ? i    -- ? is a short form for print()
end for

-- fractional numbers allowed too
for i = 10.0 to 20.5 by 0.3 do
    for j = 20 to 10 by -2 do    -- counting down
        ? {i, j}
    end for
end for
```

However, adding together floating point numbers that are not the ratio of an integer by a power of 2 -- *0.3 is not such a ratio*--leads to some "fuzz" in the value of the index. In some cases, you might get unexpected results because of this fuzz, which arises from a common hardware limitation. For instance, `floor(10*0.1)` is 1 as expected, but `floor(0.1+0.1+0.1+0.1+0.1+0.1+0.1+0.1+0.1+0.1)` is 0.

The **loop variable** is declared automatically and exists until the end of the loop. Outside of the loop the variable has no value and is not even declared. If you need its final value, copy it into another variable before leaving the loop. The compiler will not allow any assignments to a loop variable. The initial value, loop limit and increment must all be atoms. If no increment is specified then +1 is assumed. The limit and increment values are established only on entering the loop, and are not affected by anything that happens during the execution of the loop.

4.6 Flow control statements

Program execution flow refers to the order in which program statements are run in. By default, the next statement to run after the current one is the next statement *physically* located after the current one.

Example:

```
a = b + c
printf(1, "The result of adding %d and %d is %d", {b,c,a})
```

In that example, `b` is added to `c`, assigning the result to `a`, and then the information is displayed on the screen using the `printf` statement.

However, there are many times in which the order of execution needs to be different from the default order, to get the job done. EUPHORIA has a number of *flow control statements* that you can use to arrange the execution order of statements.

A set of statements that are run in their order of appearance is called a *block*. Blocks are good ways to organize code in easily identifiable chunks. However it can be desirable to leave a block before reaching the end, or slightly alter the default course of execution.

The following flow control keywords are available.

```
break retry entry exit continue return goto end
```

4.6.1 exit statement

Exiting a loop is done with the keyword **exit**. This causes flow to immediately leave the current loop and recommence with the first statement after the end of the loop.

```
for i = a to b do
  c = i
  if doSomething(i) = 0 then
    exit -- Stop executing code inside the 'for' block.
  end if
end for

-- Flow restarts here.
if c = a then ...
```

But sometimes you need to leave a block that encloses the current one. EUPHORIA has two methods available for you to do this. The safest method, in terms of future maintenance, is to name the block you want to exit from and use that name on the exit statement. The other method is to use a number on the exit statement that refers to the depth that you want to exit from.

A block's name is always a string literal and only a string literal. You cannot use a variable that contains the block's name on an exit statement. The name comes after the `label` keyword, just before the `do` keyword. Example:

```
integer b
b = 0
for i = 1 to 20 label "main" do
  for j = 1 to 20 do
    b += i + j
    ? {i, j, b}
    if b > 50 then
      b = 0
      exit "main"
    end if
  end for
end for
? b
```

The output from this is ...

```
{1, 1, 2}
{1, 2, 5}
{1, 3, 9}
{1, 4, 14}
{1, 5, 20}
{1, 6, 27}
{1, 7, 35}
{1, 8, 44}
{1, 9, 54}
0
```

The **exit "main"** causes execution flow to leave the **for** block named *main*.

The same thing could be achieved using the **exit N** format...

```
integer b
  b = 0
for i = 1 to 20 do
  for j = 1 to 20 do
    b += i + j
    ? {i, j, b}
    if b > 50 then
      b = 0
      exit 2 -- exit 2 levels of depth
    end if
  end for
end for

? b
```

But using this method means you have to take more care when changing the program so that if you change the depth, you also need to change the *exit* statement.

Note:

A special form of **exit N** is `exit 0`. This leaves all levels of loop, regardless of the depth. Control continues after the outermost loop block. Likewise, `exit -1` exits the second outermost loop, and so on.

For easier and safer program maintenance, the explicit label form is to be preferred. Other forms are variously sensitive to changes in the program organisation. Yet, they may prove more convenient in short, short lived programs, and are provided mostly for this purpose.

For information on how to associate a string to a block of code, see the section [labelled headers](#).

An **exit** without any label or number in a [while statement](#) or a [for statement](#) causes immediate termination of that loop, with control passing to the first statement after the loop.

Example:

```
for i = 1 to 100 do
  if a[i] = x then
    location = i
    exit
  end if
end for
```

It is also quite common to see something like this:

```
constant TRUE = 1

while TRUE do
  ...
  if some_condition then
    exit
  end if
  ...
end while
```

i.e. an "infinite" while-loop that actually terminates via an **exit statement** at some arbitrary point in the body of the loop.

Performance Note::

EUPHORIA optimizes this type of loop. At run-time, no test is performed at the top of the loop. There's just a simple unconditional jump from **end while** back to the first statement inside the loop.

4.6.2 break statement

Works exactly like the **exit statement**, but applies to **if statements** or **switch statements** rather than to loop statements of any kind. Example:

```
if s[1] = 'E' then
  a = 3
  if s[2] = 'u' then
    b = 1
    if s[3] = 'p' then
      break 0 -- leave topmost if block
    end if
    a = 2
  else
    b = 4
  end if
else
  a = 0
  b = 0
end if
```

This code results in:

- "Dur" -> a=0 b=0
- "Exe" -> a=3 b=4
- "Eux" -> a=2 b=1
- "Eup" -> a=3 b=1

The same optional parameters can be used with the **break** statement as with the **exit** statement, but of course apply to if and switch blocks only, instead of loops.

4.6.3 continue statement

Likewise, skipping the rest of an iteration in a single code block is done using a single keyword, **continue**. The **continue statement** continues execution of the loop it applies to by going to the next iteration now. Going to the next iteration means testing a condition (for **while** and **loop** constructs, or changing the **for** construct variable index and checking whether it is still within bounds.

```
for i = 3 to 6 do
  ? i
  if i = 4 then
    puts(1, "(2)\n")
    continue
  end if
  ? i * i
end for
```

This will print 3, 9, 4, (2), 5 25, 6 36.

```
integer b
b = 0
for i = 1 to 20 label "main" do
  for j = 1 to 20 do
    b += i + j
    if b > 50 then
      printf(1, "%d ", b)
      b = 0
      continue "main"
    end if
  end for
end for

? b
```

The same optional parameters that can be used in an **exit** statement can apply to a **continue** statement.

4.6.4 retry statement

The **retry statement** retries executing the current iteration of the loop it applies to. The statement branches to the first statement of the designated loop, without testing anything nor incrementing the for loop index.

Normally, a sub-block which contains a **retry statement** also contains another flow control keyword, since otherwise the iteration would be endlessly executed.

```
errors = 0
for i = 1 to length(files_to_open) do
  fh = open(files_to_open[i], "rb")
  if fh=-1 then
    if errors > 5 then
      exit
    else
      errors += 1
      retry
    end if
  end if
  file_handles[i] = fh
end for
```

Since **retry** does not change the value of *i* and tries again opening the same file, there has to be a way to break from the loop, which the **exit statement** provides.

The same optional parameters that can be used in an **exit** statement can apply to a **retry** statement.

4.6.5 with entry statement

It is often the case that the first iteration of a loop is somehow special. Some things have to be done before the loop starts--they are done before the statement starting the loop. Now, the problem is that, just as often, some things do not need to, or should not, be done at this initialisation stage. The **entry keyword** is an alternative to

setting flags relentlessly and forgetting to update them. Just add the **entry** keyword at the point you wish the first iteration starts.

```
public function find_all(object x, sequence source, integer from)
    sequence ret = {}

    while from > 0 with entry do
        ret &= from
        from += 1
    entry
        from = find_from(x, source, from)
    end while

    return ret
end function
```

Instead of performing an initial test, which may crash because `from` has not been assigned a value yet, the first iteration jumps at the point where `from` is being computed. The following iterations are normal. To emphasize the fact that the first iteration is not normal, the `entry` clause must be added to the loop header, after the condition.

The `entry` statement is not supported for `for` loops, because they have a more rigid nature structure than `while` or `loop` constructs.

Note on infinite loops.

With **eui.exe** or **eui**, control-c will always stop your program immediately, but with the **euiw.exe** that has not produced any console output, you will have to use the Windows process monitor to end the application.

4.6.6 goto statement

`goto` instructs the computer to resume code execution at a place which does not follow the statement. The place to resume execution is called the *target* of the statement. It is restricted to lie in the current routine, or the current file if outside any routine.

Syntax is:

```
goto "label string"
```

The target of a `goto` statement can be any accessible `label` statement:

```
label "label string"
```

Label names must be double quoted constant strings. Characters that would be illegal in an EUPHORIA identifier may appear in a label name, since it is a regular string.

Labelled headers do not count as possible `goto` targets.

Use `goto` in production code when all the following applies:

- you want to proceed with a statement which is not the following one;
- the various structured constructs wouldn't do, or very awkwardly;
- you contemplate a significant gain in speed/reliability from such a direct move;
- the code flow remains understandable for an outsider nevertheless.

During early development, it may be nice to have while the code isn't firmly structured. But most instances of `goto` should melt into structured constructs as soon as possible as code matures. You may find out that modifying a program that has `goto` statements is usually trickier than if it had not had them.

The following may be situations where `goto` can help:

- A routine has several return statements, and some processing must be done before returning, no matter from where. It may be clearer to `goto` a single return point and perform the processing only at this point.
- An exit statement in a loop corresponds to an early exit, and the normal processing that immediately follows the loop is not relevant. Replacing an exit statement followed by various flag testing by a single `goto` can help.

Explicit label names will tremendously help maintenance. Remember that there is no limit to their contents.

`goto`-ing into a scope (like an `if` block, a `for` loop,...) will just do that. Some variables may be defined only in that scope, and they may or may not have sensible values. It is up to the programmer to take appropriate action in this respect.

4.6.7 Header Labels

As shown in the above section on control flow statements, most can have their own label. To label a flow control statement, use a `label` clause immediately preceding the flow control's terminator keyword (`then / do`).

A `label` clause consists of the keyword **label** followed by a string literal. The string is the label name.

Examples:

```
if n=0 label "an_if_block" then
    ...
end if

while TRUE label "a_while_block" do
    ...
end while

loop label "a_loop_block" do
    ...
until TRUE

switch x label "a_switch_block" do
    ...
end switch
```

Note: If a flow control statement has both an `entry` clause and a `label` clause, the `entry` clause must come before the `label` clause:

```
while 1 label "top" with entry do -- WRONG
```

```
while 1 with entry label "top" do -- CORRECT
```

4.7 Short-Circuit Evaluation

When the condition tested by `if`, `elsif`, `until`, or `while` contains `and` or `or` operators, [short-circuit](#) evaluation will be used. For example,

```
if a < 0 and b > 0 then ...
```

If `a < 0` is false, then EUPHORIA will not bother to test if `b` is greater than 0. It will know that the overall result is false regardless. Similarly,

```
if a < 0 or b > 0 then ...
```

if `a < 0` is true, then EUPHORIA will immediately decide that the result true, without testing the value of `b`, since the result of this test would be irrelevant.

In general, whenever we have a condition of the form:

```
A and B
```

where `A` and `B` can be any two expressions, EUPHORIA will take a short-cut when `A` is false and immediately make the overall result false, without even looking at expression `B`.

Similarly, with:

```
A or B
```

when `A` is true, EUPHORIA will skip the evaluation of expression `B`, and declare the result to be true.

If the expression `B` contains a call to a function, and that function has possible **side-effects**, i.e. it might do more than just return a value, you will get a compile-time warning. Older versions (pre-2.1) of EUPHORIA did not use [short-circuit](#) evaluation, and it's possible that some old code will no longer work correctly, although a search of the EUPHORIA archives did not turn up any programs that depend on side-effects in this way, but other EUPHORIA code might do.

The expression, `B`, could contain something that would normally cause a run-time error. If EUPHORIA skips the evaluation of `B`, the error will not be discovered. For instance:

```
if x != 0 and 1/x > 10 then -- divide by zero error avoided
```

```
while 1 or {1,2,3,4,5} do -- illegal sequence result avoided
```

`B` could even contain uninitialized variables, out-of-bounds subscripts etc.

This may look like sloppy coding, but in fact it often allows you to write something in a simpler and more readable way. For instance:

```
if atom(x) or length(x)=1 then
```

Without short-circuiting, you would have a problem when *x* was an atom, since *length* is not defined for atoms. With short-circuiting, *length(x)* will only be checked when *x* is a sequence. Similarly:

```
-- find 'a' or 'A' in s
i = 1
while i <= length(s) and s[i] != 'a' and s[i] != 'A' do
    i += 1
end while
```

In this loop the variable *i* might eventually become greater than *length(s)*. Without short-circuit evaluation, a subscript out-of-bounds error will occur when *s[i]* is evaluated on the final iteration. With short-circuiting, the loop will terminate immediately when *i <= length(s)* becomes false. EUPHORIA will not evaluate *s[i] != 'a'* and will not evaluate *s[i] != 'A'*. No subscript error will occur.

Short-circuit evaluation of *and* and *or* takes place for *if*, *elsif*, *until* and *while* conditions only. It is not used in other contexts. For example, the assignment statement:

```
x = 1 or {1,2,3,4,5} -- x should be set to {1,1,1,1,1}
```

If short-circuiting were used here, we would set *x* to 1, and not even look at {1,2,3,4,5}. This would be wrong. Short-circuiting can be used in *if/elsif/until/while* conditions because we only care if the result is true or false, and conditions are required to produce an atom as a result.

4.8 Special Top-Level Statements

Before any of your statements are executed, the EUPHORIA front-end quickly reads your entire program. All statements are syntax checked and converted to a low-level intermediate language (IL). The interpreter immediately executes the IL after it is completely generated. The translator converts the IL to C. The binder/shrouder saves the IL on disk for later execution. These three tools all share the same front-end (written in EUPHORIA).

If your program contains only routine and variable declarations, but no top-level executable statements, then nothing will happen when you run it (other than syntax checking). You need a top-level statement to call your main routine (see [Example Program](#)). It's quite possible to have a program with nothing but top-level executable statements and no routines. For example you might want to use EUPHORIA as a simple calculator, typing just a few *print* or *?* statements into a file, and then executing it.

As we have seen, you can use any EUPHORIA statement, including *for statement*, *while statement*, *if statement*, etc... (but not *return*), at the top level i.e. *outside* of any *function* or *procedure*. In addition, the following special statements may *only* appear at the top level:

- *include*
- *with/without*

4.8.1 include

When you write a large program it is often helpful to break it up into logically separate files, by using **include statements**. Sometimes you will want to reuse some code that you have previously written, or that someone else has written. Rather than copy this code into your main program, you can use an **include statement** to refer to the file containing the code. The first form of the include statement is:

```
include filename
```

This reads in (compiles) a EUPHORIA source file.

Some Examples:

```
include std/graphics.e
include /mylib/myroutines.e
public include library.e
```

Any top-level code in the included file will be executed at start up time.

Any global identifiers that are declared in the file doing the including will also be visible in the file being included. However the situation is slightly different for an identifier declared as **public** or **export**. In these cases the file being included will **not** see public/export symbols declared in the file doing the including, unless the file being included also explicitly includes the file doing the including. Yes, you'd better read that again because its not that obvious. Here's an example...

We have two files, a.e and b.e ...

```
-- a.e --
? c -- declared as global in 'b.e'

-- b.e --
include a.e
global integer c = 0
```

This will work because being global the symbol 'c' in b.e can be seen by all files in this *include tree*.

However ...

```
-- a.e --
? c -- declared as public in 'b.e'

-- b.e --
include a.e
public integer c = 0
```

Will not work as public symbols can only be seen when their declaring file is explicitly included. So to get this to work you need to write a.e as ...

```
-- a.e --
include b.e
? c -- declared as public in 'b.e'
```

N.B. Only those symbols declared as `global` in the included file will be visible (accessible) in the remainder of the including file. Their visibility in other included files or in the main program file depends on other factors. Specifically, a global symbols can only be accessed by files in the same *include tree*. For example...

If we have `danny.e` declare a global symbol called 'foo', and `bob.e` includes `danny.e`, then code in `bob.e` can access `danny's 'foo'`. Now if we also have `cathy.e` declare a global symbol called 'foo', and `anne.e` includes `cathy.e`, then code in `anne.e` can access `cathy's 'foo'`. Nothing unusual about that situation. Now, if we have a program that includes both `bob.e` and `anne.e`, the code in `bob.e` and `anne.e` should still work even though there are now two global 'foo' symbols available. This is because the include tree for `bob.e` *only* contains `danny.e` and likewise the include tree for `anne.e` *only* contains `cathy.e`. So as the two 'foo' symbols are in separate include trees (from `bob.e` and `anne.e` perspective) code in those files continues to work correctly. A problem can occur if the main program (the one that includes both `bob.e` and `anne.e`) references 'foo'. In order for EUPHORIA to know which one the code author meant to use, the coder must use the namespace facility.

```
--- mainprog.ex ---
include anne.e as anne
include bob.e  as bob

anne:foo() -- Specify the 'foo' from anne.e.
```

If the above code did not use namespaces, EUPHORIA would not have know which 'foo' to use - the one from `bob.e` or the one in `anne.e`.

If `public` precedes the include statement, then all public identifiers from the included file will also be visible to the including file, and visible to any file that includes the current file.

If an absolute *filename* is given, EUPHORIA will open it and start parsing it. When a relative *filename* is given, EUPHORIA will try to open the file relative to the following directories, in the following order:

1. The directory containing the current source file. i.e. the source file that contains the include statement that is being processed.
2. The directory containing the main file given on the interpreter, translator or binder - see [command_line](#).
3. If you've defined an environment variable named `EUINC`, EUPHORIA will check each directory listed in `EUINC` (from left to right). `EUINC` should be a list of directories, separated by semicolons (colons on *Linux / FreeBSD*), similar in form to your `PATH` variable. `EUINC` can be added to your set of *Linux / FreeBSD* or *Windows* environment variables. (Via Control Panel / Performance & Maintenance / System / Advanced on *XP*, or `AUTOEXEC.BAT` on older versions of *Windows*). e.g. `SET EUINC=C:\EU\MYFILES;C:\EU\WIN32LIB` `EUINC` lets you organize your include files according to application areas, and avoid adding numerous unrelated files to `euphoria\include`.
4. Finally, if it still hasn't found the file, it will look in `euphoria\include`. This directory contains the standard EUPHORIA include files. The environment variable `EUDIR` tells EUPHORIA where to find your `euphoria` directory.

An included file can include other files. In fact, you can "nest" included files up to 30 levels deep.

Include file names typically end in `.e`, or sometimes `.ew` or `.eu` (when they are intended for use with *Windows* or *Unix*). This is just a convention. It is not required.

If your filename (or path) contains blanks, you must enclose it in double-quotes, otherwise quotes are optional. Note that under Windows, you can also use the forward slash '/' instead of the usually back-slash '\'. By doing this, the file paths are compatible with Unix systems and it means you don't have to 'escape' the back-slashes.

For example:

```
include "c:/program files/myfile.e"
```

Other than possibly defining a new namespace identifier (see below), an include statement will be quietly ignored if the same file has already been included.

An include statement must be written on a line by itself. Only a comment can appear after it on the same line.

The second form of the include statement is:

```
include filename as namespace_identifier:
```

This is just like the simple include, but it also defines a *namespace identifier* that can be attached to global identifiers in the included file that you want to refer to in the main file. This might be necessary to disambiguate references to those identifiers, or you might feel that it makes your code more readable. See [Using namespaces](#) for more. This namespace overrides the default namespace the file may define.

4.8.2 with / without

These special statements affect the way that EUPHORIA translates your program into internal form. Options to the `with` and `without` statement come in two flavors. One simply turns an option on or off, while the others have multiple states.

4.8.2.1 On/Off options

Default	Option
without	profile
without	profile_time
without	trace
without	batch
with	type_check
with	indirect_includes
with	inline

`with` turns **on** one of the options and `without` turns **off** one of the options.

For more information on the `profile`, `profile_time` and `trace` options, see [Debugging and Profiling](#). For more information on the `type_check` option, see [Performance Tips](#).

There is also a rarely-used special `with` option where a code number appears after `with`. In previous releases this code was used by RDS to make a file exempt from adding to the statement count in the old

"Public Domain" Edition. This is not used any longer, but does not cause an error.

You can select any combination of settings, and you can change the settings, but the changes must occur *between* subroutines, not within a subroutine. The only exception is that you can only turn on one type of profiling for a given run of your program.

An **included file** inherits the **with/without** settings in effect at the point where it is included. An included file can change these settings, but they will revert back to their original state at the end of the included file. For instance, an included file might turn off warnings for itself and (initially) for any files that it includes, but this will not turn off warnings for the main file.

`indirect_includes` This `with/without` option changes the way in which global symbols are resolved. Normally, the parser uses the way that files were included to resolve a usage of a global symbol. If `without indirect_includes` is in effect, then only direct includes are considered when resolving global symbols.

This option is especially useful when a program uses some code that was developed for a prior version of EUPHORIA that uses the pre-4.0 standard library, when all exposed symbols were global. These can often clash with symbols in the new standard library. Using `without indirect_includes` would not force a coder to use namespaces to resolve symbols that clashed with the new standard library.

Note that this setting does not propagate down to included files, unlike most `with/without` options. Each file begins with `indirect_includes` turned on.

`with_batch` Causes the program to not present the "Press Enter" prompt if an error occurs. The exit code will still be set to 1 on error. This is helpful for programs that run in a mode where no human may be directly interacting with it. For example, a CGI application or a CRON job.

You can also set this option via a [command line parameter](#).

4.8.2.2 Complex with/without options

4.8.2.2.1 with/without warning

Any warnings that are issued will appear on your screen after your program has finished execution. Warnings indicate minor problems. A warning will never terminate the execution of your program. You will simply have to hit the Enter key to keep going--which may stop the program on an unattended computer.

The forms available are ...

```
with warning
    enables all warnings
```

```
without warning
    disables all warnings
```

```
with warning (warning name list)
with warning = (warning name list)
```


enables only these warnings, and disables all other

without warning (*warning name list*)

without warning = (*warning name list*)

enables all warnings except the warnings listed

with warning &= (*warning name list*)

with warning += (*warning name list*)

enables listed warnings in addition to whichever are enabled already

##without warning &= (*warning name list*)

without warning += (*warning name list*)##

disables listed warnings and leaves any not listed in its current state.

with warning save

saves the current warning state, i.e. the list of all enabled warnings. This destroys any previously saved state.

with warning restore

causes the previously saved state to be restored.

without warning strict

overrides the -strict command line option, but only until the end of the next function or procedure.

The **with/without warnings** directives will have no effect if the -STRICT command line switch is used. The latter turns on all warnings and ignores any **with/without warnings** statement. However, it can be temporarily affected by the "without warning strict" directive.

Warning Names

Name	Meaning
none	When used with the with option, this turns off all warnings. When used with the without option, this turns on all warnings.
resolution	an identifier was used in a file, but was defined in a file this file doesn't (recursively) include.
short_circuit	a routine call may not take place because of short circuit evaluation in a conditional clause.
override	a built-in is being overridden
builtin_chosen	an unqualified call caused EUPHORIA to choose between a built-in and another global which does not override it. EUPHORIA chooses the built-in.
not_used	A variable has not been used and is going out of scope.
no_value	A variable never got assigned a value and is going out of scope.
custom	Any warning that was defined using the warning() procedure.
not_reached	After a keyword that branches unconditionally, the only thing that should appear is an end of block keyword, or possibly a label that a goto statement can target. Otherwise, there is no way that the statement can be reached at all. This warning notifies this condition.

translator	An option was given to the translator, but this option is not recognised as valid for the C compiler being used.
cmdline	A command line option was not recognised.
mixed_profile	For technical reasons, it is not possible to use both <code>with profile</code> and <code>with profile_time</code> in the same section of code. The profile statement read last is ignored, and this warning is issued.
empty_case	In <code>switch</code> that have without <code>fallthru</code> , an empty case block will result in no code being executed within the switch statement.
all	Turns all warnings on. They can still be disabled by <code>with/without warning</code> directives.

Example

```
with warning save
without warning &= (builtin_chosen, not_used)
. . . -- some code that might otherwise issue warnings
with warning restore
```

Initially, only the following warnings are enabled:

- resolution
- override
- builtin_chosen
- translator
- cmdline
- mixed_profile
- not_reached
- custom

This set can be changed using `-W` or `-X` command line switches.

`with_define`

4.8.2.2.2 with / without define

As mentioned about [ifdef statement](#), this top level statement is used to define/undefine tags which the `ifdef` statement may use.

The following tags have a predefined meaning in EUPHORIA:

- WIN32: platform is any version of Windows (tm) from '95 on to Vista and beyond
- WINDOWS: platform is any kind of Windows system
- UNIX: platform is any kind of Unix style system
- LINUX: platform is Linux
- FREEBSD: platform is FreeBSD
- OSX: platform is OS X for Macintosh
- SAFE: turns on a slower debugging version of `memory.e` called `safe.e` when defined. Switching mode by renaming files *no longer works*.
- EU4: defined on all versions of the version 4 interpreter
- EU400: defined on all versions of the interpreter from 4.0.0 to 4.0.X

- EU40000: defined only for version 4.0.0 of the interpreter

The name of a tag may contain any character that is a valid identifier character, that is A-Za-z0-9_. It is not required, but by convention defined words are upper case.

`with_inline`

4.8.2.3 `with` / `without inline`

This directive allows coders some flexibility with inlined routines. The default is for inlining to be on. Any routine that is defined when `without inline` is in effect will never be inlined.

`with inline` takes an optional integer parameter that defines the largest routine (by size of IL code) that will be considered for inlining. The default is 30.

5 Formal Syntax

basics

statements

 flow control

slice

if

switch

break

continue

retry

exit

fallthru

for

while

loop

goto

EUPHORIA Internals

 The EUPHORIA Data Structures

 MAKE_INT

 MAKE_SEQ

 IS_ATOM_INT

 DBL_PTR

 MAKE_SEQ

 IS_ATOM_DBL

 IS_ATOM

 IS_SEQUENCE

 IS_DBL_OR_SEQUENCE

The C Representations of a EUPHORIA Sequence and a EUPHORIA Double

5.1 basics

```
ALPHA ==: ('a' - 'z') | ('A' - 'Z')
```

```
DIGIT ==: ('0' - '9')
```

```
USCORE ==: '_'
```

```
IDENTIFIER ==: ( ALPHA | USCORE ) [ (ALPHA | DIGIT | USCORE) ... ]
```

```
NUMEXPR ==: (an expression that evaluates to an atom)
```

```
STREXPR ==: (an expression that evaluates to a string sequence)
```

```
SEQEXPR ==: (an expression that evaluates to an sequence)
```

```
BOOLEXPRESS ==: (an expression that evaluates to an atom in which zero represents falsehood and non-zero represents truth)
```

```
EXPR ==:
```

```
STATEMENT ==:

STMTBLK ==: STATEMENT [STATEMENT ...]

LABEL      ==:  'label' STRINGLIT

LISTDELIM  ==:  ','

STRINGLIT   ==: SIMPLESTRINGLIT | RAWSTRINGLIT

SIMPLESTRINGLIT ==: SSLITSTART [ (CHAR | ESCCHAR) ... ] SSLITEND
SSLITSTART  ==:  '""'
SSLITEND    ==:  '""'
CHAR        ==: (any byte value)
ESCCHAR     ==: ESCLEAD ( 't' | 'n' | 'r' | '\\' | '"' \ '' )
ESCLEAD     ==:  '\ '

RAWSTRINGLIT ==: DQRAWSTRING | BQRAWSTRING
DQRAWSTRING ==:  '""' [ MARGINSTR ] [CHAR ...] '""'
BQRAWSTRING ==:  `` [ MARGINSTR ] [CHAR ...] ``
MARGINSTR   ==:  ' _ ' ...
```

5.2 statements

5.2.1 flow control

IFSTMT
SWITCHSTMT
BREAKSTMT
CONTINUESTMT
RETRYSTMT
EXITSTMT
FALLTHRUSTMT
FORSTMT
WHILESTMT
LOOPSTMT
GOTOSTMT

5.3 slice

```
SLICE      ==: SLICESTART INTEXPRESSION SLICEDELIM INTEXPRESSION SLICEEND
SLICESTART ==:  '['
SLICEDELIM ==:  '...'
SLICEEND   ==:  ']'
```



5.4 if

```
IFSTMT ==: IFTEST [ ELSIF ... ] [ ELSE ] ENDIF
IFTEST ==: 'if' ATOMEXPR [ LABEL ] 'then' [ STMTBLOCK ]
ELSIF  ==: 'elsif' ATOMEXPR 'then' [ STMTBLOCK ]
ELSE   ==: 'else' [ STMTBLOCK ]
ENDIF  ==: 'end' 'if'
```

5.5 switch

```
SWITCHSTMT ==: SWITCHTEST CASE [ CASE ... ] [ CASEELSE ] [ ENDSWITCH ]
SWITCHTEST ==: 'switch' EXPRESSION [ WITHFALL ] [ LABEL ] 'do'
WITHFALL   ==: ('with' | 'without') 'fallthru'
CASE       ==: 'case' CASELIST 'then' [ STMTBLOCK ]
CASELIST   ==: EXPRESSION [ (LISTDELIM EXPRESSION) ... ]
CASEELSE   ==: 'case' 'else'
ENDSWITCH  ==: 'end' 'switch'
```

5.6 break

```
BREAKSTMT ==: 'break' [ STRINGLIT ]
```

5.7 continue

```
CONTINUESTMT ==: 'continue' [ STRINGLIT ]
```

5.8 retry

```
RETRYSTMT ==: 'retry' [ STRINGLIT ]
```

5.9 exit

```
EXITSTMT ==: 'exit' [ STRINGLIT ]
```

5.10 fallthru

```
FALLTHRUSTMT ==: 'fallthru'
```

5.11 for

```
FORSTMT ==: 'for' FORIDX [ LABEL ] 'do' [STMTBLK] 'end' 'for'
FORIDX  ==: IDENTIFIER '=' NUMEXPR 'to' NUMEXPR ['by' NUMEXPR]
```

5.12 while

```
WHILESTMT ==: 'while' BOOLEXP [WITHENTRY] [LABEL] 'do' STMTBLK [ENTRY] 'end' 'while'
WITHENTRY ==: 'with' 'entry'
ENTRY ==: 'entry' [STMTBLK]
```

5.13 loop

```
LOOPSTMT ==: 'loop' [WITHENTRY] [LABEL] 'do' STMTBLK [ENTRY] 'until' BOOLEXP 'end' 'loop'
```

5.14 goto

```
GOTOSMT ==: 'goto' LABEL
```

5.15 EUPHORIA Internals

The interpreter has four binary components:

- the translator
- library
- interpreter
- the backend.

EUPHORIA's parser first converts the code into a set of instructions that the translator, interpreter backend can process. Then the backend runs these instructions. The translator takes these same instructions and converts them into C-code. The library is called by the backend for the many builtins included in EUPHORIA.

5.15.1 The EUPHORIA Data Structures

5.15.1.1 The EUPHORIA representation of a EUPHORIA Object

Every EUPHORIA object is stored as-is. A special unlikely floating point value is used for NOVALUE. NOVALUE signifies that a variable has not been assigned a value or the end of a sequence.

5.15.1.2 The C Representation of a EUPHORIA Object

Every EUPHORIA object is either stored as is or an encoded pointer. A EUPHORIA integer is stored in a 32-bit signed integer. If the number is too big for a EUPHORIA integer or not an integer, it is assigned to a 64-bit double float in a structure and an encoded pointer to that structure is stored in the said 32-bit memory space. Sequences are stored in a similar way.

EUPHORIA integers are stored in object variables as-is. An object variable is a four byte signed integer. Legal integer values for EUPHORIA integers are between $-1,073,741,824$ ($-\text{power}(2,30)$) and $+1,073,741,823$ ($\text{power}(2,30)+1$). Unsigned hexadecimal numbers from C000_0000 to FFFF_FFFF are the negative integers and numbers from 0000_0000 to 3FFF_FFFF are the positive integers. The hexadecimal values not used as integers are thus 4000_0000 to BFFF_FFFF. Other values are for encoded pointers. Pointers are always 8 byte aligned. So a pointer is stored in 29-bits instead of 32 and can fit in a hexadecimal range 0x2000_0000 long. The other values are not stored in the same place but their encoded pointers are. The pointers are encoded in such a way that their values will never be in the range of the integers. Pointers to sequence structures (struct s1) are encoded into a range between 8000_0000 to 9FFF_FFFF. Pointers to structures for doubles (struct d) are encoded into a range between A000_0000 to BFFF_FFFF. A special value NOVALUE is at the end of the range of encoded pointers is BFFF_FFFF and it signifies that there is no value yet assigned to a variable and it also signifies the end of a sequence. These methods are how objects are stored. Values of this type are stored in the 'object' type.

A double structure 'struct d' could indeed contain a value that is legally in the range of a EUPHORIA integer. So the encoded pointer to this structure is recognized by the interpreter as an 'integer' but in this internals document when we say EUPHORIA integer we mean it actually is a C integer in the legal EUPHORIA integer range.

The macros are imperfect. For example, IS_SEQUENCE(NOVALUE) returns TRUE and IS_ATOM_DBL() will return TRUE for integer values as well as encoded pointers to 'struct d's. There is an order that these tests are made: We test IS_ATOM_INT and if that fails we can use IS_ATOM_DBL and then that will only be true if we pass an encoded pointer to a double. We must be sure that something is not NOVALUE before we use IS_SEQUENCE on it.

Often we know foo is not NOVALUE before getting into this:

```
// object foo
if (IS_ATOM_INT(foo)) {
    // some code for a EUPHORIA integer
} else if (IS_ATOM_DBL(foo)) {
    // some code for a double
} else {
    // code for a sequence foo
}
```

A sequence is held in a 'struct s1' type and a double is contained in a 'struct d'.

5.15.2 MAKE_INT

5.15.2.1 Signature:

<internal> object MAKE_INT(signed int x)

5.15.2.1.1 Returns an object with the same value as x. x must be within the integer range of a legal EUPHORIA integer type.

5.15.3 MAKE_SEQ

5.15.3.1 Signature:

<internal> object MAKE_SEQ(struct s1 * sptr)

5.15.3.1.1 Returns:

an object with an argument of a pointer to a 'struct s1'. The pointer is encoded into a range that is not a valid EUPHORIA 31-bit integer and returned.

5.15.4 IS_ATOM_INT

5.15.4.1 Signature:

<internal> int IS_ATOM_INT(object o)

5.15.4.1.1 Returns:

true if object is a EUPHORIA integer and not an encoded pointer.

5.15.4.1.2 Note:

IS_ATOM_INT() will return true even though the argument is out of the EUPHORIA integer range when the argument is positive. These values are not possible encoded pointers.

5.15.5 DBL_PTR

5.15.5.1 Signature:

<internal> struct d * DBL_PTR(object o)

5.15.5.1.1 Returns:

the pointer to a 'struct d' from the object o.

5.15.5.1.2 Assumption:

IS_ATOM_INT(o) is FALSE and IS_ATOM_DBL(o) is TRUE.

5.15.6 MAKE_SEQ**5.15.6.1 Signature:**

<internal> struct s1 * MAKE_SEQ(object o)

5.15.6.1.1 Returns:

the pointer to a 'struct s1' from the object o.

5.15.6.1.2 Assumption:

IS_SEQUENCE(o) is TRUE and /o/ is not NOVALUE.

Use MAXINT and MININT to check for overflow and underflow.

5.15.7 IS_ATOM_DBL**5.15.7.1 Signature:**

<internal> int IS_ATOM_DBL(object o)

5.15.7.1.1 Returns:

true if the object is an encoded pointer to a double struct.

5.15.7.1.2 Assumption:

o must not be a EUPHORIA integer.

5.15.8 IS_ATOM**5.15.8.1 Signature:**

<internal> int IS_ATOM(object *o*)

5.15.8.1.1 Returns:

true if the object is a EUPHORIA integer or an encoded pointer to a 'struct d'.

5.15.9 IS_SEQUENCE**5.15.9.1 Signature:**

<internal> int IS_SEQUENCE(object *o*)

5.15.9.1.1 Returns:

true if the object is an encoded pointer to a 'struct s1'.

5.15.9.1.2 Assumption:

o is not NOVALUE.

5.15.10 IS_DBL_OR_SEQUENCE**5.15.10.1 Signature:**

<internal> int IS_DBL_OR_SEQUENCE(object *o*)

5.15.10.1.1 Returns:

true if the object is an encoded pointer of either kind of structure.



5.16 The C Representations of a EUPHORIA Sequence and a EUPHORIA Double

struct s1

```
{
    object_ptr base;    // base is such that base[1] is the first element
    long length;        // this is the sequence length
    long ref;           // ref is the number of as virtual copies of this sequence
    long postfill;      // is how many extra objects could fit at the end of base
    cleanup_ptr cleanup; // this is a pointer to a EUPHORIA routine that is run
                        // just before the sequence is freed.
}
```

struct d

```
{
    double dbl;         // the actual value of a double number.
    long ref;           // ref is the number of virtual copies of this double
    cleanup_ptr cleanup; // this is a pointer to a EUPHORIA routine that is run
                        // just before the sequence is freed.
}
```

Now offset of the 'ref' in struct d must be the same as the offset of the 'ref' in struct s1. A 64bit implementation would have to reorder these members.

6 Mini-Guides

Debugging and Profiling

- Debugging
 - The Trace Screen
 - The Trace File
- Profiling
 - Some Further Notes on Time Profiling

Binding and Shrouding

- The Shroud Command
- The Bind Command

EUPHORIA To C Translator

- Introduction
- C Compilers Supported
- How to Run the Translator
- Command-Line Options
- Dynamic Link Libraries (Shared Libraries)
- Executable Size and Compression
- Interpreter vs. Translator
- Legal Restrictions
- Disclaimer:
- Frequently Asked Questions
- Common Problems

Indirect routine calling

- Indirect calling a routine coded in EUPHORIA
- Calling EUPHORIA's internals

Multitasking in EUPHORIA

- Introduction
- Why Multitask?
- Types of Tasks
- A Small Example
- Comparison with earlier multitasking schemes
- Comparison with multithreading
- Summary

EUPHORIA Database System (EDS)

- Introduction
- Structure of an EDS database
- How to access the data
- How does storage get recycled?
- Security / Multi-user Access
- Scalability
- Disclaimer
- Warning: Use the right file mode

The User Defined Pre-Processor

- A Quick Example
- Pre-process Details
- Command Line Options
- DLL/Shared Library Interface

- Advanced Examples
- EUPHORIA Trouble-Shooting Guide
 - Common Problems and Solutions
- Platform Specific Issues
 - Introduction
 - The WIN32 Platform
 - The Unix Platforms
 - Interfacing with C Code (WIN32, Linux, FreeBSD)
- Performance Tips
 - General Tips
 - Measuring Performance
 - How to Speed-Up Loops
 - Converting Multiplies to Adds in a Loop
 - Saving Results in Variables
 - In-lining of Routine Calls
 - Operations on Sequences
 - Some Special Case Optimizations
 - Assignment with Operators
 - Pixel-Graphics Tips
 - Library Routines
 - Searching
 - Sorting
 - Taking Advantage of Cache Memory
 - Using Machine Code and C
 - Using The EUPHORIA To C Translator

6.1 Debugging and Profiling

- Debugging
 - The with/without trace directive
- The Trace Screen
- The Trace File
- Profiling
 - Some Further Notes on Time Profiling

6.1.1 Debugging

Extensive run-time checking provided by the EUPHORIA interpreter catches many bugs that in other languages might take hours of your time to track down. When the interpreter catches an error, you will always get a brief report on your screen, and a detailed report in a file called `ex.err`. These reports include a full English description of what happened, along with a call-stack traceback. The file `ex.err` will also have a dump of all variable values, and optionally a list of the most recently executed statements. For extremely large sequences, only a partial dump is shown. If the name `ex.err` is not convenient, or if a nondefault path is required, you can choose another file name, anywhere on your system, by calling `crash_file()`.

In addition, you are able to create **user-defined types** that precisely determine the set of legal values for each of your variables. An error report will occur the moment that one of your variables is assigned an illegal value.

Sometimes a program will misbehave without failing any run-time checks. In any programming language it may be a good idea to simply study the source code and rethink the algorithm that you have coded. It may also be useful to insert print statements at strategic locations in order to monitor the internal logic of the program. This approach is particularly convenient in an interpreted language like EUPHORIA since you can simply edit the source and rerun the program without waiting for a re-compile/re-link. `trace`

6.1.1.1 The with/without trace directive

The interpreter provides you with additional powerful tools for debugging. Using `trace(1)` you can **trace** the execution of your program on one screen while you witness the output of your program on another. `trace(2)` is the same as `trace(1)` but the trace screen will be in monochrome. Finally, using `trace(3)`, you can log all executed statements to a file called **ctrace.out**.

The **with/without trace** special statements select the parts of your program that are available for tracing. Often you will simply insert a with trace statement at the very beginning of your source code to make it all traceable. Sometimes it is better to place the first with trace after all of your **user-defined types**, so you don't trace into these routines after each assignment to a variable. At other times, you may know exactly which routine or routines you are interested in tracing, and you will want to select only these ones. Of course, once you are in the trace window, you can skip viewing the execution of any routine by pressing down-arrow on the keyboard rather than Enter. However, once inside a routine, you must step through till it returns, even if stepping in was an mistake.

Only traceable lines can appear in **ctrace.out** or in **ex.err** as "Traced lines leading up to the failure", should a run-time error occur. If you want this information and didn't get it, you should insert a with trace and then rerun your program. Execution will be slower when lines compiled with trace are executed, especially when `trace(3)` is used.

After you have predetermined the lines that are traceable, your program must then dynamically cause the trace facility to be activated by executing a `trace()` statement. You could simply say:

```
with trace
trace(1)
```

However, you cannot dynamically set or free breakpoints while tracing. You must abort program, edit, change setting, save and run again.

at the top of your program, so you can start tracing from the beginning of execution. More commonly, you will want to trigger tracing when a certain routine is entered, or when some condition arises. e.g.

```
if x < 0 then
    trace(1)
end if
```

You can turn off tracing by executing a `trace(0)` statement. You can also turn it off interactively by typing 'q' to quit tracing. Remember that with trace must appear **outside** of any routine, whereas `trace()` can appear

inside a routine *or outside*.

You might want to turn on tracing from within a [type](#). Suppose you run your program and it fails, with the **ex.err** file showing that one of your variables has been set to a strange, although not illegal value, and you wonder how it could have happened. Simply create a type for that variable that executes `trace(1)` if the value being assigned to the variable is the strange one that you are interested in. e.g.

```
type positive_int(integer x)
  if x = 99 then
    trace(1) -- how can this be???
    return 1 -- keep going
  else
    return x > 0
  end if
end type
```

When `positive_int()` returns, you will see the exact statement that caused your variable to be set to the strange value, and you will be able to check the values of other variables. You will also be able to check the output screen to see what has happened up to this precise moment. If you define `positive_int()` so it returns 0 for the strange value (99) instead of 1, you can force a diagnostic dump into **ex.err**.

Remember that the argument to `trace()` does not need to be a constant. It only needs to be 0, 1, 2 or 3, but these values may be the result from any expression passed to `trace()`. Other values will cause `trace()` to fail.

6.1.2 The Trace Screen

When a `trace(1)` or `trace(2)` statement is executed by the interpreter, your main output screen is saved and a **trace screen** appears. It shows a view of your program with the statement that will be executed next highlighted, and several statements before and after showing as well. You cannot scroll the window further up or down though. Several lines at the bottom of the screen are reserved for displaying variable names and values. The top line shows the commands that you can enter at this point:

Command	Action
F1	display main output screen take a look at your program's output so far
F2	redisplay trace screen. Press this key while viewing the main output screen to return to the trace display.
Enter	execute the currently-highlighted statement only continue execution and break when any statement coming after this one in the source listing is about to be executed.
down-arrow	This lets you skip over subroutine calls. It also lets you stop on the first statement following the end of a loop without having to witness all iterations of the loop.
?	display the value of a variable. After hitting ? you will be prompted for the name of the variable. Many variables are displayed for you automatically as they are assigned a value. If a variable is not currently being displayed, or is only partially displayed, you can ask for it.

Large sequences are limited to one line on the trace screen, but when you ask for the value of a variable that contains a large sequence, the screen will clear, and you can scroll through a pretty-printed display of the sequence. You will then be returned to the trace screen, where only one line of the variable is displayed. Variables that are not defined at this point in the program cannot be shown. Variables that have not yet been initialized will have "< NO VALUE >" beside their name. Only variables, not general expressions, can be displayed. As you step through execution of the program, the system will update any values showing on the screen. Occasionally it will remove variables that are no longer in scope, or that haven't been updated in a long time compared with newer, recently-updated variables.

- q quit tracing and resume normal execution. Tracing will start again when the next trace(1) is executed.
- Q quit tracing and let the program run freely to its normal completion. trace() statements will be ignored.
- ! this will abort execution of your program. A traceback and dump of variable values will go to `ex.err`.

As you trace your program, variable names and values appear automatically in the bottom portion of the screen. Whenever a variable is assigned to, you will see its name and new value appear at the bottom. This value is always kept up-to-date. Private variables are automatically cleared from the screen when their routine returns. When the variable display area is full, least-recently referenced variables will be discarded to make room for new variables. The value of a long sequence will be cut off after 80 characters.

For your convenience, numbers that are in the range of printable ASCII characters (32-127) are displayed along with the ASCII character itself. The ASCII character will be in a different color (or in quotes in a mono display). This is done for all variables, since EUPHORIA does not know in general whether you are thinking of a number as an ASCII character or not. You will also see ASCII characters (in quotes) in `ex.err`. This can make for a rather "busy" display, but the ASCII information is often very useful.

The trace screen adopts the same graphics mode as the main output screen. This makes flipping between them quicker and easier.

When a traced program requests keyboard input, the main output screen will appear, to let you type your input as you normally would. This works fine for a `gets()` (read one line) input. When a `get_key()` (quickly sample the keyboard) is called you will be given 8 seconds to type a character, otherwise it is assumed that there is no input for this call to `get_key()`. This allows you to test the case of input and also the case of no input for `get_key()`.

6.1.3 The Trace File

When your program calls trace(3), tracing to a file is activated. The file, **ctrace.out** will be created in the current directory. It contains the last 500 EUPHORIA statements that your program executed. It is set up as a circular buffer that holds a maximum of 500 statements. Whenever the end of **ctrace.out** is reached, the next statement is written back at the beginning. The very last statement executed is always followed by "=== THE END ===". Because it's circular, the last statement executed could appear anywhere in **ctrace.out**. The statement coming after "=== THE END ===" is the 500th-last.

This form of tracing is supported by both the Interpreter and the the EUPHORIA To C Translator. It is particularly useful when a machine-level error occurs that prevents EUPHORIA from writing out an `ex.err`

diagnostic file. By looking at the last statement executed, you may be able to guess why the program crashed. Perhaps the last statement was a `poke()` into an illegal area of memory. Perhaps it was a call to a C routine. In some cases it might be a bug in the interpreter or the Translator.

The source code for a statement is written to **ctrace.out**, and flushed, just *before* the statement is performed, so the crash will likely have happened *during* execution of the final statement that you see in **ctrace.out**.

6.1.4 Profiling

If you specify a `with profile` or `with profile_time` (*Windows* only) directive, then a special listing of your program, called a *profile*, will be produced by the interpreter when your program finishes execution. This listing is written to the file **ex.pro** in the current directory.

There are two types of profiling available: **execution-count profiling**, and **time profiling**. You get execution-count profiling when you specify `with profile`. You get time profiling when you specify `with profile_time`. You can't mix the two types of profiling in a single run of your program. You need to make two separate runs.

We ran the **sieve8k.ex** benchmark program in **demo\bench** under both types of profiling. The results are in **sieve8k.pro** (execution-count profiling) and **sieve8k.pro2** (time profiling). `profile_file` Execution-count profiling shows precisely how many times each statement in your program was executed. If the statement was never executed the count field will be blank.

Time profiling shows an estimate of the total time spent executing each statement. This estimate is expressed as a percentage of the time spent profiling your program. If a statement was never sampled, the percentage field will be blank. If you see 0.00 it means the statement was sampled, but not enough to get a score of 0.01.

Only statements compiled with `profile` or `with profile_time` are shown in the listing. Normally you will specify either `with profile` or `with profile_time` at the top of your main **.ex** file, so you can get a complete listing. View this file with the EUPHORIA editor to see a color display.

Profiling can help you in many ways:

- it lets you see which statements are heavily executed, as a clue to speeding up your program
- it lets you verify that your program is actually working the way you intended
- it can provide you with statistics about the input data
- it lets you see which sections of code were never tested -- don't let your users be the first!

Sometimes you will want to focus on a particular action performed by your program. For example, in the **Language War** game, we found that the game in general was fast enough, but when a planet exploded, shooting 2500 pixels off in all directions, the game slowed down. We wanted to speed up the explosion routine. We didn't care about the rest of the code. The solution was to call `profile(0)` at the beginning of Language War, just after `with profile_time`, to turn off profiling, and then to call `profile(1)` at the beginning of the explosion routine and `profile(0)` at the end of the routine. In this way we could run the game, creating numerous explosions, and logging a lot of samples, just for the explosion effect. If samples were charged against other lower-level routines, we knew that those samples occurred during an explosion. If we had simply profiled the whole program, the picture would not have been clear, as the lower-level routines

would also have been used for moving ships, drawing phasors etc. `profile()` can help in the same way when you do execution-count profiling.

6.1.5 Some Further Notes on Time Profiling

With each click of the system clock, an interrupt is generated. When you specify with `profile_time` EUPHORIA will sample your program to see which statement is being executed at the exact moment that each interrupt occurs.

Each sample requires 4 bytes of memory and buffer space is normally reserved for 25000 samples. If you need more than 25000 samples you can request it:

```
with profile_time 100000
```

will reserve space for 100000 samples (for example). If the buffer overflows you'll see a warning at the top of **ex.pro**. At 100 samples per second your program can run for 250 seconds before using up the default 25000 samples. It's not feasible for EUPHORIA to dynamically enlarge the sample buffer during the handling of an interrupt. That's why you might have to specify it in your program. After completing each top-level executable statement, EUPHORIA will process the samples accumulated so far, and free up the buffer for more samples. In this way the profile can be based on more samples than you have actually reserved space for.

The percentages shown in the left margin of **ex.pro**, are calculated by dividing the number of times that a particular statement was sampled, by the total number of samples taken. e.g. if a statement were sampled 50 times out of a total of 500 samples, then a value of 10.0 (10 per cent) would appear in the margin beside that statement. When profiling is disabled with `profile(0)`, interrupts are ignored, no samples are taken and the total number of samples does not increase.

By taking more samples you can get more accurate results. However, one situation to watch out for is the case where a program synchronizes itself to the clock interrupt, by waiting for `time()` to advance. The statements executed just after the point where the clock advances might *never* be sampled, which could give you a very distorted picture. e.g.

```
while time() < LIMIT do
end while
x += 1 -- This statement will never be sampled
```

Sometimes you will see a significant percentage beside a return statement. This is usually due to time spent deallocating storage for temporary and private variables used within the routine. Significant storage deallocation time can also occur when you assign a new value to a large sequence.

If disk swapping starts to happen, you may see large times attributed to statements that need to access the swap file, such as statements that access elements of a large swapped-out sequence.

6.2 Binding and Shrouding

6.2.1 The Shroud Command

6.2.1.1 Synopsis

```
shroud [-full_debug] [-list] [-quiet] [-out shrouded_file] filename.ex[w/u]
```

The `shroud` command converts a EUPHORIA program, typically consisting of a main file plus many include files, into a single, compact file. A single file is not only convenient, it allows you to give people your program to use, without giving them your source code.

A shrouded file does not contain any EUPHORIA source code statements. Rather, it contains a low-level intermediate language (IL) that is executed by the back-end of the interpreter. A shrouded file does not require any parsing. It starts running immediately, and with large programs you will see a quicker start-up time. Shrouded files must be run using the interpreter back-end: `eubw.exe` (*Windows*) or `eub.exe` (*Linux/FreeBSD*). This backend is freely available, and you can give it to any of your users who need it. It's stored in `euphoria\bin` in the EUPHORIA interpreter package. You can run your `.il` file with:

On *Windows* use:

```
backendw myprog.il
```

On *Unix* use:

```
eub.exe myprog.il
```

Although it does not contain any source statements, a `.il` file will generate a useful `ex.err` dump in case of a run-time error.

The shrouder will remove any routines and variables that your program doesn't use. This will give you a smaller `.il` file. There are often a great number of unused routines and unused variables. For example your program might include several 3rd party include files, plus some standard files from `euphoria\include`, but only use a few items from each file. The unused items will be deleted.

6.2.1.2 Options

- **-full_debug**: Make a somewhat larger `.il` file that contains enough debug information to provide a full `ex.err` dump when a crash occurs. Normally, variable names and line-number information is stripped out of the `.il` file, so the `ex.err` will simply have "no-name" where each variable name should be, and line numbers will only be accurate to the start of a routine or the start of a file. Only the private variable values are shown, not the global or local values. In addition to saving space, some people might prefer that the shrouded file, and any `ex.err` file, not expose as much information.
- **-list**: Produce a listing in **deleted.txt** of the routines and constants that were deleted.
- **-quiet**: Suppress normal messages and statistics. Only report errors.
- **-out shrouded_file**: Write the output to `shrouded_file`.

The EUPHORIA interpreter will not perform tracing on a shrouded file. You must trace your original source.

On *Unix*, the shrouder will make your shrouded file executable, and will add a `#!` line at the top, that will run `eub.exe`. You can override this `#!` line by specifying your own `#!` line at the top of your main EUPHORIA file.

Always keep a copy of your original source. There's no way to recover it from a shrouded file.

6.2.2 The Bind Command

6.2.2.1 Synopsis:

```
bind  [-full_debug] [-list] [-quiet] [-out executable_file] [filename.ex]
bindu [-full_debug] [-list] [-quiet] [-out executable_file] [filename.ex]
bindw [-full_debug] [-list] [-quiet] [-out executable_file] [-con] [-icon filename.ico] [filename]
```

`bind` (`bindw` or `bindu`) does the same thing as `shroud`, and includes the same options. It then combines your shrouded `.il` file with the interpreter backend (`eubd.exe`, `eubw.exe` or `eub.exe`) to make a **single, stand-alone executable** file that you can conveniently use and distribute. Your users need not have EUPHORIA installed. Each time your executable file is run, a quick integrity check is performed to detect any tampering or corruption. Your program will start up very quickly since no parsing is needed.

The EUPHORIA interpreter will not perform tracing on a bound file since the source statements are not there.

6.2.2.2 Options:

- **-full_debug**: Same as `shroud` above. If EUPHORIA detects an error, your executable will generate either a partial, or a full, `ex.err` dump, according to this option.
- **-list**: Same as `shroud` above.
- **-quiet**: Same as `shroud` above.
- **-out executable_file**: This option lets you choose the name of the executable file created by the binder. Without this option, `bind` will choose a name based on the name of the main EUPHORIA source file.
- **-con: (bindw only)**: This option will create a *Windows* console program instead of a *Windows GUI* program. Console programs can access standard input and output, and they work within the current console window, rather than popping up a new one.
- **-icon filename[.ico]: (bindw only)** When you bind a program, you can patch in your own customized icon, overwriting the one in `euiw.exe`. `eui.exe` contains a 32x32 icon using 256 colors. It resembles an **E** shape. *Windows* will display this shape beside `euiw.exe`, and beside your bound program, in file listings. You can also load this icon as a resource, using the name "euiw" (see `euphoria\demo\win32>window.exw` for an example). When you bind your program, you can substitute your own 32x32 256-color icon file of size 2238 bytes or less. Other dimensions may also work as long as the file is 2238 bytes or less. The file must contain a single icon image (*Windows* will create a smaller or larger image as necessary). The default **E** icon file, `euphoria.ico`, is included in the `euphoria\bin` directory.

A one-line EUPHORIA program will result in an executable file as large as the back-end you are binding with, but the size increases very slowly as you add to your program. **When bound, the entire EUPHORIA editor, `ed.exe`, adds only 27K to the size of the back-end.** `eubw.exe` and `eub.exe` *Linux* are compressed using a [UPX](#). Note: In some very rare cases, a compressed executable may trigger a warning message from a virus scanner. This is simply because the executable file looks abnormal to the virus scanner.

The first two items returned by `command_line()` will be slightly different when your program is bound. See the procedure description for the details.

A **bound executable** file *can* handle standard input and output redirection. e.g.

```
myprog.exe < file.in > file.out
```

If you were to write a small `.bat` file, say `myprog.bat`, that contained the line `"eui myprog.ex"` you would *not* be able to redirect input and output. The following will not work:

```
myprog.bat < file.in > file.out
```

You *could* however use redirection on individual lines *within* the `.bat` file.

6.3 EUPHORIA To C Translator

6.3.1 Introduction

The EUPHORIA to C Translator will translate any EUPHORIA program into equivalent C source code.

There are versions of the translator for *Windows*, *Linux*, *FreeBSD* and *Mac OS X*. After translating a EUPHORIA program to C, you can compile and link using one of the supported C compilers. This will give you an executable file that will typically run much faster than if you used the EUPHORIA interpreter.

The translator can translate/compile **itself** into an executable file for each platform. The translator is also used in translating/compiling the front-end portion of the interpreter. The source code for the Translator is in `euphoria\source`. It's written 100% in EUPHORIA.

6.3.2 C Compilers Supported

The **Translator** currently works with GNU C on *Linux* or *FreeBSD*, and with [Watcom C](#) on *Windows*. These are all **free** compilers.

GNU C will exist already on your Linux or FreeBSD system. The others can be downloaded from their respective Web sites.

6.3.2.1 Notes:

- Warnings are turned off when compiling with `emake.bat`. If you turn them on, you may see some harmless messages about variables declared but not used, labels defined but not used, function prototypes not declared etc.

6.3.3 How to Run the Translator

Running the **Translator** is similar to running the **Interpreter**. On Windows you would type:

```
euc taskwire.exw
```

On Linux/FreeBSD you would type:

```
euc qsort.ex
```

but instead of running the `allsorts.ex` program, the **Translator** will create several C source files. Anyone can run the **Translator**. It's included in `euphoria\bin` along with the interpreter. To compile and link the C files, you need to install one of the supported C compilers. The **Translator** creates a batch file called `emake.bat` that does all the compiling and linking steps for you, so you don't actually have to know anything about C or C compilers. Just type:

```
emake
```

When the C compiling and linking is finished, you will have a file called:

```
allsorts.exe
```

and the C source files will have been removed to avoid clutter.

When you run `allsorts.exe`, it should run the same as if you had typed:

```
eui allsorts
```

to run it with the **Interpreter**, except that it should run faster, showing reduced times for the various sorting algorithms in `euphoria\demo\allsorts.ex`.

After creating your executable file, `emake` removes all the C files that were created. If you want to look at these files, run the translator again and look at the files before running `emake`.

6.3.3.1 Note to Linux, FreeBSD and OSX users:

The files will be called `emake` and `shell`, and you type `./emake` to perform the compiles and link, and `./shell` to run the shell sort program.

6.3.4 Command-Line Options

If you happen to have more than one C compiler for a given platform, you can select the one you want to use with a command-line option:

```
-wat
```

on the command line to **euc** or **euc**. e.g.

```
euc -bor pretend.exw
```

Normally, after building your `.exe` file, the **emake** batch file will delete all C files and object files produced by the Translator. If you want **emake** to keep these files, add the `-keep` option to the Translator command-line. e.g.

```
euc -wat -keep sanity.ex
```

To make a *Windows* `.dll` file, or *Linux* or *FreeBSD* `.so` file, just add `-dll` to the command line. e.g.

```
euc -bor -dll mylib.ew
```

To make a *Windows* console program instead of a *Windows* GUI program, add `-con` to the command line. e.g.

```
euc -bor -con myprog.exw
```

To increase or decrease the total amount of stack space reserved for your program, add `-stack nnnn` to the command line. e.g.

```
euc -stack 100000 myprog.ex
```

The total stack space (in bytes) that you specify will be divided up among all the tasks that you have running (assuming you have more than one). Each task has its own private stack space. If it exceeds its allotment, you'll get a run-time error message identifying the task and giving the size of its stack space. Most non-recursive tasks can run with call stacks as small as 2000 bytes, but to be safe, you should allow more than this. A deeply-recursive task could use a great deal of space. It all depends on the maximum levels of calls that a task might need. At run-time, as your program creates more simultaneously-active tasks, the stack space allotted to each task will tend to decrease.

To compile your program with debugging information, usable with a debugger compatible with your compiler, use the `-debug` option:

```
euc -debug myapp.ex
```

It is sometimes useful to link your translated code to a EUPHORIA runtime library other than the default supplied library. This ability is probably mostly useful for testing and debugging the runtime library itself, or to give additional debugging information when debugging translated EUPHORIA code. Note that only the default library is supplied. Use the `-lib {library}` option:

```
euc -lib decu.a myapp.ex
```



```
-plat FREEBSD
-plat LINUX
-plat OSX
-plat WIN
-plat NETBSD
-plat OPENBSD
-plat SUNOS
```

Use one of these options to translate code for the specified platform, even if running on a different platform. The default will always be the native platform of the translator that is executed, so `euc.exe` will default to Windows, and `euc` will default to the platform upon which it was built.

The resulting output can be compiled by the appropriate compiler on the specified platform, or, possibly a cross platform compiler, if you have one configured.

6.3.5 Dynamic Link Libraries (Shared Libraries)

Simply by adding `-dll` to the command line, the **Translator** will build a *Windows* `.dll` (*Linux/FreeBSD* `.so`) file instead of an executable program.

You can translate and compile a set of useful EUPHORIA routines, and share them with other people, without giving them your source. Furthermore, your routines will likely run much faster when translated and compiled. Both translated/compiled and interpreted programs will be able to use your library.

Only the global EUPHORIA procedures and functions, i.e. those declared with the "global", "public" or "export" keyword, will be exported from the `.dll` (`.so`).

Any EUPHORIA program, whether translated/compiled or interpreted, can link with a EUPHORIA `.dll` (`.so`) using the same mechanism that lets you link with a `.dll` (`.so`) written in C. The program first calls `open_dll()` to open the `.dll` or `.so` file, then it calls `define_c_func()` or `define_c_proc()` for any routines that it wants to call. It calls these routines using `c_func()` and `c_proc()`.

The routine names exported from a EUPHORIA `.dll` will vary depending on which C compiler you use.

GNU C on *Linux* or *FreeBSD* exports the names exactly as they appear in the C code produced by the **Translator**, e.g. a EUPHORIA routine

```
global procedure foo(integer x, integer y)
```

would be exported as `"_0foo"` or maybe `"_1foo"` etc. The underscore and digit are added to prevent naming conflicts. The digit refers to the EUPHORIA file where the identifier is defined. The main file is numbered as 0. The include files are numbered in the order they are encountered by the compiler. You should check the C source to be sure.

For Watcom the **Translator** also creates an `EXPORT` command, added to `objfiles.lnk` for each exported identifier, so `foo()` would be exported as `"foo"`.

With Watcom you can edit the `objfiles.lnk` file, and rerun `emake.bat`, to rename the exported identifiers, or remove ones that you don't want to export.

Having nice exported names is not critical, since the name need only appear once in each EUPHORIA program that uses the .dll, i.e. in a single `define_c_func()` or `define_c_proc()` statement. The author of a .dll should probably provide his users with a EUPHORIA include file containing the necessary `define_c_func()` and `define_c_proc()` statements, and he might even provide a set of EUPHORIA "wrapper" routines to call the routines in the .dll.

When you call `open_dll()`, any top-level EUPHORIA statements in the .dll or .so will be executed automatically, just like a normal program. This gives the library a chance to initialize its data structures prior to the first call to a library routine. For many libraries no initialization is required.

To pass EUPHORIA data (atoms and sequences) as arguments, or to receive a EUPHORIA object as a result, you will need to use the following constants in `euphoria\include\dll.e`:

```
-- EUPHORIA types for .dll (.so) arguments and return values:
global constant
    E_INTEGER = #06000004,
    E_ATOM    = #07000004,
    E_SEQUENCE= #08000004,
    E_OBJECT   = #09000004
```

Use these in `define_c_proc()` and `define_c_func()` just as you currently use `C_INT`, `C_UINT` etc. to call C .dll's and .so's.

Currently, file numbers returned by `open()`, and routine id's returned by `routine_id()`, can be passed and returned, but the library and the main program each have their own separate ideas of what these numbers mean. Instead of passing the file number of an open file, you could instead pass the file name and let the .dll (.so) open it. Unfortunately there is no simple solution for passing routine id's. This might be fixed in the future.

A EUPHORIA .dll or .so currently may not execute any multitasking operations. The Translator will give you an error message about this.

EUPHORIA .dlls (.so's) can also be used by C programs as long as only 31-bit integer values are exchanged. If a 32-bit pointer or integer must be passed, and you have the source to the C program, you could pass the value in two separate 16-bit integer arguments (upper 16 bits and lower 16 bits), and then combine the values in the EUPHORIA routine into the desired 32-bit atom.

6.3.6 Executable Size and Compression

emake does not include a command to compress your executable file. If you want to do this we suggest you try the free [UPX](#) compressor.

Large Win32Lib-based .exe's produced by the Translator can be compressed by UPX to about 15% of their original size, and you won't notice any difference in start-up time.

The **Translator** deletes routines that are not used, including those from the standard EUPHORIA include files. After deleting unused routines, it checks again for more routines that have now become unused, and so on. This can make a big difference, especially with Win32Lib-based programs where a large file is included,

but many of the included routines are not used in a given program.

Nevertheless, your compiled executable file will likely be larger than the same EUPHORIA program bound with the interpreter **back-end**. This is partly due to the **back-end** being a compressed executable. Also, EUPHORIA statements are extremely compact when stored in a bound file. They need more space after being translated to C, and compiled into machine code. Future versions of the **Translator** will produce faster and smaller executables.

6.3.7 Interpreter vs. Translator

All EUPHORIA programs can be translated to C, and with just a few exceptions noted below, will run the same as with the **Interpreter** (but hopefully faster).

The **Interpreter** and **Translator** share the same parser, so you will get the same syntax errors, variable not declared errors etc. with either one.

The **Interpreter** automatically expands the call stack (until memory is exhausted), so you can have a huge number of levels of nested calls. Most C compilers, on most systems, have a pre-set limit on the size of the stack. Consult your compiler or linker manual if you want to increase the limit, for example if you have a recursive routine that might need thousands of levels of recursion. Modify the link command in `emake.bat`. For Watcom C, use `OPTION STACK=nnnn`, where `nnnn` is the number of bytes of stack space.

6.3.7.1 Note:

The **Translator** assumes that your program has no run-time errors in it that would be caught by the **Interpreter**. The **Translator** does not check for: subscript out of bounds, variable not initialized, assigning the wrong type of data to a variable, etc.

You should **debug** your program with the **Interpreter**. The **Translator** checks for certain run-time errors, but in the interest of speed, most are not checked. When translated C code crashes you'll typically get a very cryptic machine exception. In most cases, the first thing you should do is run your program with the **Interpreter**, using the same inputs, and preferably with `type_check` turned on. If the error only shows up in translated code, you can use `with trace` and `trace(3)` to get a `ctrace.out` file showing a circular buffer of the last 500 EUPHORIA statements executed. If a translator-detected error message is displayed (and stored in `ex.err`), you will also see the offending line of EUPHORIA source whenever `with trace` is in effect. `with trace` will slow your program down, and the slowdown can be extreme when `trace(3)` is also in effect.

6.3.8 Legal Restrictions

As far as RDS is concerned, any executable programs or .dll's that you create with this **Translator** without modifying an RDS translator library file, may be distributed royalty-free. You are free to incorporate any EUPHORIA files provided by RDS into your application.

In general, if you wish to use EUPHORIA code written by 3rd parties, you had better honor any restrictions that apply. If in doubt, you should ask for permission.

On *Linux*, *FreeBSD*, the GNU Library licence will normally not affect programs created with this **Translator**. Simply compiling with GNU C does not give the Free Software Foundation any jurisdiction over your program. If you statically link their libraries you will be subject to their Library licence, but the standard compile/link procedure in `emake` does not statically link any FSF libraries, so there should be no problem.

The Allegro graphics library, used by DJGPP, is referred to as "Giftware" in their documentation, and they allow you to redistribute it as part of your program. They ask for, but do not require, some acknowledgement.

6.3.9 Disclaimer:

This is what we believe to be the case. We are not lawyers. If it's important to you, you should read **all** licences and the legal comments in them, to form your own judgement. You may need to get professional legal opinion as well.

6.3.10 Frequently Asked Questions

6.3.10.1 How much of a speed-up should I expect?

It all depends on what your program spends its time doing. Programs that use mainly integer calculations, don't call run-time routines very often, and don't do much I/O will see the greatest improvement, currently up to about 5x faster. Other programs may see only a few percent improvement.

The various C compilers are not equal in optimization ability.

6.3.10.2 What if I want to change the compile or link options in `emake.bat`?

Feel free to do so, however you should copy `emake.bat` to your own file called (say) `mymake.bat`, then run `mymake.bat` after running the Translator. Occasionally the number of `.c` files produced by the Translator could change.

6.3.10.3 How can I make my program run even faster?

It's important to declare variables as integer where possible. In general, it helps if you choose the most restrictive type possible when declaring a variable.

Typical user-defined types will not slow you down. Since your program is supposed to be free of `type_check` errors, types are ignored by the Translator, unless you call them directly with normal function calls. The one

exception is when a user-defined type routine has side-effects (i.e. it sets a global variable, performs pokes into memory, I/O etc.). In that case, if **with type_check** is in effect, the Translator will issue code to call the type routine and report any type_check failure that results.

On *Windows* we have left out the `/o1` loop optimization for Watcom's `wcc386`. We found in a couple of rare cases that this option led to incorrect machine code being emitted by the Watcom C compiler. If you add it back in to your own version of `emake.bat` you might get a slight improvement in speed, with a slight risk of buggy code.

On *Linux* or *FreeBSD* you could try the `-O3` option of `gcc` instead of `-O2`. It will "in-line" small routines, improving speed slightly, but creating a larger executable. You could also try the [Intel C++ Compiler for Linux](#). It's compatible with GNU C, but some adjustments to `emake` might be required.

6.3.11 Common Problems

Many large programs have been successfully translated and compiled using each of the supported C compilers, and the Translator is now quite stable.

6.3.11.1 Note:

On *Windows*, if you call a C routine that uses the `cdecl` calling convention (instead of `stdcall`), you must specify a '+' character at the start of the routine's name in `define_c_proc()` and `define_c_func()`. If you don't, the call may not work when running the `eui` Interpreter.

In some cases a huge EUPHORIA routine is translated to C, and it proves to be too large for the C compiler to process. If you run into this problem, make your EUPHORIA routine smaller and simpler. You can also try turning off C optimization in `emake.bat` for just the `.c` file that fails. Breaking up a single constant declaration of many variables into separate constant declarations of a single variable each, may also help. EUPHORIA has no limits on the size of a routine, or the size of a file, but most C compilers do. The Translator will automatically produce multiple small `.c` files from a large EUPHORIA file to avoid stressing the C compiler. It won't however, break a large routine into smaller routines.

Post bug reports on EUforum.

In particular, report any program that does not run the same when compiled as it does when interpreted.

6.4 Indirect routine calling

EUPHORIA does not have function pointers. However, it enables you to call any routine, including some internal to the interpreter, in an indirect way, using two different sets of identifiers.

6.4.1 Indirect calling a routine coded in EUPHORIA

The following applies to any routine coded in EUPHORIA that your program uses, whether it is defined in the standard library, any third party library or your own code. It does not apply to routines implemented in the backend.

6.4.1.1 Getting a routine identifier

Whenever a routine is in scope, you can supply its name to the builtin `routine_id()` function, which returns a small integer:

```
include get.e
constant value_id = routine_id("value")
```

Because `value()` is defined as public, that routine is in scope. This ensures the call succeeds. A failed call returns -1, else a small nonnegative integer.

You can then feed this integer to `call_func()` or `call_proc()` as appropriate. It doesn't matter whether the routine is still in scope at the time you make that call. Once the id is gotten, it's valid.

6.4.1.2 Calling EUPHORIA routines by id.

This is very similar to using `c_func()` or `c_proc()` to interface with external code.

6.4.1.2.1 Calling a function

This is done as follows:

```
result = call_func(id_of_the_routine, argument_sequence)
```

where

- `id_of_the_routine` is an id you obtained from `routine_id()`.
- `argument_sequence` is the list of the parameters to pass, enclosed into curly braces.

```
include get.e

constant value_id = routine_id("value")
result = call_func(value_id, {"Model 36A", 6, GET_LONG_ANSWER})
-- result is {GET_SUCCESS, 36, 4, 1}
```

This is equivalent to

```
result = value("Model 36A", 6, GET_LONG_ANSWER)
```

6.4.1.2.2 Calling a procedure

The same formalism applies, but using `call_proc()` instead. The differences are almost the same as between `c_func()` and `c_proc()`.

```
include std/pretty.e

constant pretty_id = routine_id("pretty_print")

call_proc(pretty_id, {1, some_object, some_options})
```

This does the same as a straightforward

```
include std/pretty.e

pretty_print(1, some_object, some_options)
```

The difference with `c_proc()` is that you can call an external function using `c_proc()` and thus ignore its return value, like in C. Note that you cannot use `call_proc()` to invoke a EUPHORIA function, only C functions.

6.4.1.3 Why call indirectly?

The above examples do not look quite convincing, because it seems that calling indirectly is a more complicated way to call routines. And slower.

However, when the name of the routine you want to call might not be known until run-time, the indirect method will solve the issue for you.

```
integer foo_id

function bar(integer x)
    return call_func(foo_id, {x})
end function

function foo_dev1(integer y)
    yreturn
end function

function foo_dev2(integer y)
    yreturn
end function

function foo_dev3(integer y)
    yreturn 3
end function

function user_opt(object x)
    ...
end function

-- Initialize foo ID
switch user_opt("dev") do
    1 then
        routinefoo_id(foo_dev1)
```



```
2 these
  routine_f0d(f0d00_dev2")
  case else
    routine_f0d(f0d00_dev3")
end switch
```

One last word: when calling a routine indirectly, its **full** parameter list must be passed, even if some of its parameters are defaulted. This limitation may be overcome in future versions.

6.4.2 Calling EUPHORIA's internals

A number of EUPHORIA routines are defined in different ways depending on the platform they will run on. It would be cumbersome, and at times downright impossible, to put such code in include files, nor to make the routine fully builtin.

A response to this is provided by `machine_func()` and `machine_proc()`. User code normally doesn't ever need to use these. Various examples are to be found in the standard library.

These primitives are called like this:

```
machine_proc(id, argument)
result = machine_func(id, argument)
```

`argument` is either an atom, or a sequence standing for one or more parameters. Since the first parameter doesn't need to be a constant, you may use some sort of dynamic calling. The circumstances where it is useful are rare.

The complete list of known values for `id` is to be found in the file `source/execute.h`.

Defining new identifiers and overriding `machine_func/proc()` to handle them is an easy way to extend the capabilities of the interpreter.

6.5 Multitasking in EUPHORIA

6.5.1 Introduction

EUPHORIA allows you to set up multiple, independent tasks. Each task has its own current statement that it is executing, its own call stack, and its own set of private variables. Tasks run in parallel with each other. That is, before any given task completes its work, other tasks can be given a chance to execute. EUPHORIA's task scheduler decides which task should be active at any given time.

6.5.2 Why Multitask?

Most programs do not need to use multitasking and would not benefit from it. However it is very useful in some cases:

- Action games where numerous characters, projectiles etc. need to be displayed in a realistic way, as if they are all independent of one another. Language War is a good example.
- Situations where your program must sometimes wait for input from a human or other computer. While one task in your program is waiting, another separate task could be doing some computation, disk search, etc.
- Windows, Linux and FreeBSD all have special API routines that let you initiate some I/O, and then proceed without waiting for it to finish. A task could check periodically to see if the I/O is finished, while another task is performing some useful computation, or is perhaps starting another I/O operation.
- Situations where your program might be called upon to serve many users simultaneously. With multiple tasks, it's easy to keep track of the state of your interaction with all these separate users.
- Perhaps you can divide your program into two logical processes, and have a task for each. One produces data and stores it, while the other reads the data and processes it. Maybe the first process is time-critical, since it interacts with the user, while the second process can be executed during lulls in the action, where the user is thinking or doing something that doesn't require quick response.

6.5.3 Types of Tasks

EUPHORIA supports two types of tasks: real-time tasks, and time-share tasks.

Real-time tasks are scheduled at intervals, specified by a number of seconds or fractions of a second. You might schedule one real-time task to be activated every 3 seconds, while another is activated every 0.1 seconds. In Language War, when the EUPHORIA ship moves at warp 4, or a torpedo flies across the screen, it's important that they move at a steady, timed pace.

Time-share tasks need a share of the CPU but they needn't be rigidly scheduled according to any clock.

It's possible to reschedule a task at any time, changing its timing or its slice of the CPU. You can even convert a task from one type to the other dynamically.

6.5.4 A Small Example

This example shows the main task (which all EUPHORIA programs start off with) creating two additional real-time tasks. We call them real-time because they are scheduled to get control every few seconds.

You should try copy/pasting and running this example. You'll see that task 1 gets control every 2.5 to 3 seconds, while task 2 gets control every 5 to 5.1 seconds. In between, the main task (task 0), has control as it checks for a 'q' character to abort execution.

```
constant TRUE = 1, FALSE = 0

type boolean(integer x)
    xret0r0r x = 1
end type

boolean t1_running, t2_running

procedure task1(sequence message)
```

```

        i = foto 10 do
            (1, "task1pridf%s\n", {i, message})
            () task_yield
        end for
        t1_running = FALSE
end procedure

procedure task2(sequence message)
    i = foto 10 do
        (1, "task2pridf%s\n", {i, message})
        () task_yield
    end for
    t2_running = FALSE
end procedure

puts(1, "main task: start\n")

atom t1, t2

t1 = task_create(routine_id("task1"), {"Hello"})
t2 = task_create(routine_id("task2"), {"Goodbye"})

task_schedule(t1, {2.5, 3})
task_schedule(t2, {5, 5.1})

t1_running = TRUE
t2_running = TRUE

while t1_running or t2_running do
    get_k%f() = 'q' then
        exit
    end if
    task_yield
end while

puts(1, "main task: stop\n")
-- program ends when main task is finished

```

6.5.5 Comparison with earlier multitasking schemes

In earlier releases of EUPHORIA, Language War already had a mechanism for multitasking, and some people submitted to User Contributions their own multitasking schemes. These were all implemented using plain EUPHORIA code, whereas this new multitasking feature is built into the interpreter. Under the old Language War tasking scheme a scheduler would **call** a task, which would eventually have to **return** to the scheduler, so it could then dispatch the next task.

In the new system, a task can call the built-in procedure `task_yield()` at any point, perhaps many levels deep in subroutine calls, and the scheduler, which is now part of the interpreter, will be able to transfer control to any other task. When control comes back to the original task, it will resume execution at the statement after `task_yield()`, with its call stack and all private variables intact. Each task has its own call stack, program counter (i.e. current statement being executed), and private variables. You might have several tasks all executing a routine at the same time, and each task will have its own set of private variable values for that routine. Global and local variables are shared between tasks.

It's fairly easy to take any piece of code and run it as a task. Just insert a few `task_yield()` statements so it won't hog the CPU.

6.5.6 Comparison with multithreading

When people talk about threads, they are usually referring to a mechanism provided by the operating system. That's why we prefer to use the term "multitasking". Threads are generally "preemptive", whereas EUPHORIA multitasking is "cooperative". With preemptive threads, the operating system can force a switch from one thread to another at virtually any time. With cooperative multitasking, each task decides when to give up the CPU and let another task get control. If a task were "greedy" it could keep the CPU for itself for long intervals. However since a program is written by one person or group that wants the program to behave well, it would be silly for them to favor one task like that. They will try to balance things in a way that works well for the user. An operating system might be running many threads, and many programs, that were written by different people, and it would be useful to enforce a reasonable degree of sharing on these programs. Preemption makes sense across the whole operating system. It makes far less sense within one program.

Furthermore, threading is notorious for causing subtle bugs. Nasty things can happen when a task loses control at just the wrong moment. It may have been updating a global variable when it loses control and leaves that variable in an inconsistent state. Something as trivial as incrementing a variable can go awry if a thread-switch happens at the wrong moment. e.g. consider two threads. One has:

```
x = x + 1
```

and the other also has:

```
x = x + 1
```

At the machine level, the first task loads the value of `x` into a register, then loses control to the second task which increments `x` and stores the result back into `x` in memory. Eventually control goes back to the first task which also increments `x` *using the value of `x` in the register*, and then stores it into `x` in memory. So `x` has only been incremented once instead of twice as was intended. To avoid this problem, each thread would need something like:

```
lock x
x = x + 1
unlock x
```

where `lock` and `unlock` would be special primitives that are safe for threading. It's often the case that programmers forget to lock data, but their program seems to run ok. Then one day, many months after they've written the code, the program crashes mysteriously.

Cooperative multitasking is much safer, and requires far fewer expensive locking operations. Tasks relinquish control at safe points once they have completed a logical operation.

6.5.7 Summary

For a complete function reference, refer to the Library Documentation [Multitasking](#).

6.6 EUPHORIA Database System (EDS)

6.6.1 Introduction

Many people have expressed an interest in accessing databases using EUPHORIA programs. Those people have either wanted to access a name-brand database management system from EUPHORIA, or they've wanted a simple, easy-to-use, EUPHORIA-oriented database for storing data. EDS is the latter. **It provides a simple, extremely flexible, database system for use by EUPHORIA programs.**

6.6.2 Structure of an EDS database

In EDS, a **database** is a single file with `.edb` file type. An EDS database contains 0 or more **tables**. Each table has a **name**, and contains 0 or more **records**. Each record consists of a **key** part, and a **data** part. The key can be **any** EUPHORIA object - an atom, a sequence, a deeply-nested sequence, whatever. Similarly the data can be **any** EUPHORIA object. There are **no** constraints on the size or structure of the key or data. Within a given table, the keys are all unique. That is, no two records in the same table can have the same key part.

The records of a table are stored in ascending order of key value. An efficient binary search is used when you refer to a record by key. You can also access a record directly, with no search, if you know its current **record number** within the table. Record numbers are integers from 1 to the length (current number of records) of the table. By incrementing the record number, you can efficiently step through all the records, in order of key. Note however that a record's number can change whenever a new record is inserted, or an existing record is deleted.

The keys and data parts are stored in a compact form, but **no** accuracy is lost when saving or restoring floating-point numbers or **any** other EUPHORIA data.

`database.e` will work as is, on all platforms. EDS database files can be copied and shared between programs running on all platforms as well. Be sure to make an exact byte for byte copy using "binary" mode copying, rather than "text" or "ASCII" mode which could change the line terminators.

Example:

```
database: "mydata.edb"
  first table: "passwords"
    record #1: key: "jones"    data: "euphor123"
    record #2: key: "smith"   data: "billgates"

    second table: "parts"
      record #1: key: 134525    data: {"hammer", 15.95, 500}
      record #2: key: 134526    data: {"saw", 25.95, 100}
```

```
record #3:  key: 134530    data: {"screw driver", 5.50, 1500}
```

It's up to you to interpret the meaning of the key and data. **In keeping with the spirit of EUPHORIA, you have total flexibility.** Unlike most other database systems, an EDS record is *not* required to have either a fixed number of fields, or fields with a preset maximum length.

In many cases there will not be any natural key value for your records. In those cases you should simply create a meaningless, but unique, integer to be the key. Remember that you can always access the data by record number. It's easy to loop through the records looking for a particular field value.

6.6.3 How to access the data

To reduce the number of parameters that you have to pass, there is a notion of the **current database**, and **current table**.

6.6.3.1 The current database.

Any data operation or table operation assumes there is a current database being defined. You set the current database by opening, creating or selecting a database. Deleting the current database leaves the current database undefined.

6.6.3.2 The current table.

All data operations assume there is a current table being defined. You must create, select or rename a table in order to make it current. Deleting the current table leaves the current table undefined.

6.6.3.3 Accessing data

Most routines use these **current** values automatically. You normally start by opening (or creating) a database file, then selecting the table that you want to work with.

You can map a key to a record number using [db_find_key](#). It uses an efficient binary search. Most of the other record-level routines expect the record number as a parameter. You can very quickly access any record, given it's number. You can access all the records by starting at record number 1 and looping through to the record number returned by [db_table_size](#).

6.6.4 How does storage get recycled?

When you delete something, such as a record, the space for that item gets put on a free list, for future use. Adjacent free areas are combined into larger free areas. When more space is needed, and no suitable space is found on the free list, the file will grow in size. Currently there is no automatic way that a file will shrink in size, but you can use a [db_compress](#) to completely rewrite a database, removing the unused spaces.

6.6.5 Security / Multi-user Access

This release provides a simple way to lock an entire database to prevent unsafe access by other processes.

6.6.6 Scalability

Internal pointers are 4 bytes. In theory that limits the size of a database file to 4 Gb. In practice, the limit is 2 Gb because of limitations in various C file functions used by EUPHORIA. Given enough user demand, EDS databases could be expanded well beyond 2 Gb in the future.

The current algorithm allocates 4 bytes of memory per record in the current table. So you'll need at least 4Mb RAM per million records on disk.

The binary search for keys should work reasonably well for large tables.

Inserts and deletes take slightly longer as a table gets larger.

At the low end of the scale, it's possible to create extremely small databases without incurring much disk space overhead.

6.6.7 Disclaimer

Do not store valuable data without a backup. RDS will not be responsible for any damage or data loss.

6.6.8 Warning: Use the right file mode

.edb files are binary files, not text files. You **must** use `BINARY` mode when transferring a .edb file via FTP from one machine to another. You must also avoid loading a .edb file into an editor and saving it. If you open a .edb file directly using EUPHORIA's `open()`, which is not recommended, you must use binary mode, not text mode. Failure to follow these rules could result in 10 (line-feed) and 13 (carriage-return) bytes being changed, leading to subtle and not-so-subtle forms of corruption in your database.

6.7 The User Defined Pre-Processor

Introduced in 4.0 beta 1, the user defined pre-processor opens a world of possibilities to the EUPHORIA programmer. In a sentence, it allows one to create (or use) a translation process that occurs transparently when a program is run. This mini-guide is going to explore the pre-processor interface by first giving a quick example, then explaining it in detail and finally by writing a few useful pre-processors that can be put immediately to work.

Any program can be used as a pre-processor. It must, however, adhere to a simple specification:

1. Accept a parameter "-i filename" which specifies which file to read and process.

2. Accept a parameter "-o filename" which specifies which file to write the result to.
3. Exit with a zero error code on success or a non-zero error code on failure.

It does not matter what type of program it is. It can be a EUPHORIA script, an executable written in the C programming language, a script/batch file or anything else that can read one file and write to another file. As EUPHORIA programmers, however, we are going to focus on writing pre-processors in the EUPHORIA programming language. As a benefit, we will describe later on how you can easily convert your pre-processor to a shared library that EUPHORIA can make use of directly thus improving performance.

6.7.1 A Quick Example

The problem in this case is that you want the copyright statement and the about screen to show what date the program was compiled on but you do not want to manually maintain this date. So, we are going to create a simple pre-processor that will read a source file, replace all instances of @DATE@ with the current date and then write the output back out.

Before we get started, let me say that we will expand on this example later on. Up front, we are going to do almost no error checking for the purpose of showing off the pre-processor not for the sake of making a production quality application.

We are going to name this file `datesub.ex`.

```
-- datesub.ex
include std/datetime.e -- now() and format()
include std/io.e       -- read_file() and write_file()
include std/search.e   -- find_replace()

sequence cmds = command_line()
sequence inFileName, outFileName

for i = 3 to length(cmds) do
    switch cmds[i] do
        case "-i" then
            inFileName = cmds[i+1]
        case "-o" then
            outFileName = cmds[i+1]
        end switch
    end for

sequence content = read_file(inFileName)

content = find_replace("@DATE@", content, format(now()))

write_file(outFileName, content)

-- programs automatically exit with ZERO error code, if you want
-- non-zero, you exit with abort(1), for example.
```

So, that is our pre-processor. Now, how do we make use of it? First let's create a simple test program that we can watch it work with. Name this file `thedata.ex`.

```
-- thedate.ex
```



```
puts(1, "The date this was run is @DATE@\n")
```

Rather simple, but it shows off the pre-processor we have created. Now, let's run it, but first without a pre-processor hook defined.

NOTE: Through this document I am going to assume that you are working in Windows. If not, you can make the appropriate changes to the shell type examples.

```
C:\MyProjects\datesub> eui thedate.ex
The date this was run is @DATE@
```

Not very helpful? Ok, let's tell EUPHORIA how to use the pre-processor that we just created and then see what happens.

```
C:\MyProjects\datesub> eui -p eui:datesub.ex thedate.ex
The date this was run is 2009-08-05 19:36:22
```

If you got something similar to the above output, good job, it worked! If not, go back up and check your code for syntax errors or differences from the examples above.

What is this -p paramater? In short, -p tells eui or euc that there is a pre-processor. The definition of the pre-processor comes next and can be broken into 2 required sections and 1 optional section. Each section is divided by a colon (:).

For example, -p e,ex:datesub.ex

1. e, ex tells EUPHORIA that when it comes across a file with the extension e or ex that it should run a pre-processor
2. datesub.ex tells EUPHORIA which pre-processor should be run. This can be a .ex file or any other executable command.
3. An optional section exists to pass options to the pre-processor but we will go into this later.

That's it for the quick introduction. I hope that the wheels are turning in your head already as to what can be accomplished with such a system. If you are interested, please continue reading and see where things will get very interesting!

6.7.2 Pre-process Details

EUPHORIA manages when the pre-processor should be called and with what arguments. The pre-processor does not need to concern itself as to if it should run, what filename it is reading or what filename it will be writing to. It should simply do as EUPHORIA tells it to do. This is because EUPHORIA monitors what the modification time is on the source file and what time the last pre-process call was made on the file. If nothing has changed in the source file then the pre-processor is not called again. Pre-processing does have a slight penalty in speed as the file is processed twice. For example, the datesub.ex pre-processor read the entire file, searched for @DATE@, wrote the file and then EUPHORIA picked up from there reading the output file, parsing it and finally executing it. To minimize the time taken, EUPHORIA caches the output of the pre-processor so that the interim process is not normally needed after it has been run once.

6.7.3 Command Line Options

6.7.3.1 -p - Define a pre-processor

The primary command line option that you will use is the `-p` option which defines the pre-processor. It is a two or three section option. The first section is a comma delimited list of file extensions to associate with the pre-processor, the second is the actual pre-processor script/command and the optional third is parameters to send to the pre-processor in addition to the `-i` and `-o` parameters.

Let's go over some examples:

- `-p e:datesub.ex` - This will be executed for every `.e` file and the command to call is `datesub.ex`.
- `-p "de,dex,dew:dot4.dll:-verbose -no-dbc"` - Files with `de`, `dex`, `dew` extensions will be passed to the `dot4.dll` process. `dot4.dll` will get the optional parameters `-verbose -no-dbc` passed to it.

Multiple pre-processors can be defined at the same time. For instance,

```
C:\MyProjects\datesub> eui -p e,ex:datesub.ex -p de,dex:dot4.dll \  
-p le,lex:lex.literate.ex hello.ex
```

is a valid command line. It's possible that `hello.ex` may include a file named `greeter.le` and that file may include a file named `person.de`. Thus, all three pre-processors will be called upon even though the main file is only processed by `datesub.ex`

6.7.3.2 -pf - Force pre-processing

When writing a pre-processor you may run into the problem that your source file did not change, therefore, EUPHORIA is not calling your pre-processor. However, your pre-processor has changed and you want EUPHORIA to re-process your unchanged source file. This is where `-pf` comes into play. `-pf` causes EUPHORIA to force the pre-processing, regardless of the cached state of any file. When used, EUPHORIA will always call the pre-processor for all files with a matching pre-processor definition.

6.7.3.3 Use of a configuration file

Ok, so who wants to type these pre-processor definitions in all the time? I don't either. That's where the standard EUPHORIA configuration file comes into play. You can simply create a file named `eu.cfg` and place something like this into it.

```
-p le,lex:lex.literate.ex  
-p e,ex:datesub.ex  
... etc ...
```

Then you can execute any of those files directly without the `-p` parameters on the command line. This `eu.cfg` file can be local to a project, local to a user or global on a system. Please read about the `[[eu.cfg]`

file for more information.

6.7.4 DLL/Shared Library Interface

A pre-processor may be a EUPHORIA file, ending with an extension of `.ex`, a compiled EUPHORIA program, `.exe` or even a compiled EUPHORIA DLL file, `.dll`. The only requirements are that it must accept the two command line options, `-i` and `-o` described above and exit with a ZERO status code on success or non-ZERO on failure.

The DLL file (or shared library on Unixes) has a real benefit in that with each file that needs to be pre-processed does not require a new process to be spawned as with an executable or a EUPHORIA script. Once you have the pre-processor written and functioning, it's easy to convert your script to use the more advanced, better performing shared library method. Let's do that now with our `datesub.ex` pre-processor. Take a moment to review the code above for the `datesub.ex` program before continuing. This will allow you to more easily see the changes that we make here.

```
-- datesub.ex
include std/datetime.e -- now() and format()
include std/io.e       -- read_file() and write_file()
include std/search.e   -- find_replace()

public function preprocess(sequence inFileName, sequence outFileName,
                           sequence options={})

    sequence content = read_file(inFileName)

    content = find_replace("@DATE@", content, format(now()))

    write_file(outFileName, content)

    return 0
end function

ifdef not EUC_DLL then
    sequence cmds = command_line()
    sequence inFileName, outFileName

    for i = 3 to length(cmds) do
        switch cmds[i] do
            case "-i" then
                inFileName = cmds[i+1]
            case "-o" then
                outFileName = cmds[i+1]
            end switch
        end for

        preprocess(inFileName, outFileName)
    end ifdef
```

It's beginning to look a little more like a well structured program. You'll notice that we took the actual pre-processing functionality out the the top level program making it into an exported function named `preprocess`. That function takes three parameters:

1. `inFileName` - filename to read from
2. `outFileName` - filename to write to
3. `options` - options that the user may wish to pass on verbatim to the pre-processor

It should return 0 on no error and non-zero on an error. This is to keep a standard with the way error levels from executables function. In that convention, it's suggested that 0 be OK and 1, 2, 3, etc... indicate different types of error conditions. Although the function could return a negative number, the main routine cannot exit with a negative number.

To use this new process, we simply translate it through `euc`,

```
C:\MyProjects\datesub> euc -dll datesub.ex
```

If all went correctly, you now have a `datesub.dll` file. I'm sure you can guess on how it should be used, but for the sake of being complete,

```
C:\MyProjects\datesub> eui -p e,ex:datesub.dll thedate.ex
```

On such a simple file and such a simple pre-processor, you probably are not going to notice a speed difference but as things grow and as the pre-processor gets more complicated, compiling to a shared library is your best option.

6.7.5 Advanced Examples

6.7.5.1 Finish `datesub.ex`

Before we move totally away from our `datesub.ex` example, let's finish it off by adding some finishing touches and making use of optional parameters. Again, please go back and look at the Shared Library version of `datesub.ex` before continuing so that you can see how we have changed things.

```
-- datesub.ex
include std/cmdline.e -- command line parsing
include std/datetime.e -- now() and format()
include std/io.e      -- read_file() and write_file()
include std/map.e     -- map accessor functions (get())
include std/search.e  -- find_replace()

sequence cmdopts = {
    { "f", 0, "Date format", { NO_CASE, HAS_PARAMETER, "format" } }
}

public function preprocess(sequence inFileName, sequence outFileName,
    sequence options={})
    map opts = cmd_parse(cmdopts, options)
    sequence content = read_file(inFileName)

    content = find_replace("@DATE@", content, format(now(), map:get(opts,
    "f")))

    write_file(outFileName, content)
```

```

    return 0
end function

ifndef not EUC_DLL then
    cmdopts = {
        { "i", 0, "Input filename", { NO_CASE, MANDATORY, HAS_PARAMETER,
"filename" } },
        { "o", 0, "Output filename", { NO_CASE, MANDATORY, HAS_PARAMETER,
"filename" } }
    } & cmdopts

    map opts = cmd_parse(cmdopts)
    preprocess(map:get(opts, "i"), map:get(opts, "o"),
        "-f " & map:get(opts, "f", "%Y-%m-%d"))
end ifdef

```

Here we simply used `cmdline.e` to handle the command line parsing for us giving out command line program a nice interface, such as parameter validation and an automatic help screen. At the same time we also added a parameter for the date format to use. This is optional and if not supplied, `%Y-%m-%d` is used.

The final version of `datesub.ex` and `thedata.ex` are located in the `demo/preproc` directory of your EUPHORIA installation.

6.7.5.2 Others

TODO: this needs done still.

EUPHORIA includes two more demos of pre-processors. They are ETML and `literate`. Please explore `demo/preproc` for these examples and explanations.

6.7.5.2.1 Other examples of pre-processors include:

- `eSQL` - Allows you to include a `.sql` file directly. It parses `CREATE TABLE` and `CREATE INDEX` statements building common methods to create, destroy, get by id, find by any index, add, remove and save entities.
- `make40` - Will process any 3.x script on the fly making sure that it will run in 4.x. It does this by converting variables, constants and routine names that are the same as new 4.x keywords into something acceptable to 4.x. Thus, 3.x programs can run in the 4.x interpreter and translator with out any user intervention.
- `dot4` - Adds all sorts of syntax goodies to EUPHORIA such as structured sequence access, one line if statements, DOT notation for any function/routine call, design by contract and more.

6.7.5.2.2 Other Ideas

- Include a Windows `.RC` file that defines a dialog layout and generate code that will create the dialog and interact with it.
- Object Oriented system for EUPHORIA that translates into pure EUPHORIA code, thus has the raw speed of EUPHORIA.

- Include a Yacc, Lex, ANTLR parser definition directly that then generates a EUPHORIA parser for the given syntax.
- Instead of writing interpreters such as a QBasic clone, simply write a pre-processor that converts QBasic code into EUPHORIA code, thus you can run `eui -p bas:qbasic.ex hello.bas` directly.
- Include a XML specification, which in turn, gives you nice accessory functions for working with XML files matching that schema.

If you have ideas of helpful pre-processors, please put the idea out on the forum for discussion.

6.8 EUPHORIA Trouble-Shooting Guide

If you get stuck, here are some things you can do:

1. Type: `guru` followed by some keywords associated with your problem. For example, `guru declare global include`
2. Check the list of common problems ([Common Problems and Solutions](#)).
3. Read the relevant parts of the documentation, i.e. [EUPHORIA Programming Language v4.0](#) or [General Routine Reference](#).
4. Try running your program with trace:

```
with trace
trace(1)
```

1. The [EUPHORIA Mailing List](#) has a search facility. You can search the archive of all previous messages. There's a good chance that your question has already been discussed.
2. Post a message on the mailing list.
3. Visit the EUPHORIA IRC channel, <irc://irc.freenode.net/#euphoria>.

6.8.1 Common Problems and Solutions

Here are some commonly reported problems and their solutions.

6.8.1.1 Console window disappeared

I ran my program with `euiw.exe` and the console window disappeared before I could read the output.

The console window will only appear if required, and will disappear immediately when your program finishes execution. Perhaps you should code something like:

```
puts(1, "\nPress Enter\n")
if getc(0) then
end if
```

at the end of your program.

You may also run your console program with `eui.exe`.

6.8.1.2 Press Enter

At the end of execution of my program, I see "Press Enter" and I have to hit the Enter key. How do I get rid of that?

Call `free_console` just before your program terminates.

```
include dll.e  
  
free_console()
```

6.8.1.3 CGI Program Hangs/No Output

My EUPHORIA CGI program hangs or has no output

1. Make sure that you are using the `-batch` parameter to `eui`. This causes EUPHORIA to not present the normal "Press any key to continue..." prompt when a warning or error occurs. The web server will not respond to this prompt and your application will hang waiting for ENTER to be pressed.
2. Use the `-wf` parameter to write all warnings to a file instead of the console. The warnings that EUPHORIA will write to the console may interfere with the actual output of your web content.
3. Look for an `ex.err` file in your `cgi-bin` directory. Turn on with `trace/trace(3)` to see what statements are executed (see `ctrace.out` in your `cgi-bin`). On Windows you should always use `eui.exe` to run CGI programs, or you may have problems with standard output. With Apache Web Server, you can have a first line in your program of:
4. `!.eui.exe` to run your program using `eui.exe` in the current (`cgi-bin`) directory. Be careful that your first line ends with the line breaking characters appropriate for your platform, or the `#!` won't be handled correctly. You must also set the execute permissions on your program correctly, and `ex.err` and `ctrace.out` must be writeable by the server process or they won't be updated.

6.8.1.4 Read/Write Ports?

How do I read/write ports?

There are collections of machine-level routines from the [EUPHORIA Web Page](#).

6.8.1.5 Program has no errors, no output

When I run my program there are no errors but nothing happens.

You probably forgot to call your main procedure. You need a top-level statement that comes after your main procedure to call the main procedure and start execution.

6.8.1.6 Routine not declared

I'm trying to call a routine documented in `library.doc`, but it keeps saying the routine has not been declared.

Did you remember to include the necessary `.e` file from the `euphoria\include` directory? If the syntax of the routine says for example, `"include graphics.e"`, then your program must have `"include graphics.e"` (without the quotes) before the place where you first call the routine.

6.8.1.7 Routine not declared, my file

*I have an include file with a routine in it that I want to call, but when I try to call the routine it says the routine has not been declared. But it **has** been declared.*

Did you remember to define the routine as `public`, `export` or possibly `global`? If not, the routine is not visible outside of its own file.

6.8.1.8 After user input, left margin problem

After inputting a string from the user with `gets()`, the next line that comes out on the screen does not start at the left margin.

Your program should output a *new-line* character e.g.

```
input = gets()
puts(SCREEN, '\n')
```

6.8.1.9 Floating-point calculations not exact

Why aren't my floating-point calculations coming out exact?

Intel CPU's, and most other CPU's, use binary numbers to represent fractions. Decimal fractions such as 0.1, 0.01 and similar numbers can't be represented precisely. For example, 0.1 might be stored internally as 0.09999999999999999. That means that `10 * 0.1` might come out as 0.9999999999999999, and `floor(10 * 0.1)` might be 0, not 1 as you'd expect. This can be a nuisance when you are dealing with money calculations, but it's not a EUPHORIA problem. It's a general problem that you must face in most programming languages. Always remember: floating-point numbers are just an approximation to the "real" numbers in mathematics. Assume that any floating-point calculation might have a tiny bit of error in the result. Sometimes you can solve the problem by rounding, e.g. `x = round(x, 100)` would round `x` off to the nearest hundredth. Storing money values as an integer number of pennies, rather than a fractional number of dollars (or similar currency) will help, but some calculations could still cause problems.

6.8.1.10 Number to a string?

How do I convert a number to a string?

Use `sprintf`:

```
string = eu:sprintf("%d", 10) -- string is "10"
```

or use `number`:

```
include std/locale.e as locale
string = locale:number(10) -- string is probably "10.00" if called in the U.S.
                           -- It depends on the locale preferences set on your computer.
```

Number formats according to the locale setting on your computer and strangely, this means to give you two decimal places whether or not you supply an integer value for the U.S. locale.

Besides `%d`, you can also try other formats, such as `%x` (Hex) or `%f` (floating-point).

6.8.1.11 String to a number?

How do I convert a string to a number?

Use `value`.

6.8.1.12 Redefine my for-loop variable?

It says I'm attempting to redefine my for-loop variable.

For-loop variables are declared automatically. Apparently you already have a declaration with the same name earlier in your routine or your program. Remove that earlier declaration or change the name of your loop variable.

6.8.1.13 Unknown Escape Character

I get the message "unknown escape character" on a line where I am trying to specify a file name.

Do not say `"C:\TMP\MYFILE"`. You need to say `"C:\\TMP\\MYFILE"`.

Backslash is used for escape characters such as `\n` or `\t`. To specify a single backslash in a string you need to type `\\`. Therefore, say `"C:\\TMP\\MYFILE"` instead of `"C:\TMP\MYFILE"`

6.8.1.14 Only first character in printf

I'm trying to print a string using `printf` but only the first character comes out.

You need to put braces around the parameters sequence to `printf`. You probably wrote:

```
printf(1, "Hello, %s!\n", mystring)
```

but you need:

```
printf(1, "Hello, %s!\n", {mystring})
```

6.8.1.15 Only 10 significant digits during printing

When I print numbers using `printf` or `?` only 10 significant digits are displayed.

EUPHORIA normally only shows about 10 digits. Internally, all calculations are performed using at least 15 significant digits. To see more digits you have to use `printf`. For example,

```
printf(1, "%.15f", 1/3)
```

This will display 15 digits.

6.8.1.16 A type is expected here

It complains about my routine declaration, saying, "a type is expected here."

When declaring subroutine parameters, EUPHORIA requires you to provide an explicit type for each individual parameter. e.g.

```
procedure foo(integer x, y)      -- WRONG
procedure foo(integer x, integer y) -- RIGHT
```

In all other contexts it is ok to make a list:

```
atom a, b, c, d, e
```

6.8.1.17 Expected to see...

It says: Syntax Error - expected to see possibly 'xxx', not 'yyy'

At this point in your program you have typed a variable, keyword, number or punctuation symbol, yyy, that does not fit syntactically with what has come before it. The compiler is offering you one example, xxx, of something that would be accepted at this point in place of yyy. Note that there may be many other legal (and much better) possibilities at this point than xxx, but xxx might at least give you a clue as to what the compiler is "thinking."

6.9 Platform Specific Issues

6.9.1 Introduction

OpenEUPHORIA currently supports EUPHORIA on many different *platforms*. More platforms will be added in the future.

Windows ^(tm) in particular, the 32-bit version of Windows that is used on Windows 95/98/ME, as well as NT/2000/XP and later systems.

Linux. Linux is based on the UNIX operating system. It has recently become very popular on PCs. There are many distributors of Linux, including Red Hat, Debian, Caldera, etc. Linux can be obtained on a CD for a very low price. Linux is an open-source operating system.

FreeBSD. FreeBSD is also based on the UNIX operating system. It is very popular on Internet server machines. It's also open source.

Apple's **OS X.** OS X is also based on the UNIX operating system. While it is closed source, it is gaining a wide following due to it's ease of use and power.

Sun OS. Sun OS was developed by and is owned by SUN Microsystems Inc. It is also based on UNIX.

Open BSD. Open BSD is also a UNIX-like Operating System and is developed by volunteers.

Net BSD. Net BSD is also a UNIX-like Operating System and is designed to be easily portable to other hardware platforms.

EUPHORIA source files use various file extensions. The common extensions are:

- .e EUPHORIA include file
- .ew EUPHORIA include file for a Windowed (GUI) application only
- .ex Console main program file
- .exw Windowed (GUI) main program file

The EUPHORIA for Windows installation file contains **eui.exe**. It runs EUPHORIA programs on the Windows 32bit platform.

The EUPHORIA for Linux .tar file contains only **eui**. It runs EUPHORIA programs on the Linux platform.

The FreeBSD version of EUPHORIA is installed by first installing the Linux version of EUPHORIA, and then replacing eui, by the version of eui for FreeBSD.

Sometimes you'll find that the majority of your code will be the same on all platforms, but some small parts will have to be written differently for each platform. Use the [platform](#) built-in function to tell you which platform you are currently running on. Note that platform() returns the same value (3) on both Linux and FreeBSD, since those systems are so similar. OS X returns (4) as it is different enough to warrant it's own identifier.

6.9.2 The WIN32 Platform

With WIN32 you also have access to all of the memory on your machine. Most library routines work the same way on each platform. Many existing text mode programs written for MS-DOS can be run using **eui** without any change. With **eui** you can run programs from the command line, and display text on a standard (typically 25 line x 80 column) DOS window. The DOS window is known as the *console* in Windows terminology. EUPHORIA makes the transition from DOS32 text mode programming, to WIN32 console programming, trivial.

You can add calls to WIN32 C functions and later, if desired, you can create real Windows GUI windows.

A console window will be created automatically when a WIN32 EUPHORIA program first outputs something to the screen or reads from the keyboard. You will also see a console window when you read standard input or write to standard output, even when these have been redirected to files. The console will disappear when your program finishes execution, or via a call to [free_console](#). If there is something on the console that you want your user to read, you should prompt him and wait for his input before terminating. To prevent the console from quickly disappearing you might include a statement such as:

```
if getc(0) then
end if
```

which will wait for the user enter something.

If you want to run EUPHORIA programs without popping up a new console window use `eui.exe` otherwise use `euiw.exe`. `eui.exe` uses the current console window.

Under WIN32, long filenames are fully supported for reading and writing and creating.

6.9.2.1 High-Level WIN32 Programming

Thanks to **David Cuny**, **Derek Parnell**, **Judith Evans** and many others, there's a package called **Win32Lib** that you can use to develop Windows GUI applications in EUPHORIA. It's remarkably easy to learn and use, and comes with good documentation and many small example programs. You can download Win32Lib and Judith's IDE from the [EUPHORIA Web site](#). **Andrea Cini** has also developed a similar, somewhat smaller package called **EuWinGUI**. It's also available from our site.

6.9.2.2 Low-Level WIN32 Programming

To allow access to WIN32 at a lower level, EUPHORIA provides a mechanism for calling any C function in any WIN32 API .dll file, or indeed in any 32-bit Windows .dll file that you create or someone else creates. There is also a call-back mechanism that lets Windows call your EUPHORIA routines. Call-backs are necessary when you create a graphical user interface.

To make full use of the WIN32 platform, you need documentation on 32-bit Windows programming, in particular the WIN32 Application Program Interface (API), including the C structures defined by the API. There is a large WIN32.HLP file (c) Microsoft that is available with many programming tools for Windows.

There are numerous books available on the subject of WIN32 programming for C/C++. You can adapt most of what you find in those books to the world of EUPHORIA programming for WIN32. A good book is *Programming Windows by Charles Petzold*.

A WIN32 API Windows help file (8 Mb) can be downloaded from <ftp://ftp.borland.com/pub/delphi/techpubs/delphi2/win32.zip>, Borland's Web site.

6.9.3 The Unix Platforms

As with WIN32, you can write text on a console, or xterm window, in multiple colors and at any line or column position.

Just as in WIN32, you can call C routines in shared libraries and C code can call back to your EUPHORIA routines.

EUPHORIA for Unix does not have integrated support for pixel graphics, but Pete Eberlein has created a EUPHORIA interface to **svglib**.

Easy X windows GUI programming is available using either Irv Mullin's EuGTK interface to the GTK GUI library, or wxEUPHORIA developed by Matt Lewis. wxEUPHORIA also runs on Windows.

When porting code from Windows to Unix, you'll notice the following differences:

- Some of the numbers assigned to the 16 main colors in graphics.e are different. If you use the constants defined in graphics.e you won't have a problem. If you hard-code your color numbers you will see that blue and red have been switched etc.
- The key codes for special keys such as Home, End, arrow keys are different, and there are some additional differences when you run under XTERM.
- The Enter key is code 10 (line-feed) on Linux, where on Windows it was 13 (carriage-return).
- Linux and FreeBSD use '/' (slash) on file paths. Windows use '\' (backslash).
- Calls to system() and system_exec() that contain Windows commands will obviously have to be changed to the corresponding Linux or FreeBSD command. e.g. "DEL" becomes "rm", and "MOVE" becomes "mv".

6.9.4 Interfacing with C Code (WIN32, Linux, FreeBSD)

On WIN32 and Unix it's possible to interface EUPHORIA code with C code. Your EUPHORIA program can call C routines and read and write C variables. C routines can even call ("callback") your EUPHORIA routines. The C code must reside in a WIN32 dynamic link library (.dll file), a Linux or FreeBSD shared library (.so file) or an OS X shared library (.dylib file). By interfacing with .dll libraries and shared libraries, you can access the full programming interface on these systems.

Using the EUPHORIA to C Translator, you can translate EUPHORIA routines to C, and compile them into a shared library file. You can pass EUPHORIA atoms and sequences to these compiled EUPHORIA routines, and receive EUPHORIA data as a result. Translated/compiled routines typically run much faster than interpreted routines. For more information, see the [Translator](#).

6.9.4.1 Calling C Functions

To call a C function in a shared library file you must perform the following steps:

1. Open the shared library file that contains the C function by calling [open_dll](#).
2. Define the C function, by calling [define_c_func](#) or [define_c_proc](#). This tells EUPHORIA the number and type of the arguments as well as the type of value returned.
EUPHORIA currently supports all C integer and pointer types as arguments and return values. It also supports floating-point arguments and return values (C double type). It is currently not possible to pass C structures by value or receive a structure as a function result, although you can certainly pass a pointer to a structure and get a pointer to a structure as a return value. Passing C structures by value is rarely required for operating system calls.
EUPHORIA also supports all forms of EUPHORIA data - atoms and arbitrarily-complex sequences, as arguments to translated/compiled EUPHORIA routines.
3. Call the C function by calling [c_func](#) or [c_proc](#)

```
include dll.e

atom user32
integer LoadIcon, icon

user32 = open_dll("user32.dll")

-- The name of the routine in user32.dll is "LoadIconA".
-- It takes a pointer and an 32-bit integers as arguments,
-- and it returns a 32-bit integer.
LoadIcon = define_c_func(user32, "LoadIconA", {C_POINTER, C_INT}, C_INT)

icon = c_func(LoadIcon, {NULL, IDI_APPLICATION})
```

See [c_func](#), [c_proc](#), [define_c_func](#), [define_c_proc](#), [open_dll](#)

See **demo\win32** or **demo/linux** for example programs.

On Windows there is more than one C calling convention. The Windows API routines all use the `{{__stdcall}}` convention. Most C compilers however have `__cdecl` as their default. `__cdecl` allows for variable numbers of arguments to be passed. EUPHORIA assumes `__stdcall`, but if you need to call a C routine that uses `__cdecl`, you can put a '+' sign at the start of the routine name in `define_c_proc()` and `define_c_func()`. In the example above, you would have `"+LoadIconA"`, instead of `"LoadIconA"`.

You can examine a `dll` file by right-clicking on it, and choosing "QuickView" (if it's on your system). You will see a list of all the C routines that the `dll` exports.

To find out which `.dll` file contains a particular WIN32 C function, run **euphoria\demo\win32\dsearch.exw**

6.9.4.2 Accessing C Variables

You can get the address of a C variable using [define_c_var](#). You can then use [poke](#) and [peek](#) to access the value of the variable.

6.9.4.3 Accessing C Structures

Many C routines require that you pass pointers to structures. You can simulate C structures using allocated blocks of memory. The address returned by `allocate` can be passed as if it were a C pointer.

You can read and write members of C structures using `peek` and `poke`, or `peek4u`, `peek4s`, and `poke4`. You can allocate space for structures using `allocate`.

You must calculate the offset of a member of a C structure. This is usually easy, because anything in C that needs 4 bytes will be assigned 4 bytes in the structure. Thus C int's, char's, unsigned int's, pointers to anything, etc. will all take 4 bytes. If the C declaration looks like:

```
// Warning C code ahead!

struct example {
    int a;           // offset  0
    char *b;         // offset  4
    char c;          // offset  8
    long d;          // offset 12
};
```

To allocate space for "struct example" you would need:

```
atom p = allocate(16) -- size of "struct example"
```

The address that you get from `allocate` is always at least 4-byte aligned. This is useful, since WIN32 structures are supposed to start on a 4-byte boundary. Fields within a C structure that are 4-bytes or more in size must start on a 4-byte boundary in memory. 2-byte fields must start on a 2-byte boundary. To achieve this you may have to leave small gaps within the structure. In practice it is not hard to align most structures since 90% of the fields are 4-byte pointers or 4-byte integers.

You can set the fields using something like:

```
poke4(p + 0, a)
poke4(p + 4, b)
poke4(p + 8, c)
poke4(p +12, d)
```

You can read a field with something like:

```
d = peek4(p+12)
```

Tip:

For readability, make up EUPHORIA constants for the field offsets. See Example below.

```
constant RECT_LEFT = 0,
RECT_TOP    = 4,
RECT_RIGHT  = 8,
RECT_BOTTOM = 12,
RECT_SIZEOF = 16

atom rect = allocate(RECT_SIZEOF)

poke4(rect + RECT_LEFT, 10)
```



```
poke4(rect + RECT_TOP,      20)
poke4(rect + RECT_RIGHT,    90)
poke4(rect + RECT_BOTTOM, 100)

-- pass rect as a pointer to a C structure
-- hWnd is a "handle" to the window
if not c_func(InvalidateRect, {hWnd, rect, 1}) then
    puts(2, "InvalidateRect failed\n")
end if
```

The EUPHORIA code that accesses C routines and data structures may look a bit ugly, but it will typically form just a small part of your program, especially if you use Win32Lib, EuWinGUI, or Irv Mullin's X Windows library. Most of your program will be written in pure EUPHORIA, which will give you a big advantage over those forced to code in C.

6.9.4.4 Call-backs to your EUPHORIA routines

When you create a window, the Windows operating system will need to call your EUPHORIA routine. To set this up, you must get a 32-bit "call-back" address for your routine and give it to Windows. For example (taken from **demo\win32\window.exw**):

```
integer id
atom WndProcAddress

id = routine_id("WndProc")

WndProcAddress = call_back(id)
```

`routine_id` uniquely identifies a EUPHORIA procedure or function by returning an integer value. This value can be used later to call the routine. You can also use it as an argument to the `call_back` function.

In the example above, The 32-bit *call-back address*, `~WndProcAddress`, can be stored in a C structure and passed to Windows via the `~RegisterClass()` C API function.

This gives Windows the ability to call the EUPHORIA routine, `~WndProc()`, whenever the user performs an action on a certain class of window. Actions include clicking the mouse, typing a key, resizing the window etc.

See the *window.exw* demo program for the whole story.

Note:

It is possible to get a *call-back address* for *any* EUPHORIA routine that meets the following conditions: * the routine must be a function, not a procedure * it must have from 0 to 9 parameters * the parameters should all be of type atom (or integer etc.), not sequence * the return value should be an integer value up to 32-bits in size

You can create as many call-back addresses as you like, but you should not call `call_back` for the same EUPHORIA routine multiple times - each call-back address that you create requires a small block of memory.

The values that are passed to your EUPHORIA routine can be any 32-bit unsigned atoms, i.e. non-negative. Your routine could choose to interpret large positive numbers as negative if that is desirable. For instance, if a C routine tried to pass you -1, it would appear as hex FFFFFFFF. If a value is passed that

does not fit the type you have chosen for a given parameter, a EUPHORIA type-check error may occur (depending on [type_check](#))

No error will occur if you declare all parameters as `atom`.

Normally, as in the case of `~WndProc()` above, Windows initiates these call-backs to your routines. **It is also possible for a C routine in any .dll to call one of your EUPHORIA routines.** You just have to declare the C routine properly, and pass it the call-back address.

Here's an example of a WATCOM C routine that takes your call-back address as its only parameter, and then calls your 3-parameter EUPHORIA routine:

```
/* 1-parameter C routine that you call from EUPHORIA */
unsigned EXPORT APIENTRY test1(
    LRESULT CALLBACK (*eu_callback)(unsigned a,
    unsigned b,
    unsigned c))
{
    /* Your 3-parameter EUPHORIA routine is called here
    via eu_callback pointer */
    return (*eu_callback)(111, 222, 333);
}
```

The C declaration above declares `test1` as an externally-callable C routine that takes a single parameter. The single parameter is a pointer to a routine that takes 3 unsigned parameters - i.e. your EUPHORIA routine.

In WATCOM C, "CALLBACK" is the same as "`__stdcall`". This is the calling convention that's used to call WIN32 API routines, and the C pointer to your EUPHORIA routine should be declared this way too, or you'll get an error when your EUPHORIA routine tries to return to your .DLL.

If you need your EUPHORIA routine to be called using the `__cdecl` convention, you must code the call to `call_back()` as:

```
myroutineaddr = call_back({'+', id})
```

The plus sign and braces indicate the `__cdecl` convention. The simple case, with no braces, is `__stdcall`.

In the example above, your EUPHORIA routine will be passed the three values 111, 222 and 333 as arguments. Your routine will return a value to `test1`. That value will then be immediately returned to the caller of `test1` (which could be at some other place in your EUPHORIA program).

A call-back address can be passed to the Linux or FreeBSD `signal()` function to specify a EUPHORIA routine to handle various signals (e.g. `SIGTERM`). It can also be passed to C routines such as `qsort()`, to specify a EUPHORIA comparison function.

6.10 Performance Tips

6.10.1 General Tips

- If your program is fast enough, forget about speeding it up. Just make it simple and readable.
- If your program is way too slow, the tips below will probably not solve your problem. You should find a better overall algorithm.
- The easiest way to gain a bit of speed is to turn off run-time type-checking. Insert the line:

```
without type_check
```

at the top of your main `.ex` file, ahead of any include statements. You'll typically gain between 0 and 20 percent depending on the types you have defined, and the files that you are including. Most of the standard include files do some user-defined type-checking. A program that is completely without user-defined type-checking might still be speeded up slightly.

Also, be *sure* to remove, or comment-out, any

```
with trace
with profile
with profile_time
```

statements. **with trace** (even without any calls to `trace`), and **with profile** can easily slow you down by 10% or more. **with profile_time** might slow you down by 1%. Each of these options will consume extra memory as well.

- Calculations using integer values are faster than calculations using floating-point numbers
- Declare variables as integer rather than atom where possible, and as sequence rather than object where possible. This usually gains you a few percent in speed.
- In an expression involving floating-point calculations it's usually faster to write constant numbers in floating point form, e.g. when `x` has a floating-point value, say, `x = 9.9`

change:

```
x = x * 5
```

to:

```
x = x * 5.0
```

This saves the interpreter from having to convert integer 5 to floating-point 5.0 each time.

- EUPHORIA does *short-circuit* evaluation of `if`, `elsif`, and `while` conditions involving `and` and `or`. EUPHORIA will stop evaluating any condition once it determines if the condition is true or not. For instance in the *if-statement*:

```
if x > 20 and y = 0 then
...
end if
```

The "`y = 0`" test will only be made when "`x > 20`" is true.

For maximum speed, you can order your tests. Do "`x > 20`" first if it is more likely to be false than "`y = 0`".

In general, with a condition "`A and B`", EUPHORIA will not evaluate the expression `B`, when `A` is false

(zero). Similarly, with a condition like "A or B", B will not be evaluated when A is true (non-zero). Simple if-statements are highly optimized. With the current version of the interpreter, nested simple if's that compare integers are usually a bit faster than a single short-circuit if-statement e.g.:

```
if x > 20 then
  if y = 0 then
    ...
  end if
end if
```

- The speed of access to private variables, local variables and global variables is the same.
 - There is no performance penalty for defining constants versus plugging in hard-coded literal numbers.
- The speed of:

```
y = x * MAX
```

is exactly the same as:

```
y = x * 1000
```

where you've previously defined:

```
constant MAX = 1000
```

- There is no performance penalty for having lots of comments in your program. Comments are completely ignored. They are not executed in any way. It might take a few milliseconds longer for the initial load of your program, but that's a very small price to pay for future maintainability, and when you **bind** your program, or **translate** your program to C, all comments are stripped out, so the cost becomes absolute zero.

6.10.2 Measuring Performance

In any programming language, and especially in EUPHORIA, **you really have to make measurements before drawing conclusions about performance.**

EUPHORIA provides both **execution-count profiling**, as well as **time profiling**. You will often be surprised by the results of these profiles. Concentrate your efforts on the places in your program that are using a high percentage of the total time (or at least are executed a large number of times.) There's no point to rewriting a section of code that uses 0.01% of the total time. Usually there will be one place, or just a few places where code tweaking will make a significant difference.

You can also measure the speed of code by using the `time` function. e.g.

```
atom t
t = time()
for i = 1 to 10000 do
  -- small chunk of code here
end for
? time() - t
```

You might rewrite the small chunk of code in different ways to see which way is faster.

6.10.3 How to Speed-Up Loops

Profiling will show you the *hot spots* in your program. These are usually inside loops. Look at each calculation inside the loop and ask yourself if it really needs to happen every time through the loop, or could it be done just once, prior to the loop.

6.10.4 Converting Multiplies to Adds in a Loop

Addition is faster than multiplication. Sometimes you can replace a multiplication by the loop variable, with an addition. Something like:

```
for i = 0 to 199 do
    poke(screen_memory+i*320, 0)
end for
```

becomes:

```
x = screen_memory
for i = 0 to 199 do
    poke(x, 0)
    x = x + 320
end for
```

6.10.5 Saving Results in Variables

- It's faster to save the result of a calculation in a variable, than it is to recalculate it later. Even something as simple as a subscript operation, or adding 1 to a variable is worth saving.
- When you have a sequence with multiple levels of subscripting, it is faster to change code like:

```
for i = 1 to 1000 do
    y[a][i] = y[a][i]+1
end for
```

to:

```
ya = y[a]
for i = 1 to 1000 do
    ya[i] = ya[i] + 1
end for
y[a] = ya
```

So you are doing 2 subscript operations per iteration of the loop, rather than 4. The operations, `ya = y[a]` and `y[a] = ya` are very cheap. **They just copy a pointer.** They don't copy a whole sequence.

- There is a slight cost when you create a new sequence using **{a,b,c}**. If possible, move this operation out of a critical loop by storing it in a variable before the loop, and referencing the variable inside the loop.

6.10.6 In-lining of Routine Calls

If you have a routine that is rather small and fast, but is called a huge number of times, you will save time by doing the operation *in-line*, rather than calling the routine. Your code may become less readable, so it might be better to in-line only at places that generate a lot of calls to the routine.

6.10.7 Operations on Sequences

EUPHORIA lets you operate on a large sequence of data using a single statement. This saves you from writing a loop where you process one element at-a-time. e.g.

```
x = {1, 3, 5, 7, 9}
y = {2, 4, 6, 8, 10}
z = x + y
```

versus:

```
z = repeat(0, 5) -- if necessary
for i = 1 to 5 do
    z[i] = x[i] + y[i]
end for
```

In most interpreted languages, it is much faster to process a whole sequence (array) in one statement, than it is to perform scalar operations in a loop. This is because the interpreter has a large amount of overhead for each statement it executes.

EUPHORIA is different. EUPHORIA is very lean, with little interpretive overhead, so operations on sequences don't always win. The only solution is to time it both ways. The per-element cost is usually lower when you process a sequence in one statement, but there are overheads associated with allocation and deallocation of sequences that may tip the scale the other way.

6.10.8 Some Special Case Optimizations

EUPHORIA automatically optimizes certain special cases. *x* and *y* below could be variables or arbitrary expressions.

```
x + 1      -- faster than general x + y
1 + x      -- faster than general y + x
x * 2      -- faster than general x * y
2 * x      -- faster than general y * x
x / 2      -- faster than general x / y
floor(x/y) -- where x and y are integers, is faster than x/y
floor(x/2) -- faster than floor(x/y)
```

x below is a simple variable, *y* is any variable or expression:

```
x = append(x, y)  -- faster than general z = append(x, y)
x = prepend(x, y) -- faster than general z = prepend(x, y)

x = x & y          -- where x is much larger than y,
                   -- is faster than general z = x & y
```

When you write a loop that "grows" a sequence, by appending or concatenating data onto it, the time will, in general, grow in proportion to the **square** of the number (N) of elements you are adding. However, if you can use one of the special optimized forms of `append()`, `prepend()` or concatenation listed above, the time will grow in proportion to just N (roughly). This could save you a **huge** amount of time when creating an extremely long sequence.

(You could also use `repeat()` to establish the maximum size of the sequence, and then fill in the elements in a loop, as discussed below.)

6.10.9 Assignment with Operators

For greater speed, convert:

```
**left-hand-side = left-hand-side op expression**
```

to:

```
**left-hand-side op= expression**
```

For example:

```
-- Instead of ...
some_val = some_val * 3
-- Use ...
some_val *= 3
```

whenever left-hand-side contains at least 2 subscripts, or at least one subscript and a slice. In all simpler cases the two forms run at the same speed (or very close to the same).

6.10.10 Pixel-Graphics Tips

- Mode 19 is the fastest mode for **animated graphics** and **games**.
- The video memory (in mode 19) is not cached by the CPU. It usually takes longer to read or write data to the screen than to a general area of memory that you allocate. This adds to the efficiency of *virtual screens*, where you do all of your image updating in a **block of memory** that you get from **allocate()**, and then you periodically **mem_copy()** the resulting image to the real **screen memory**. In this way you never have to read the (slow) screen memory.
- When plotting pixels, you may find that modes 257 and higher are fast near the top of the screen, but slow near the bottom.

6.10.11 Library Routines

Some common routines are extremely fast. You probably couldn't do the job faster any other way, even if you used C or assembly language. Some of these are:

- `mem_copy()`
- `mem_set()`

- `repeat()`

Other routines are reasonably fast, but you might be able to do the job faster in some cases if speed was crucial.

```
x = repeat(0,100) -- Pre-allocate all the elements first.
for i = 1 to 100 do
    x[i] = i
end for
```

is somewhat faster than:

```
x = {}
for i = 1 to 100 do
    x = append(x, i)
end for
```

because `append()` has to allocate and reallocate space as `x` grows in size. With `repeat()`, the space for `x` is allocated once at the beginning. (`append()` is smart enough not to allocate space with *every* `append` to `x`. It will allocate somewhat more than it needs, to reduce the number of reallocations.)

You can replace:

```
remainder(x, p)
```

with:

```
and_bits(x, p-1)
```

for greater speed when `p` is a positive power of 2. `x` must be a non-negative integer that fits in 32-bits.

`arctan` is faster than `arccos` or `arcsin`.

6.10.12 Searching

EUPHORIA's `find` is the fastest way to search for a value in a sequence up to about 50 elements. Beyond that, you might consider a *hash table* (`demo\hash.ex`) or a *binary tree* (`demo\tree.ex`).

6.10.13 Sorting

In most cases you can just use the *shell sort* routine in `sort.e`.

If you have a huge amount of data to sort, you might try one of the sorts in `demo\allsorts.e` (e.g. *great sort*). If your data is too big to fit in memory, don't rely on EUPHORIA's automatic memory swapping capability. Instead, sort a few thousand records at a time, and write them out to a series of temporary files. Then merge all the sorted temporary files into one big sorted file.

If your data consists of integers only, and they are all in a fairly narrow range, try the *bucket sort* in `demo\allsorts.e`.

6.10.14 Taking Advantage of Cache Memory

As CPU speeds increase, the gap between the speed of the on-chip cache memory and the speed of the main memory or DRAM (dynamic random access memory) becomes ever greater. You might have 256 Mb of DRAM on your computer, but the on-chip cache is likely to be only 8K (data) plus 8K (instructions) on a Pentium, or 16K (data) plus 16K (instructions) on a Pentium with MMX or a Pentium II/III. Most machines will also have a "level-2" cache of 256K or 512K.

An algorithm that steps through a long sequence of a couple of thousand elements or more, many times, from beginning to end, performing one small operation on each element, will not make good use of the on-chip data cache. It might be better to go through once, applying several operations to each element, before moving on to the next element. The same argument holds when your program starts swapping, and the least-recently-used data is moved out to disk.

These cache effects aren't as noticeable in EUPHORIA as they are in lower-level compiled languages, but they are measurable.

6.10.15 Using Machine Code and C

EUPHORIA lets you call routines written in 32-bit Intel machine code. On **WIN32** and **Linux** and **FreeBSD** you can call C routines in `dll` or `so` files, and these C routines can call your EUPHORIA routines. You might need to call C or machine code because of something that can't be done directly in EUPHORIA, or you might do it for improved speed.

To boost speed, the machine code or C routine needs to do a significant amount of work on each call, otherwise the overhead of setting up the arguments and making the call will dominate the time, and it might not gain you much.

Many programs have some inner core operation that consumes most of the CPU time. If you can code this in C or machine code, while leaving the bulk of the program in EUPHORIA, you might achieve a speed comparable to C, without sacrificing EUPHORIA's safety and flexibility.

6.10.16 Using The EUPHORIA To C Translator

From version 3.0, the full EUPHORIA To C Translator is included in the installation package. It will translate any EUPHORIA program into a set of C source files that you can compile using a C compiler.

The executable file that you get using the Translator should run the same, but faster than when you use the interpreter. The speed-up can be anywhere from a few percent to a factor of 5 or more.

7 Included Tools

EuTEST - Unit Testing

- Introduction
- The eutest Program
- The Unit Test Files
- The Error Control Files

EuDOC - Source Documentation Tool

- Documentation tags
- Generic documentation
- Source documentation
- Assembly file
- Creole markup
- Documentation software

Ed - EUPHORIA Editor

- Introduction
- Summary
- Special Keys
- Escape Commands
- Recalling Previous Strings
- Cutting and Pasting
- Use of Tabs
- Long Lines
- Maximum File Size
- Non-text Files
- Line Terminator
- Source Code
- Platform Issues

7.1 EuTEST - Unit Testing

- Introduction
- The eutest Program
 - Synopsis for running the tests
 - Synopsis for creating report from the log
 - General behavior
 - Options detail
- The Unit Test Files
 - A trivial example
 - You can generate these files by
- The Error Control Files

7.1.1 Introduction

The testing system gives you the ability to check if the library, interpreter and translator works properly by use of *unit tests*. The unit tests are **EUPHORIA** include files that include `unittest.e` at the top, several test-routines for comparison between expected value and true value and at the end of the program a call to `test_report()`. There are error control files for when we expect the interpreter to fail but we want it to fail with a particular error message.

7.1.2 The eutest Program

7.1.2.1 Synopsis for running the tests

```
stable-interpreter-path [-D REC] eutest.ex
    [-verbose] [-log] [-i include path] [-cc wat|gcc] [-exe interpreter]
    [-ec translator] [-lib binary library path]
    [optional list of unit test files]
```

7.1.2.2 Synopsis for creating report from the log

```
stable-interpreter-path eutest.ex -process-log [-html]
```

7.1.2.3 General behavior

For your *stable-interpreter-path* please use a stable interpreter, to run `eutest.ex`. You should use a alpha or beta release copy as your '*stable-interpreter-path*', then you can use `-exe` to specify your '*test interpreter path*' which could be your build of an SVN version or your own modified copy. If you want to test translation as well, you can specify it in a similar way with `-ec`. Developers, **please** do not modify `eutest.ex` to use features not available in the last two alpha releases. If you don't specify unit tests on the command line `eutest` will scan the directory for unit test files using the pattern `t_*.e`. If you specify a pattern it will interpret the pattern as some shells do not do this for programs.

7.1.2.4 Options detail

- `-D REC`: Is for creating control files, use only when on tests that work already on an interpreter that correctly works or correctly
- `fails*` with them. This option must come before the `eutest.ex` program itself in the command line and is the option with that requirement.
- `-log`: Is for creating a log file for later processing
- `-verbose`: Is for `eutest.ex` to give you detail of what it is doing
- `-i`: is for specify the include path which will be passed to both the interpreter and the translator when interpreting and translating the test.
- `-cc`: is for specifying the compiler. This can be any one of `-wat`, `djg`, or `lcc`. Each of these represent the kind of compiler we will request the translator to use.

- -process-log: Is for processing a log created by a previous invocation of eutest.ex output is sent to standard output as a report of how the tests went. By default this is in ascii format. Use -html to make it HTML format.
- -html: Is for making the report creation to be in HTML format

7.1.3 The Unit Test Files

Unit test files must match a pattern `t_* .e`. If the unit test file matches `t_c_* .e` the test program will expect the program to fail, if there is an error control file in a directory with its same name and 'd' extension it will also expect it to fail according to the control file's contents. Found in the said directory.

7.1.3.1 A trivial example

The following is a minimal unit test [file](#):

```
include std/unittest.e

test_report()
```

Please see "Unit Test Framework", currently 13.5, for information on how to construct these files.

7.1.3.2 You can generate these files by

7.1.4 The Error Control Files

There are times when we expect something to fail. We want good EUPHORIA code to do the correct thing and there is a correct thing to do also for *bad* code. The interpreter must return with an error message of why it failed and the error must be correct and it must get written to ex.err. We must thus check the ex.err file to see if it has the correct error message.

If the unit test is `t_foo.e` then the location for its control file can be in the following locations:

- `t_foo.d/interpreter/OSNAME/control.err`
- `t_foo.d/OSNAME/control.err`
- `t_foo.d/control.err`

The *OSNAME* is the name of the operating system. Which is either UNIX or Win32.

Now, if `t_foo.d/Win32/control.err` exists, then the testing program eutest.ex expects `t_foo.e` to fail when run with the WIN32 interpreter. However, this is not necessarily true for other platforms. In WIN32, eutest runs it, watches it fail, then compares the ex.err file to `t_foo.d/Win32/control.err`. If they ex.err is different from control.err an error message is written to the log. Now on, say NetBSD, `t_file.e` is tested with the expectation it will return 0 and the tests will all pass unless `t_foo.d/UNIX/control.err` or `t_foo.d/control.err` also exist. Thus you can have different expectations for differing platforms. Some feature that is not possible to implement under WIN32 can be put

into a unit test and the resulting `ex.err` file can be put into a control file for WIN32. This means we do not need to have all of these errors that we expect to get drawing our attention away from errors that need our attention. On the other hand, if an unexpected error message not like `t_foo.d/Win32/control.err` gets generated in the Windows case then eutest will tell us that.

How do we construct these control files? You don't really need to, you can take an `ex.err` file that results from running a stable interpreter on a test and rename it and move it to the appropriate place.

7.2 EuDOC - Source Documentation Tool

- [Documentation tags](#)
- [Generic documentation](#)
- [Source documentation](#)
- [Assembly file](#)
- [Creole markup](#)
- [Documentation software](#)

Write your documentation, in the form of comments, within the source-code of your program. EUPHORIA comments do not slow down the execution of programs. Documentation written inside source-code introduces no speed penalty, but is very convenient.

The **eudoc** program extracts documentation written inside source-code, and will extract information about routines and identifiers. `eudoc` can also incorporate documentation written externally from your source-code. This automates the creation of documentation and makes it easy to maintain.

Write your material using *Creole* style markup to format documentation. This gives you creative control using elements like headers, fonts, cross-references, tables... The **creolehtml** program takes the output of `eudoc` and produces html formatted documentation.

A third party program like `htmldoc` may then be used to convert html to pdf.

7.2.1 Documentation tags

Documentation is embedded in source-code using the EUPHORIA line (`--`) comments. Two special tags, `--****` and `--**` distinguish documentation from comments that will not be extracted.

7.2.2 Generic documentation

"Generic" documentation starts with the (`--****`) tag, continues with lines starting with `--` in the first column, and ends with the next blank line. The tags and `--` will not appear in the documentation.

```
--****
-- generic text, thus tagged, will be extracted by eudoc
-- write your documentation here...
--
```

-- blank line is a terminator, this line is not included

Produces...

generic text, thus tagged, will be extracted by eudoc
write your documentation here...

7.2.3 Source documentation

"Source" documentation starts with the (`--**`) tag. Locate them before a routine or identifier that you wish to be described in your documentation. The `eudoc` program will extract the "signature" of a routine and combines it with the comments that you write after this tag.

- Starting with the source-code file **favorite.ex** ...

```
--**  
-- this is my favorite routine  
  
procedure hello( sequence name )  
    printf(1, "Hello %s!", {name} )  
end procedure
```

- Executing **eui eudoc favorite.ex -o foo.txt** produces...

```
%%disallow={camelcase}  
  
!!CONTEXT:favorite.ex  
  
@[hello|]  
==== hello  
<eucode>  
include favorite.ex  
procedure hello(sequence name)  
</eucode>
```

This is my favorite routine.

- Process with **eui creolehtml foo.txt** ...

hello

```
include favorite.ex  
procedure hello(sequence name)
```

This is my favorite routine.

- If you examine the source-code included with EUPHORIA you will realize how these steps were used to create the documentation you are reading now.

7.2.4 Assembly file

Large projects are managed using an *assembly file*, which is a list of files (source-code, and external) that will be incorporated into one output file. Look at `euphoria/docs/manual.af` for the file used to produce this documentation.

7.2.5 Creole markup

Creole is a text markup language used in wikis, such as the EUPHORIA Wiki, and for documenting source-code.

- Common Creole tags are:

= Title

<<LEVELTOC depth=2>>

== Section

//italic// **bold** ##fixed##

* bullet
* lists are
* easy to produce

|| tables || are |
| easy to produce | //with bold headers// |

```
<eucode>
-- euphoria code is colorized
for i=1 to 5 do
    ? i
end for
</eucode>
```

- The previous tags will produce html that looks like...

Title
Section

italic **bold** fixed

- bullet
- lists are

- easy to produce

tables	are
easy to produce	<i>with bold headers</i>

```
-- euphoria code is colorized
for i=1 to 5 do
    ? i
end for
```

- More choices are described at the EUPHORIA Wiki.

7.2.6 Documentation software

The programs eudoc ... <https://rapideuphoria.svn.sourceforge.net/svnroot/rapideuphoria/tools/eudoc/trunk>

and creolehtml ... <https://rapideuphoria.svn.sourceforge.net/svnroot/rapideuphoria/tools/creole/trunk>

are hosted at Source-Forge ... <http://sourceforge.net/projects/rapideuphoria/>

More on using eudoc... <http://openeuphoria.org/wiki/euwiki.cgi?Documenting40>

More on using Creole markup... <http://openeuphoria.org/wiki/euwiki.cgi?CreoleHelp>

The program htmldoc is found at... <http://www.htmldoc.org/> and <http://htmldoc-binaries.org/>

7.3 Ed - EUPHORIA Editor

- [Introduction](#)
- [Summary](#)
- [Special Keys](#)
- [Escape Commands](#)
- [Recalling Previous Strings](#)
- [Cutting and Pasting](#)
- [Use of Tabs](#)
- [Long Lines](#)
- [Maximum File Size](#)
- [Non-text Files](#)
- [Line Terminator](#)
- [Source Code](#)
- [Platform Issues](#)

7.3.1 Introduction

The EUPHORIA download package includes a handy, text-mode editor, `ed`, that's written completely in EUPHORIA. Many people find `ed` convenient for editing EUPHORIA programs and other files, but there is no requirement that you use it.

If you don't like `ed`, you have many alternatives. David Cuny's **EE editor** is a text-based editor for EUPHORIA that's also written in EUPHORIA. It has a friendly mouse-based user interface with drop down menus etc. It's available from the RDS Web site. There are several other EUPHORIA-oriented editors that run on Windows and support files for use with other editors with EUPHORIA.

If you search for **Editor** in our Archive, you'll find many editors written **in** EUPHORIA or **for** EUPHORIA. Some written **in** EUPHORIA and **for** EUPHORIA. In fact, *any* text editor can be used to edit a EUPHORIA program, such as `vi`, `jEdit`, `Emacs`, and even dinosaur applications like Windows `NotePad`.

Among the editors you will find Judith's IDE and `wxIDE`: Integrated Development Environments.

Do you still want to use this `ed`? Then read on...

7.3.2 Summary

Usage:

1. `ed filename`
2. `ed`

After any error, just type `ed`, and you'll be placed in the editor, at the line and column where the error was detected. The error message will be at the top of your screen.

EUPHORIA-related files are displayed in color. Other text files are in mono. You'll know that you have misspelled something when the color does not change as you expect. Keywords are blue. Names of routines that are built in to the interpreter appear in magenta. Strings are green, comments are red, most other text is black. Balanced brackets (on the same line) have the same color. You can change these colors as well as several other parameters of `ed`. See "user-modifiable parameters" near the top of `ed.ex`.

The arrow keys move the cursor left, right, up or down. Most other characters are immediately inserted into the file.

In Windows, you can "associate" various types of files with `ed.bat`. You will then be put into `ed` when you *double-click* on these types of files - e.g. `.e`, `.pro`, `.doc` etc. Main EUPHORIA files ending in `.ex`, `.exd` or `.exw` might better be associated with `eui.exe`, `eid.exe`, or `euiw.exe`, respectively.

`ed` is a multi-file/multi-window text-based editor. `Esc c` will split your screen so you can view and edit up to 10 files simultaneously, with cutting and pasting between them. You can also use multiple edit windows to view and edit different parts of a single file.

7.3.3 Special Keys

Some PC keys do not work in a Linux or FreeBSD text console, or in Telnet, and some keys do not work in an xterm under X windows. Alternate keys have been provided. In some cases on Linux/FreeBSD you might have to edit `ed.exe` to map the desired key to the desired function.

Delete	Delete the current character above the cursor
Backspace	Move the cursor to the left and delete a character
C-Delete	Delete the current line (not available on all platforms)
C-d	Delete the current line (same as C-Delete)
Insert	re-insert the preceding series of Deletes before the current line/character
C-Left	Move to the start of the previous word. On Unix, use C-l
C-Right	Move to the start of the next word. On Unix, use C-r
Home	Move to the beginning of the current line
End	Move to the end of the current line
C-Home	Move to the beginning of the file (<code>euide.exe</code> only, others use C-t)
C-End	Move to the end of the file (<code>euide.exe</code> only, others use C-b)
PgUp	Move up one screen. On Unix use C-u
PgDn	Move down one screen. On Unix use C-p
F1..F10	Select a new window. The windows are numbered from top to bottom with the top window on the screen being <i>F1</i>
F12	User definable key (see <code>CUSTOM_KEYSTROKES</code> near top of <code>ed.exe</code> . Default action is to insert <code>--</code> for a EUPHORIA comment

7.3.4 Escape Commands

Press and release the *Esc* key, then press one of the following keys:

- h Get help text for the editor, or EUPHORIA. The screen is split so you can view your program and the help text at the same time.
"Clone" the current window, i.e. make a new edit window that is initially viewing the same file at the same position as the current window. The sizes of all windows are adjusted to make room for the new window. You might want to use **Esc l** to get more lines on the screen. Each window that you create can
- c be scrolled independently and each has its own menu bar. The changes that you make to a file will initially appear only in the current window. When you press an **F-key** to select a new window, any changes will appear there as well. You can use **Esc n** to read a new file into any window.
- q Quit (delete) the current window and leave the editor if there are no more windows. You'll be warned if this is the last window used for editing a modified file. Any remaining windows are given more space.
- s Save the file being edited in the current window, then quit the current window as **Esc q** above.
- w Save the file but do not quit the window.
Save the file, and then execute it with `euide`, `euidew` or `euil`. When the program finishes execution you'll
- e hear a beep. Hit *Enter* to return to the editor. This operation may not work if you are very low on extended memory. You can't supply any command-line arguments to the program.
- d

Run an operating system command. After the beep, hit *Enter* to return to the editor. You could also use this command to edit another file and then return, but *Esc c* is probably more convenient.

n Start editing a new file in the current window. Deleted lines/chars and search strings are available for use in the new file. You must type in the path to the new file. Alternatively, you can drag a file name from a Windows file manager window into the console window for *ed*. This will type the full path for you.

f Find the next occurrence of a string in the current window. When you type in a new string there is an option to "match case" or not. Press *y* if you require upper/lower case to match. Keep hitting *Enter* to find subsequent occurrences. Any other key stops the search. To search from the beginning, press *C-Home* before *Esc f*. The default string to search for, if you don't type anything, is shown in double quotes.

r Globally replace one string by another. Operates like *Esc f* command. Keep hitting *Enter* to continue replacing. Be careful - *there is no way to skip over a possible replacement*.

l Change the number of lines displayed on the screen. Only certain values are allowed, depending on your video card. Many cards will allow 25, 28, 43 and 50 lines. In a Linux/FreeBSD text console you're stuck with the number of lines available (usually 25). In a Linux/FreeBSD xterm window, *ed* will use the number of lines initially available when *ed* is started up. Changing the size of the window will have no effect after *ed* is started.

m Show the modifications that you've made so far. The current edit buffer is saved as *editbuff.tmp*, and is compared with the file on disk using the Windows *fc* command, or the Linux/FreeBSD *diff* command. *Esc m* is very useful when you want to quit the editor, but you can't remember what changes you made, or whether it's ok to save them. It's also useful when you make an editing mistake and you want to see what the original text looked like.

ddd Move to line number *ddd*. e.g. *Esc 1023 Enter* would move to line 1023 in the file.

CR *Esc Carriage-Return*, i.e. *Esc Enter*, will tell you the name of the current file, as well as the line and character position you are on, and whether the file has been modified since the last save. If you press *Esc* and then change your mind, it is harmless to just hit *Enter* so you can go back to editing.

7.3.5 Recalling Previous Strings

The *Esc n*, *Esc d*, *Esc r* and *Esc f* commands prompt you to enter a string. You can recall and edit these strings just as you would at the command line. Type up-arrow or down-arrow to cycle through strings that you previously entered for a given command, then use left-arrow, right-arrow and the delete key to edit the strings. Press *Enter* to submit the string.

7.3.6 Cutting and Pasting

When you *C-Delete* (or *C-d*) a series of consecutive lines, or *Delete* a series of consecutive characters, you create a "kill-buffer" containing what you just deleted. This kill-buffer can be re-inserted by moving the cursor and then pressing *Insert*.

A new kill-buffer is started, and the old buffer is lost, each time you move away and start deleting somewhere else. For example, cut a series of *lines* with *C-Delete*. Then move the cursor to where you want to paste the lines and press *Insert*. If you want to copy the lines, without destroying the original text, first *C-Delete* them, then immediately press *Insert* to re-insert them. Then move somewhere else and press *Insert* to insert them again, as many times as you like. You can also *Delete* a series of individual *characters*, move the cursor, and

then paste the deleted characters somewhere else. Immediately press *Insert* after deleting if you want to copy without removing the original characters.

Once you have a kill-buffer, you can type *Esc n* to read in a new file, or you can press an *F-key* to select a new edit window. You can then insert your kill-buffer.

7.3.7 Use of Tabs

The standard *tab* width is 8 spaces. The editor assumes `tab=8` for most files. However, it is more convenient when editing a program for a tab to equal the amount of space that you like to indent. Therefore you will find that tabs are set to 4 when you edit EUPHORIA files (or `.c`, or `.h` or `.bas` files). The editor converts from `tab=8` to `tab=4` when reading your *program* file, and converts back to `tab=8` when you save the file. Thus your file remains compatible with the `tab=8` world. **If you would like to choose a different number of spaces to indent**, change the line at the top of `ed.ex` that says `"constant PROG_INDENT = 4"`.

7.3.8 Long Lines

Lines that extend beyond the right edge of the screen are marked with an *inverse video* character in the 80th column. This warns you that there is more text "out there" that you can't see. You can move the cursor beyond the 80th column. The screen will scroll left or right so the cursor position is always visible.

7.3.9 Maximum File Size

Like any EUPHORIA program, `ed` can access all the memory on your machine. It can edit huge files, and unless disk swapping occurs, most operations will be very fast.

7.3.10 Non-text Files

`ed` is designed for editing pure text files, although you can use it to view other files. As `ed` reads in a file, it replaces certain non-printable characters (less than ASCII 14) with ASCII 254 - small square. *If you try to save a non-text file you will be warned about this.* Since `ed` opens all files as "text" files, a *control-z* character (26) embedded in a file will appear to `ed` to be the *end of the file*.

7.3.11 Line Terminator

The end-of-line terminator on Linux/FreeBSD/OSX/SUNOS/OPENBSD/NETBSD is simply `\n`. On Windows, text files have lines ending with `\r\n`. If you copy a Windows file to Linux/FreeBSD and try to modify it, **ed** will give you a choice of either keeping the `\r\n` terminators, or saving the file with `\n` terminators.

7.3.12 Source Code

The complete source code to this editor is in `bin\ed.exe` and `bin\syncolor.e`. You are welcome to make improvements. There is a section at the top of `ed.exe` containing "user-modifiable" configuration parameters that you can adjust. The colors and the cursor size may need adjusting for some operating environments.

7.3.13 Platform Issues

`euphoria\bin\ed.bat` can be set up to run `ed.exe` using `eui.exe` or `euid.exe`. You are better off running **ed** with `euid.exe` on Windows 95/98/ME. You'll get much quicker screen updates than with `eui.exe`. On Windows XP you'll be a bit better off using `eui.exe`. You'll get slightly quicker screen updates, and you'll be able to create files with long names, not just open existing ones. However some special keys won't work with `eui.exe`, e.g. you'll have to use *C-t* and *C-b* instead of *C-Home* and *C-End*. On other platforms there are no problems with long filenames, and the keyboard response is always fast.

8 General Routine Reference

Command Line Handling

- Constants

- Routines

Console

- Cursor Style Constants

- Keyboard related routines

- Cross Platform Text Graphics

Date/Time

- Localized Variables

- Constants

- Types

- Routines

File System

- Constants

- Directory Handling

- File name parsing

- File Types

- File Handling

I/O

- Constants

- Read/Write Routines

- Low Level File/Device Handling

- File Reading/Writing

Math

- Sign and comparisons

- Roundings and remainders

- Trigonometry

- Logarithms and powers.

- Hyperbolic trigonometry

- Accumulation

- Bitwise operations

Math Constants

- Constants

Random Numbers

Mouse

- Requirements

- Constants

- Routines

Operating System Helpers

- Operating System Constants

- Environment.

- Interacting with the OS

- Miscellaneous

- Pipe Input/Output

- Accessor Constants

- Opening/Closing

- Read/Write Process
- Pretty Printing
 - Routines
- Statistics
 - Routines
- Multi-tasking
 - General Notes
 - Warning
 - Routines
- Types - Extended
 - Support Functions
 - Types

8.1 Command Line Handling

Constants

- NO_PARAMETER
- HAS_PARAMETER
- NO_CASE
- HAS_CASE
- MANDATORY
- OPTIONAL
- ONCE
- MULTIPLE
- HELP
- NO_HELP
- HELP_RID
- VALIDATE_ALL
- NO_VALIDATION
- NO_VALIDATION_AFTER_FIRST_EXTRA
- SHOW_ONLY_OPTIONS
- AT_EXPANSION
- NO_AT_EXPANSION
- PAUSE_MSG
- OPT_IDX
- OPT_CNT
- OPT_VAL
- OPT_REV

Routines

- command_line
- option_switches
- show_help
- cmd_parse
- build_commandline
- parse_commandline

8.1.1 Constants

8.1.1.1 NO_PARAMETER

```
include std/cmdline.e
public constant NO_PARAMETER
```

This option switch does not have a parameter. See [cmd_parse](#)

8.1.1.2 HAS_PARAMETER

```
include std/cmdline.e
public constant HAS_PARAMETER
```

This option switch does have a parameter. See [cmd_parse](#)

8.1.1.3 NO_CASE

```
include std/cmdline.e
public constant NO_CASE
```

This option switch is not case sensitive. See [cmd_parse](#)

8.1.1.4 HAS_CASE

```
include std/cmdline.e
public constant HAS_CASE
```

This option switch is case sensitive. See [cmd_parse](#)

8.1.1.5 MANDATORY

```
include std/cmdline.e
public constant MANDATORY
```

This option switch must be supplied on command line. See [cmd_parse](#)

8.1.1.6 OPTIONAL

```
include std/cmdline.e
public constant OPTIONAL
```

This option switch does not have to be on command line. See [cmd_parse](#)

8.1.1.7 ONCE

```
include std/cmdline.e
public constant ONCE
```

This option switch must only occur once on the command line. See [cmd_parse](#)

8.1.1.8 MULTIPLE

```
include std/cmdline.e
public constant MULTIPLE
```

This option switch may occur multiple times on a command line. See [cmd_parse](#)

8.1.1.9 HELP

```
include std/cmdline.e
public constant HELP
```

This option switch triggers the 'help' display. See [cmd_parse](#)

8.1.1.10 NO_HELP

```
include std/cmdline.e
public constant NO_HELP
```

8.1.1.11 HELP_RID

```
include std/cmdline.e
public enum HELP_RID
```

Additional help routine id. See [cmd_parse](#)

8.1.1.12 VALIDATE_ALL

```
include std/cmdline.e
public enum VALIDATE_ALL
```

Validate all parameters (default). See [cmd_parse](#)

8.1.1.13 NO_VALIDATION

```
include std/cmdline.e
public enum NO_VALIDATION
```

Do not cause an error for an invalid parameter. See [cmd_parse](#)

8.1.1.14 NO_VALIDATION_AFTER_FIRST_EXTRA

```
include std/cmdline.e
public enum NO_VALIDATION_AFTER_FIRST_EXTRA
```

Do not cause an error for an invalid parameter after the first extra item has been found. This can be helpful for processes such as the Interpreter itself that must deal with command line parameters that it is not meant to handle. At expansions after the first extra are also disabled.

8.1.1.14.1 For instance:

`eui -D TEST greet.ex -name John -greeting Bye -D TEST` is meant for `eui`, but `-name` and `-greeting` options are meant for `greet.ex`. See [cmd_parse](#)

`eui @euopts.txt greet.ex @hotmail.com here 'hotmail.com'` is not expanded into the command line but `'euopts.txt'` is.

8.1.1.15 SHOW_ONLY_OPTIONS

```
include std/cmdline.e
public enum SHOW_ONLY_OPTIONS
```

Only display the option list in `show_help`. Do not display other information such as program name, options, etc... See [cmd_parse](#)

8.1.1.16 AT_EXPANSION

```
include std/cmdline.e
public enum AT_EXPANSION
```

Expand arguments that begin with '@' into the command line. (default) For example, @filename will expand the contents of file named 'filename' as if the file's contents were passed in on the command line. Arguments that come after the first extra will not be expanded when NO_VALIDATION_AFTER_FIRST_EXTRA is specified.

8.1.1.17 NO_AT_EXPANSION

```
include std/cmdline.e
public enum NO_AT_EXPANSION
```

Do not expand arguments that begin with '@' into the command line. Normally @filename will expand the file names contents as if the file's contents were passed in on the command line. This option supresses this behavior.

8.1.1.18 PAUSE_MSG

```
include std/cmdline.e
public enum PAUSE_MSG
```

Supply a message to display and pause just prior to abort() being called.

8.1.1.19 OPT_IDX

```
include std/cmdline.e
public enum OPT_IDX
```

An index into the `opts` list. See [cmd_parse](#)

8.1.1.20 OPT_CNT

```
include std/cmdline.e
public enum OPT_CNT
```

The number of times that the routine has been called by `cmd_parse` for this option. See [cmd_parse](#)

8.1.1.21 OPT_VAL

```
include std/cmdline.e
public enum OPT_VAL
```

The option's value as found on the command line. See [cmd_parse](#)

8.1.1.22 OPT_REV

```
include std/cmdline.e
public enum OPT_REV
```

The value 1 if the command line indicates that this option is to remove any earlier occurrences of it. See [cmd_parse](#)

8.1.2 Routines

8.1.2.1 command_line

```
<built-in> function command_line()
```

A **sequence**, of strings, where each string is a word from the command-line that started your program.

8.1.2.1.1 Returns:

1. The path, to either the EUPHORIA executable, (eui, eui.exe, euid.exe euiw.exe) or to your bound executable file.
2. The next word, is either the name of your EUPHORIA main file, or (again) the path to your bound executable file.
3. Any extra words, typed by the user. You can use these words in your program.

There are as many entries as words, plus the two mentioned above.

The EUPHORIA interpreter itself does not use any command-line options. You are free to use any options for your own program. It does have [command line switches](#) though.

The user can put quotes around a series of words to make them into a single argument.

If you convert your program into an executable file, either by binding it, or translating it to C, you will find that all command-line arguments remain the same, except for the first two, even though your user no longer types "eui" on the command-line (see examples below).

8.1.2.1.2 Example 1:

```
-- The user types:  eui myprog myfile.dat 12345 "the end"

cmd = command_line()

-- cmd will be:
{ "C:\EUPHORIA\BIN\EUI.EXE",
  "myprog",
  "myfile.dat",
  "12345",
  "the end" }
```

8.1.2.1.3 Example 2:

```
-- Your program is bound with the name "myprog.exe"
-- and is stored in the directory c:\myfiles
-- The user types:  myprog myfile.dat 12345 "the end"

cmd = command_line()

-- cmd will be:
{ "C:\MYFILES\MYPROG.EXE",
  "C:\MYFILES\MYPROG.EXE", -- place holder
  "myfile.dat",
  "12345",
  "the end"
}

-- Note that all arguments remain the same as example 1
-- except for the first two. The second argument is always
-- the same as the first and is inserted to keep the numbering
-- of the subsequent arguments the same, whether your program
-- is bound or translated as a .exe, or not.
```

8.1.2.1.4 See Also:

[build_commandline](#), [option_switches](#), [getenv](#), [cmd_parse](#), [show_help](#)

8.1.2.2 option_switches

```
<built-in> function option_switches()
```

Retrieves the list of switches passed to the interpreter on the command line.

8.1.2.2.1 Returns:

A **sequence**, of strings, each containing a word related to switches.

8.1.2.2.2 Comments:

All switches are recorded in upper case.

8.1.2.2.3 Example 1:

```
euiw -d helLo
-- will result in
-- option_switches() being {"-D", "helLo"}
```

8.1.2.2.4 See Also:

[Command line switches](#)

8.1.2.3 show_help

```
include std/cmdline.e
public procedure show_help(sequence opts, object add_help_rid = - 1, sequence cmds = command_line())
```

Show help message for the given opts.

8.1.2.3.1 Parameters:

1. `opts` : a sequence of options. See the [cmd_parse](#) for details.
2. `add_help_rid` : an object. Either a `routine_id` or a set of text strings. The default is -1 meaning that no additional help text will be used.
3. `cmds` : a sequence of strings. By default this is the output from [command_line\(\)](#)

8.1.2.3.2 Comments:

- `opts` is identical to the one used by [cmd_parse](#)
- `add_help_rid` can be used to provide additional help text. By default, just the option switches and their descriptions will be displayed. However you can provide additional text by either supplying a `routine_id` of a procedure that accepts no parameters; this procedure is expected to write text to the stdout device. Or you can supply one or more lines of text that will be displayed.

8.1.2.3.3 Example 1:

```
-- in myfile.ex
constant description = {
    "Creates a file containing an analysis of the weather.",
    "The analysis includes temperature and rainfall data",
    "for the past week."
}
```

```
show_help({
    {"q", "silent", "Suppresses any output to console", NO_PARAMETER, -1},
    {"r", 0, "Sets how many lines the console should display", {HAS_PARAMETER,"lines"}, -1}},
    description)
```

8.1.2.3.4 Outputs:

myfile.ex options:

```
-q, --silent      Suppresses any output to console
-r lines          Sets how many lines the console should display
```

Creates a file containing an analysis of the weather.
The analysis includes temperature and rainfall data
for the past week.

8.1.2.3.5 Example 2:

```
-- in myfile.ex
constant description = {
    "Creates a file containing an analysis of the weather.",
    "The analysis includes temperature and rainfall data",
    "for the past week."
}
procedure sh()
    for i = 1 to length(description) do
        printf(1, " >> %s <<\n", {description[i]})
    end for
end procedure

show_help({
    {"q", "silent", "Suppresses any output to console", NO_PARAMETER, -1},
    {"r", 0, "Sets how many lines the console should display", {HAS_PARAMETER,"lines"}, -1}},
    routine_id("sh"))
```

8.1.2.3.6 Outputs:

myfile.ex options:

```
-q, --silent      Suppresses any output to console
-r lines          Sets how many lines the console should display
```

```
>> Creates a file containing an analysis of the weather. <<
>> The analysis includes temperature and rainfall data <<
>> for the past week. <<
```

8.1.2.4 cmd_parse

```
include std/cmdline.e
public function cmd_parse(sequence opts, object parse_options = {}, sequence cmds = command_line)
```

Parse command line options, and optionally call procedures that relate to these options

8.1.2.4.1 Parameters:

1. `opts` : a sequence of valid option records: See Comments: section for details
2. `parse_options` : an optional sequence of parse options: See Parse Options section for details
3. `cmds` : an optional sequence of command line arguments. If omitted the output from `command_line()` is used.

8.1.2.4.2 Returns:

A **map**, containing the options set. The returned map has one special key named "extras" which are values passed on the command line that are not part of any option, for instance a list of files `myprog -verbose file1.txt file2.txt`. If any command element begins with an @ symbol then that file will be opened and its contents used to add to the command line.

8.1.2.4.3 Parse Options:

`parse_options` can be a sequence of options that will affect the parsing of the command line options. Options can be:

1. `VALIDATE_ALL` -- The default. All options will be validated for all possible errors.
2. `NO_VALIDATION` -- Do not validate any parameter.
3. `NO_VALIDATION_AFTER_FIRST_EXTRA` -- Do not validate any parameter after the first extra was encountered. This is helpful for programs such as the Interpreter itself: `eui -D TEST greet.ex -name John. -D TEST` should be validated but anything after "greet.ex" should not as it is meant for greet.ex to handle, not eui.
4. `HELP_RID` -- Specify a routine id to call in the event of a parse error (invalid option given, mandatory option not given, no parameter given for an option that requires a parameter, etc...) or a set of text strings. This can be used to provide additional help text. By default, just the option switches and their descriptions will be displayed. However you can provide additional text by either supplying a routine_id of a procedure that accepts no parameters; this procedure is expected to write text to the stdout device. Or you can supply one or more lines of text that will be displayed.
5. `NO_AT_EXPANSION` -- Do not expand arguments that begin with '@'.
6. `AT_EXPANSION` -- Expand arguments that begin with '@'. The name that follows @ will be opened as a file, read, and each trimmed non-empty line that does not begin with a '#' character will be inserted as arguments in the command line. These lines replace the original '@' argument as if they had been entered on the original command line.
 - ◆ If the name following the '@' begins with another '@', the extra '@' is removed and the remainder is the name of the file. However, if that file cannot be read, it is simply ignored. This allows *optional* files to be included on the command line. Normally, with just a single '@', if the file cannot be found the program aborts.
 - ◆ Lines whose first non-whitespace character is '#' are treated as a comment and thus ignored.
 - ◆ Lines enclosed with double quotes will have the quotes stripped off and the result is used as an argument. This can be used for arguments that begin with a '#' character, for example.
 - ◆ Lines enclosed with single quotes will have the quotes stripped off and the line is then further split up use the space character as a delimiter. The resulting 'words' are then all treated as individual arguments on the command line.

8.1.2.4.4 An example of parse options:

```
{ HELP_RID, routine_id("my_help"), NO_VALIDATION }
```

8.1.2.4.5 Comments:

8.1.2.4.6 Token types recognized on the command line:

1. a single '-'. Simply added to the 'extras' list
2. a single "--". This signals the end of command line options. What remains of the command line is added to the 'extras' list, and the parsing terminates.
3. -shortName. The option will be looked up in the short name field of `opts`.
4. /shortName. Same as -shortName.
5. -!shortName. If the 'shortName' has already been found the option is removed.
6. /!shortName. Same as -!shortName
7. --longName. The option will be looked up in the long name field of `opts`.
8. --!longName. If the 'longName' has already been found the option is removed.
9. anything else. The word is simply added to the 'extras' list.

For those options that require a parameter to also be supplied, the parameter can be given as either the next command line argument, or by appending '=' or ':' to the command option then appending the parameter data. For example, **-path=/usr/local** or as **-path /usr/local**.

On a failed lookup, the program shows the help by calling `show_help(opts, add_help_rid, cmds)` and terminates with status code 1.

8.1.2.4.7 Option records have the following structure:

1. a sequence representing the (short name) text that will follow the "-" option format. Use an atom if not relevant
2. a sequence representing the (long name) text that will follow the "--" option format. Use an atom if not relevant
3. a sequence, text that describes the option's purpose. Usually short as it is displayed when "-h"/"--help" is on the command line. Use an atom if not required.
4. An object ...
 - ◆ If an **atom** then it can be either `HAS_PARAMETER` or anything else if there is no parameter for this option. This format also implies that the option is optional, case-sensitive and can only occur once.
 - ◆ If a **sequence**, it can contain zero or more processing flags in any order ...
 - ◇ `MANDATORY` to indicate that the option must always be supplied.
 - ◇ `HAS_PARAMETER` to indicate that the option must have a parameter following it. You can optionally have a name for the parameter immediately follow the `HAS_PARAMETER` flag. If one isn't there, the help text will show "x" otherwise it shows the supplied name.
 - ◇ `NO_CASE` to indicate that the case of the supplied option is not significant.
 - ◇ `ONCE` to indicate that the option must only occur once on the command line.

◇ MULTIPLE to indicate that the option can occur any number of times on the command line.

- ◆ If both ONCE and MULTIPLE are omitted then switches that also have HAS_PARAMETER are only allowed once but switches without HAS_PARAMETER can have multiple occurrences but only one is recorded in the output map.

5. an integer; a [routine_id](#). This function will be called when the option is located on the command line and before it updates the map.

Use -1 if cmd_parse is not to invoke a function for this option.

The user defined function must accept a single sequence parameter containing four values. If the function returns 1 then the command option does not update the map. You can use the predefined index values OPT_IDX, OPT_CNT, OPT_VAL, OPT_REV when referencing the function's parameter elements.

1. An index into the `opts` list.
2. The number of times that the routine has been called by `cmd_parse` for this option
3. The option's value as found on the command line
4. 1 if the command line indicates that this option is to remove any earlier occurrences of it.

When assigning a value to the resulting map, the key is the long name if present, otherwise it uses the short name. For options, you must supply a short name, a long name or both.

If you want `cmd_parse()` to call a user routine for the extra command line values, you need to specify an Option Record that has neither a short name or a long name, in which case only the `routine_id` field is used.

For more details on how the command line is being pre-parsed, see [command_line](#).

8.1.2.4.8 Example:

```
sequence option_definition
integer gVerbose = 0
sequence gOutFile = {}
sequence gInFile = {}
procedure opt_verbose( sequence value)
    if value[OPT_VAL] = -1 then -- (!v used on command line)
        gVerbose = 0
    else
        if value[OPT_CNT] = 1 then
            gVerbose = 1
        else
            gVerbose += 1
        end if
    end if
end if
end procedure

procedure opt_output_filename( sequence value)
    gOutFile = value[OPT_VAL]
end procedure

procedure opt_extras( sequence value)
    if not file_exists(value[OPT_VAL]) then
        show_help(option_definitions, sprintf("Cannot find '%s'", value[OPT_VAL]))
        abort(1)
    end if
end if
```

```

    gInFile = append(gInFile, value[OPT_VAL])
end procedure

option_definition = {
    { "v", "verbose", "Verbose output", {NO_PARAMETER}, routine_id("opt_verbose") },
    { "h", "hash", "Calculate hash values", {NO_PARAMETER}, -1 },
    { "o", "output", "Output filename", {MANDATORY, HAS_PARAMETER, ONCE}, routine_id("opt_output") },
    { "i", "import", "An import path", {HAS_PARAMETER, MULTIPLE}, -1 },
    { 0, 0, 0, 0, routine_id("opt_extras") }
}

map:map opts = cmd_parse(option_definition)

-- When run as: eui myprog.ex -v @output.txt -i /etc/app input1.txt input2.txt
-- and the file "output.txt" contains the two lines ...
-- --output=john.txt
-- '-i /usr/local'
--
-- map:get(opts, "verbose") --> 1
-- map:get(opts, "hash") --> 0 (not supplied on command line)
-- map:get(opts, "output") --> "john.txt"
-- map:get(opts, "import") --> {"/usr/local", "/etc/app"}
-- map:get(opts, "extras") --> {"input1.txt", "input2.txt"}

```

8.1.2.4.9 See Also:

[show_help, command_line](#)

8.1.2.5 build_commandline

```

include std/cmdline.e
public function build_commandline(sequence cmds)

```

Returns a text string based on the set of supplied strings. Typically, this is used to ensure that arguments on a command line are properly formed before submitting it to the shell.

8.1.2.5.1 Parameters:

1. `cmds` : A sequence. Contains zero or more strings.

8.1.2.5.2 Returns:

A **sequence**, which is a text string. Each of the strings in `cmds` is quoted if they contain spaces, and then concatenated to form a single string.

8.1.2.5.3 Comments:

Though this function does the quoting for you it is not going to protect your programs from globbing *, ?. And it is not specied here what happens if you pass redirection or piping characters.

8.1.2.5.4 Example 1:

```
s = build_commandline( { "-d", "/usr/my docs/" } )  
-- s now contains '-d "/usr/my docs/'
```

8.1.2.5.5 Example 2:

8.1.2.5.6 You can use this to run things that might be difficult to quote out:

Suppose you want to run a program that requires quotes on its command line? Use this function to pass quotation marks:

```
s = build_commandline( { "awk", "-e", "'{ print $1\"x\"$2; }'" } )  
system(s,0)
```

8.1.2.5.7 See Also:

[parse_commandline](#), [system](#), [system_exec](#), [command_line](#)

8.1.2.6 parse_commandline

```
include std/cmdline.e  
public function parse_commandline(sequence cmdline)
```

Parse a command line string breaking it into a sequence of command line options.

8.1.2.6.1 Parameters:

1. `cmdline` : Command line sequence (string)

8.1.2.6.2 Returns:

A **sequence**, of command line options

8.1.2.6.3 Example 1:

```
sequence opts = parse_commandline("-v -f '%Y-%m-%d %H:%M'")
-- opts = { "-v", "-f", "%Y-%m-%d %H:%M" }
```

8.1.2.6.4 See Also:

[build_commandline](#)

8.2 Console

Cursor Style Constants

NO_CURSOR
UNDERLINE_CURSOR
THICK_UNDERLINE_CURSOR
HALF_BLOCK_CURSOR
BLOCK_CURSOR

Keyboard related routines

get_key
allow_break
check_break
wait_key
any_key
maybe_any_key
prompt_number
prompt_string

Cross Platform Text Graphics

positive_int
clear_screen
get_screen_char
put_screen_char
attr_to_colors
colors_to_attr
display_text_image
save_text_image
text_rows
cursor
free_console
display

8.2.1 Cursor Style Constants

In the cursor constants below, the second and fourth hex digits (from the left) determine the top and bottom row of pixels in the cursor. The first digit controls whether the cursor will be visible or not. For example, #0407 turns on the 4th through 7th rows.

8.2.1.1 See Also:

[cursor](#)

8.2.1.2 NO_CURSOR

```
include std/console.e
public constant NO_CURSOR
```

8.2.1.3 UNDERLINE_CURSOR

```
include std/console.e
public constant UNDERLINE_CURSOR
```

8.2.1.4 THICK_UNDERLINE_CURSOR

```
include std/console.e
public constant THICK_UNDERLINE_CURSOR
```

8.2.1.5 HALF_BLOCK_CURSOR

```
include std/console.e
public constant HALF_BLOCK_CURSOR
```

8.2.1.6 BLOCK_CURSOR

```
include std/console.e
public constant BLOCK_CURSOR
```

8.2.2 Keyboard related routines

8.2.2.1 get_key

```
<built-in> function get_key()
```

Return the key that was pressed by the user, without waiting. Special codes are returned for the function keys, arrow keys etc.

8.2.2.1.1 Returns:

An **integer**, either -1 if no key waiting, or the code of the next key waiting in keyboard buffer.

8.2.2.1.2 Comments:

The operating system can hold a small number of key-hits in its keyboard buffer. `get_key()` will return the next one from the buffer, or -1 if the buffer is empty.

Run the `key.bat` program to see what key code is generated for each key on your keyboard.

8.2.2.1.3 Example 1:

```
integer n = get_key()
if n=-1 then
    puts(1, "No key waiting.\n")
end if
```

8.2.2.1.4 See Also:

[wait_key](#)

8.2.2.2 allow_break

```
include std/console.e
public procedure allow_break(boolean b)
```

Set behavior of CTRL+C/CTRL+Break

8.2.2.2.1 Parameters:

1. `b` : a boolean, TRUE (`!= 0`) to enable the trapping of Ctrl-C/Ctrl-Break, FALSE (`0`) to disable it.

8.2.2.2.2 Comments:

When `b` is 1 (true), CTRL+C and CTRL+Break can terminate your program when it tries to read input from the keyboard. When `b` is 0 (false) your program will not be terminated by CTRL+C or CTRL+Break.

Initially your program can be terminated at any point where it tries to read from the keyboard.

You can find out if the user has pressed Control-C or Control-Break by calling [check_break\(\)](#).

8.2.2.2.3 Example 1:

```
allow_break(0)  -- don't let the user kill the program!
```

8.2.2.2.4 See Also:

[check_break](#)

8.2.2.3 check_break

```
include std/console.e
public function check_break()
```

Returns the number of Control-C/Control-BREAK key presses.

8.2.2.3.1 Returns:

An **integer**, the number of times that CTRL+C or CTRL+Break have been pressed since the last call to `check_break()`, or since the beginning of the program if this is the first call.

8.2.2.3.2 Comments:

This is useful after you have called [allow_break\(0\)](#) which prevents CTRL+C or CTRL+Break from terminating your program. You can use `check_break()` to find out if the user has pressed one of these keys. You might then perform some action such as a graceful shutdown of your program.

Neither CTRL+C or CTRL+Break will be returned as input characters when you read the keyboard. You can only detect them by calling `check_break()`.

8.2.2.3.3 Example 1:

```
k = get_key()
if check_break() then  -- ^C or ^Break was hit once or more
    temp = graphics_mode(-1)
    puts(STDOUT, "Shutting down...")
    save_all_user_data()
    abort(1)
end if
```

8.2.2.3.4 See Also:

[allow_break](#)

8.2.2.4 wait_key

```
include std/console.e
public function wait_key()
```

Waits for user to press a key, unless any is pending, and returns key code.

8.2.2.4.1 Returns:

An **integer**, which is a key code. If one is waiting in keyboard buffer, then return it. Otherwise, wait for one to come up.

8.2.2.4.2 See Also:

[get_key](#), [getc](#)

8.2.2.5 any_key

```
include std/console.e
public procedure any_key(sequence prompt = "Press Any Key to continue...", integer con = 1)
```

Display a prompt to the user and wait for any key.

8.2.2.5.1 Parameters:

1. `prompt` : Prompt to display, defaults to "Press Any Key to continue..."
2. `con` : Either 1 (stdout), or 2 (stderr). Defaults to 1.

8.2.2.5.2 Comments:

This wraps [wait_key](#) by giving a clue that the user should press a key, and perhaps do some other things as well.

8.2.2.5.3 Example 1:

```
any_key() -- "Press Any Key to continue..."
```

8.2.2.5.4 Example 2:

```
any_key("Press Any Key to quit")
```

8.2.2.5.5 See Also:

[wait_key](#)

8.2.2.6 maybe_any_key

```
include std/console.e
public procedure maybe_any_key(sequence prompt = "Press Any Key to continue...", integer con =
```

Display a prompt to the user and wait for any key **only** if the user is running under a GUI environment.

8.2.2.6.1 Parameters:

1. `prompt` : Prompt to display, defaults to "Press Any Key to continue..."
2. `con` : Either 1 (stdout), or 2 (stderr). Defaults to 1.

8.2.2.6.2 Comments:

This wraps [wait_key](#) by giving a clue that the user should press a key, and perhaps do some other things as well.

8.2.2.6.3 Example 1:

```
any_key() -- "Press Any Key to continue..."
```

8.2.2.6.4 Example 2:

```
any_key("Press Any Key to quit")
```

8.2.2.6.5 See Also:

[wait_key](#)

8.2.2.7 prompt_number

```
include std/console.e
public function prompt_number(sequence prompt, sequence range)
```

Prompts the user to enter a number, and returns only validated input.

8.2.2.7.1 Parameters:

1. `st` : is a string of text that will be displayed on the screen.
2. `s` : is a sequence of two values {lower, upper} which determine the range of values that the user may enter. `s` can be empty, {}, if there are no restrictions.

8.2.2.7.2 Returns:

An **atom**, in the assigned range which the user typed in.

8.2.2.7.3 Errors:

If `puts()` cannot display `st` on standard output, or if the first or second element of `s` is a sequence, a runtime error will be raised.

If user tries cancelling the prompt by hitting Ctrl-Z, the program will abort as well, issuing a type check error.

8.2.2.7.4 Comments:

As long as the user enters a number that is less than lower or greater than upper, the user will be prompted again.

If this routine is too simple for your needs, feel free to copy it and make your own more specialized version.

8.2.2.7.5 Example 1:

```
age = prompt_number("What is your age? ", {0, 150})
```

8.2.2.7.6 Example 2:

```
t = prompt_number("Enter a temperature in Celcius:\n", {})
```

8.2.2.7.7 See Also:

[puts](#), [prompt_string](#)

8.2.2.8 prompt_string

```
include std/console.e
public function prompt_string(sequence prompt)
```

Prompt the user to enter a string of text.

8.2.2.8.1 Parameters:

1. `st` : is a string that will be displayed on the screen.

8.2.2.8.2 Returns:

A **sequence**, the string that the user typed in, stripped of any new-line character.

8.2.2.8.3 Comments:

If the user happens to type control-Z (indicates end-of-file), "" will be returned.

8.2.2.8.4 Example 1:

```
name = prompt_string("What is your name? ")
```

8.2.2.8.5 See Also:

[prompt_number](#)

8.2.3 Cross Platform Text Graphics

8.2.3.1 positive_int

```
include std/console.e
public type positive_int(integer x)
```

8.2.3.2 clear_screen

```
<built-in> procedure clear_screen()
```

Clear the screen using the current background color (may be set by [bk_color\(\)](#)).

8.2.3.2.1 See Also:

[bk_color](#)

8.2.3.3 get_screen_char

```
include std/console.e
public function get_screen_char(positive_atom line, positive_atom column, integer fgbg = 0)
```

Get the value and attribute of the character at a given screen location.

8.2.3.3.1 Parameters:

1. `line` : the 1-base line number of the location
2. `column` : the 1-base column number of the location
3. `fgbg` : an integer, if 0 (the default) you get an `attribute_code` returned otherwise you get a foreground and background color number returned.

8.2.3.3.2 Returns:

- If `fgbg` is zero then a **sequence** of *two* elements, `{character, attribute_code}` for the specified location.
- If `fgbg` is not zero then a **sequence** of *three* elements, `{characterfg_color, bg_color}`

8.2.3.3.3 Comments:

- This function inspects a single character on the *active page*.
- The `attribute_code` is an atom that contains the foreground and background color of the character, and possibly other operating-system dependant information describing the appearance of the character on the screen.
- The `fg_color` and `bg_color` are integers in the range 0 to 15, which correspond to...

color number	name
0	black
1	dark blue
2	green
3	cyan
4	crimson
5	purple
6	brown
7	light gray
8	dark gray
9	blue
10	bright green
11	light blue
12	red
13	magenta
14	yellow

15 white

- With `get_screen_char()` and `put_screen_char()` you can save and restore a character on the screen along with its `attribute_code`.

8.2.3.3.4 Example 1:

```
-- read character and attributes at top left corner
s = get_screen_char(1,1)
-- s could be {'A', 92}
-- store character and attributes at line 25, column 10
put_screen_char(25, 10, s)
```

8.2.3.3.5 Example 2:

```
-- read character and colors at line 25, column 10.
s = get_screen_char(25,10, 1)
-- s could be {'A', 12, 5}
```

8.2.3.3.6 See Also:

[put_screen_char](#), [save_text_image](#)

8.2.3.4 put_screen_char

```
include std/console.e
public procedure put_screen_char(positive_atom line, positive_atom column, sequence char_attr)
```

Stores/displays a sequence of characters with attributes at a given location.

8.2.3.4.1 Parameters:

1. `line` : the 1-based line at which to start writing
2. `column` : the 1-based column at which to start writing
3. `char_attr` : a sequence of alternated characters and attribute codes.

8.2.3.4.2 Comments:

`char_attr` must be in the form {character, attribute code, character, attribute code, ...}.

8.2.3.4.3 Errors:

The length of `char_attr` must be a multiple of 2.

8.2.3.4.4 Comments:

The attributes atom contains the foreground color, background color, and possibly other platform-dependent information controlling how the character is displayed on the screen. If `char_attr` has 0 length, nothing will be written to the screen. The characters are written to the *active page*. It's faster to write several characters to the screen with a single call to `put_screen_char()` than it is to write one character at a time.

8.2.3.4.5 Example 1:

```
-- write AZ to the top left of the screen
-- (attributes are platform-dependent)
put_screen_char(1, 1, {'A', 152, 'Z', 131})
```

8.2.3.4.6 See Also:

[get_screen_char](#), [display_text_image](#)

8.2.3.5 attr_to_colors

```
include std/console.e
public function attr_to_colors(integer attr_code)
```

Converts an attribute code to its foreground and background color components.

8.2.3.5.1 Parameters:

1. `attr_code` : integer, an attribute code.

8.2.3.5.2 Returns:

A sequence of two elements - {fgcolor, bgcolor}

8.2.3.5.3 Example 1:

```
? attr_to_colors(92) --> {12, 5}
```

8.2.3.5.4 See Also:

[get_screen_char](#), [colors_to_attr](#)

8.2.3.6 colors_to_attr

```
include std/console.e
public function colors_to_attr(object fgbg, integer bg = 0)
```

Converts a foreground and background color set to its attribute code format.

8.2.3.6.1 Parameters:

1. `fgbg` : Either a sequence of {fgcolor, bgcolor} or just an integer fgcolor.
2. `bg` : An integer bgcolor. Only used when `fgbg` is an integer.

8.2.3.6.2 Returns:

An integer attribute code.

8.2.3.6.3 Example 1:

```
? colors_to_attr({12, 5}) --> 92
? colors_to_attr(12, 5) --> 92
```

8.2.3.6.4 See Also:

[get_screen_char](#), [put_screen_char](#), [attr_to_colors](#)

8.2.3.7 display_text_image

```
include std/console.e
public procedure display_text_image(text_point xy, sequence text)
```

Display a text image in any text mode.

8.2.3.7.1 Parameters:

1. `xy` : a pair of 1-based coordinates representing the point at which to start writing
2. `text` : a list of sequences of alternated character and attribute.

8.2.3.7.2 Comments:

This routine displays to the active text page, and only works in text modes.

You might use [save_text_image\(\)/display_text_image\(\)](#) in a text-mode graphical user interface, to allow "pop-up" dialog boxes, and drop-down menus to appear and disappear without losing what was previously on the screen.

8.2.3.7.3 Example 1:

```
clear_screen()
display_text_image({1,1}, {{ 'A', WHITE, 'B', GREEN},
                           { 'C', RED+16*WHITE},
                           { 'D', BLUE}})

-- displays:
--      AB
--      C
--      D
-- at the top left corner of the screen.
-- 'A' will be white with black (0) background color,
-- 'B' will be green on black,
-- 'C' will be red on white, and
-- 'D' will be blue on black.
```

8.2.3.7.4 See Also:

[save_text_image](#), [put_screen_char](#)

8.2.3.8 save_text_image

```
include std/console.e
public function save_text_image(text_point top_left, text_point bottom_right)
```

Copy a rectangular block of text out of screen memory

8.2.3.8.1 Parameters:

1. `top_left` : the coordinates, given as a pair, of the upper left corner of the area to save.
2. `bottom_right` : the coordinates, given as a pair, of the lower right corner of the area to save.

8.2.3.8.2 Returns:

A **sequence**, of {character, attribute, character, ...} lists.

8.2.3.8.3 Comments:

The returned value is appropriately handled by [display_text_image](#).

This routine reads from the active text page, and only works in text modes.

You might use this function in a text-mode graphical user interface to save a portion of the screen before displaying a drop-down menu, dialog box, alert box etc.

8.2.3.8.4 Example 1:

```
-- Top 2 lines are: Hello and World
s = save_text_image({1,1}, {2,5})

-- s is something like: {"H-e-l-l-o-", "W-o-r-l-d-"}
```

8.2.3.8.5 See Also:

[display_text_image](#), [get_screen_char](#)

8.2.3.9 text_rows

```
include std/console.e
public function text_rows(positive_int rows)
```

Set the number of lines on a text-mode screen.

8.2.3.9.1 Parameters:

1. `rows` : an integer, the desired number of rows.

8.2.3.9.2 Platforms:

Not *Unix*

8.2.3.9.3 Returns:

An **integer**, the actual number of text lines.

8.2.3.9.4 Comments:

Values of 25, 28, 43 and 50 lines are supported by most video cards.

8.2.3.9.5 See Also:

[graphics_mode](#), [video_config](#)

8.2.3.10 cursor

```
include std/console.e
public procedure cursor(integer style)
```

Select a style of cursor.

8.2.3.10.1 Parameters:

1. `style` : an integer defining the cursor shape.

8.2.3.10.2 Platform:

Not *Unix*

8.2.3.10.3 Comments:

In pixel-graphics modes no cursor is displayed.

8.2.3.10.4 Example 1:

```
cursor(BLOCK_CURSOR)
```

8.2.3.10.5 Cursor Type Constants:

- [NO_CURSOR](#)
- [UNDERLINE_CURSOR](#)
- [THICK_UNDERLINE_CURSOR](#)
- [HALF_BLOCK_CURSOR](#)
- [BLOCK_CURSOR](#)

8.2.3.10.6 See Also:

[graphics_mode](#), [text_rows](#)

8.2.3.11 free_console

```
include std/console.e
public procedure free_console()
```

Free (delete) any console window associated with your program.

8.2.3.11.1 Comments:

EUPHORIA will create a console text window for your program the first time that your program prints something to the screen, reads something from the keyboard, or in some way needs a console. On WIN32 this window will automatically disappear when your program terminates, but you can call `free_console()` to make it disappear sooner. On Linux or FreeBSD, the text mode console is always there, but an xterm window will disappear after EUPHORIA issues a "Press Enter" prompt at the end of execution.

On Unix-style systems, `free_console()` will set the terminal parameters back to normal, undoing the effect that curses has on the screen.

In an xterm window, a call to `free_console()`, without any further printing to the screen or reading from the keyboard, will eliminate the "Press Enter" prompt that EUPHORIA normally issues at the end of execution.

After freeing the console window, you can create a new console window by printing something to the screen, or simply calling `clear_screen()`, `position()` or any other routine that needs a console.

When you use the trace facility, or when your program has an error, EUPHORIA will automatically create a console window to display trace information, error messages etc.

There's a WIN32 API routine, `FreeConsole()` that does something similar to `free_console()`. You should use `free_console()` instead, because it lets the interpreter know that there is no longer a console to write to or read from.

8.2.3.11.2 See Also:

[clear_screen](#)

8.2.3.12 display

```
include std/console.e
public procedure display(object data_in, object args = 1, integer finalnl = - 918273645)
```

Displays the supplied data on the console screen at the current cursor position.

8.2.3.12.1 Parameters:

1. `data_in` : Any object.
2. `args` : Optional arguments used to format the output. Default is 1.
3. `finalnl` : Optional. Determines if a new line is output after the data. Default is to output a new line.

8.2.3.12.2 Comments:

- If `data_in` is an atom or integer, it is simply displayed.
- If `data_in` is a simple text string, then `args` can be used to produce a formatted output with `data_in` providing the `format` string and `args` being a sequence containing the data to be formatted.
 - ◆ If the last character of `data_in` is an underscore character then it is stripped off and `finalnl` is set to zero. Thus ensuring that a new line is **not** output.
 - ◆ The formatting codes expected in `data_in` are the ones used by `format`. It is not mandatory to use formatting codes, and if `data_in` does not contain any then it is simply displayed and anything in `args` is ignored.
- If `data_in` is a sequence containing floating-point numbers, sub-sequences or integers that are not characters, then `data_in` is forwarded on to the `pretty_print()` to display.
 - ◆ If `args` is a non-empty sequence, it is assumed to contain the `pretty_print` formatting options.
 - ◆ if `args` is an atom or an empty sequence, the assumed `pretty_print` formatting options are assumed to be `{2}`.

After the data is displayed, the routine will normally output a New Line. If you want to avoid this, ensure that the last parameter is a zero. Or to put this another way, if the last parameter is zero then a New Line will **not** be output.

8.2.3.12.3 Examples:

```
display("Some plain text") -- Displays this string on the console plus a new line.
display("Your answer:",0) -- Displays this string on the console without a new line.
display("cat")
display("Your answer:",,0) -- Displays this string on the console without a new line.
display("")
display("Your answer:_") -- Displays this string, except the '_', on the console without a new line.
display("dog")
display({"abc", 3.44554}) -- Displays the contents of 'res' on the console.
display("The answer to [1] was [2]", {"'why'", 42}) -- formats these with a new line.
display("",2)
display({51,362,71}, {1})
```

Output would be ...

```
Some plain text
Your answer:cat
===== Your answer:
Your answer:dog
{
  "abc",
  3.44554
```

```
}  
The answer to 'why' was 42  
""  
{51'3',362,71'G'}
```

8.3 Date/Time

Localized Variables

- month_names
- month_abbrs
- day_names
- day_abbrs
- ampm

Constants

- YEAR
- YEARS

Types

- datetime

Routines

- time
- date
- from_date
- now
- now_gmt
- new
- new_time
- weeks_day
- years_day
- is_leap_year
- days_in_month
- days_in_year
- to_unix
- from_unix
- format
- parse
- add
- subtract
- diff

8.3.1 Localized Variables

8.3.1.1 month_names

```
include std/datetime.e  
public sequence month_names
```

Names of the months

8.3.1.2 month_abbrs

```
include std/datetime.e
public sequence month_abbrs
```

Abbreviations of month names

8.3.1.3 day_names

```
include std/datetime.e
public sequence day_names
```

Names of the days

8.3.1.4 day_abbrs

```
include std/datetime.e
public sequence day_abbrs
```

Abbreviations of day names

8.3.1.5 ampm

```
include std/datetime.e
public sequence ampm
```

AM/PM

8.3.2 Constants

8.3.2.1 YEAR

```
include std/datetime.e
public enum YEAR
```

Accessors



- YEAR
- MONTH
- DAY
- HOUR
- MINUTE
- SECOND

8.3.2.2 YEARS

```
include std/datetime.e
public enum YEARS
```

Intervals

- YEARS
- MONTHS
- WEEKS
- DAYS
- HOURS
- MINUTES
- SECONDS
- DATE

8.3.3 Types

8.3.3.1 datetime

```
include std/datetime.e
public type datetime(object o)
```

datetime type

8.3.3.1.1 Parameters:

1. `obj` : any object, so no crash takes place.

8.3.3.1.2 Comments:

A datetime type consists of a sequence of length 6 in the form {year, month, day_of_month, hour, minute, second}. Checks are made to guarantee those values are in range.

8.3.3.1.3 Note:

All components must be integers except seconds, as those can also be floating point values.

8.3.4 Routines

8.3.4.1 time

<built-in> `function time()`

Return the number of seconds since some fixed point in the past.

8.3.4.1.1 Returns:

An **atom**, which represents an absolute number of seconds.

8.3.4.1.2 Comments:

Take the difference between two readings of `time()`, to measure, for example, how long a section of code takes to execute.

On some machines, `time()` can return a negative number. However, you can still use the difference in calls to `time()` to measure elapsed time.

8.3.4.1.3 Example 1:

```
constant ITERATIONS = 1000000
integer p
atom t0, loop_overhead

t0 = time()
for i = 1 to ITERATIONS do
    -- time an empty loop
end for
loop_overhead = time() - t0

t0 = time()
for i = 1 to ITERATIONS do
    p = power(2, 20)
end for
? (time() - t0 - loop_overhead)/ITERATIONS
-- calculates time (in seconds) for one call to power
```

8.3.4.1.4 See Also:

[date](#), [now](#)

8.3.4.2 date

```
<built-in> function date()
```

Return a sequence with information on the current date.

8.3.4.2.1 Returns:

A **sequence** of length 8, laid out as follows:

1. year -- since 1900
2. month -- January = 1
3. day -- day of month, starting at 1
4. hour -- 0 to 23
5. minute -- 0 to 59
6. second -- 0 to 59
7. day of the week -- Sunday = 1
8. day of the year -- January 1st = 1

8.3.4.2.2 Comments:

The value returned for the year is actually the number of years since 1900 (not the last 2 digits of the year). In the year 2000 this value was 100. In 2001 it was 101, etc.

8.3.4.2.3 Example 1:

```
now = date()
-- now has: {95,3,24,23,47,38,6,83}
-- i.e. Friday March 24, 1995 at 11:47:38pm, day 83 of the year
```

8.3.4.2.4 See Also:

[time](#), [now](#)

8.3.4.3 from_date

```
include std/datetime.e
public function from_date(sequence src)
```

Convert a sequence formatted according to the built-in `date()` function to a valid `datetime` sequence.

8.3.4.3.1 Parameters:

1. `src` : a sequence which `date()` might have returned

8.3.4.3.2 Returns:

A **sequence**, more precisely a **datetime** corresponding to the same moment in time.

8.3.4.3.3 Example 1:

```
d = from_date(date())  
-- d is the current date and time
```

8.3.4.3.4 See Also:

[date](#), [from_unix](#), [now](#), [new](#)

8.3.4.4 now

```
include std/datetime.e  
public function now()
```

Create a new `datetime` value initialized with the current date and time

8.3.4.4.1 Returns:

A **sequence**, more precisely a **datetime** corresponding to the current moment in time.

8.3.4.4.2 Example 1:

```
dt = now()  
-- dt is the current date and time
```

8.3.4.4.3 See Also:

[from_date](#), [from_unix](#), [new](#), [new_time](#), [now_gmt](#)

8.3.4.5 now_gmt

```
include std/datetime.e
public function now_gmt()
```

Create a new datetime value that falls into the Greenwich Mean Time (GMT) timezone. This function will return a datetime that is GMT, no matter what timezone the system is running under.

8.3.4.5.1 Example 1:

```
dt = now_gmt()
-- If local time was July 16th, 2008 at 10:34pm CST
-- dt would be July 17th, 2008 at 03:34pm GMT
```

8.3.4.5.2 See Also:

[now](#)

8.3.4.6 new

```
include std/datetime.e
public function new(integer year = 0, integer month = 0, integer day = 0, integer hour = 0, integer minute = 0, integer second = 0)
```

Create a new datetime value.

8.3.4.6.1 Parameters:

1. `year` -- the full year.
2. `month` -- the month (1-12).
3. `day` -- the day of the month (1-31).
4. `hour` -- the hour (0-23) (defaults to 0)
5. `minute` -- the minute (0-59) (defaults to 0)
6. `second` -- the second (0-59) (defaults to 0)

8.3.4.6.2 Example 1:

```
dt = new(2010, 1, 1, 0, 0, 0)
-- dt is Jan 1st, 2010
```

8.3.4.6.3 See Also:

[from_date](#), [from_unix](#), [now](#), [new_time](#)

8.3.4.7 new_time

```
include std/datetime.e
public function new_time(integer hour, integer minute, atom second)
```

Create a new datetime value with a date of zeros.

8.3.4.7.1 Parameters:

1. `hour` : is the hour (0-23)
2. `minute` : is the minute (0-59)
3. `second` : is the second (0-59)

8.3.4.7.2 Example 1:

```
dt = new_time(10, 30, 55)
dt is 10:30:55 AM
```

8.3.4.7.3 See Also:

[from_date](#), [from_unix](#), [now](#), [new](#)

8.3.4.8 weeks_day

```
include std/datetime.e
public function weeks_day(datetime dt)
```

Get the day of week of the datetime dt.

8.3.4.8.1 Parameters:

1. `dt` : a datetime to be queried.

8.3.4.8.2 Returns:

An **integer**, between 1 (Sunday) and 7 (Saturday).

8.3.4.8.3 Example 1:

```
d = new(2008, 5, 2, 0, 0, 0)
day = weeks_day(d) -- day is 6 because May 2, 2008 is a Friday.
```

8.3.4.9 years_day

```
include std/datetime.e
public function years_day(datetime dt)
```

Get the Julian day of year of the supplied date.

8.3.4.9.1 Parameters:

1. dt : a datetime to be queried.

8.3.4.9.2 Returns:

An **integer**, between 1 and 366.

8.3.4.9.3 Comments:

For dates earlier than 1800, this routine may give inaccurate results if the date applies to a country other than United Kingdom or a former colony thereof. The change from Julian to Gregorian calendar took place much earlier in some other European countries.

8.3.4.9.4 Example 1:

```
d = new(2008, 5, 2, 0, 0, 0)
day = years_day(d) -- day is 123
```

8.3.4.10 is_leap_year

```
include std/datetime.e
public function is_leap_year(datetime dt)
```

Determine if dt falls within leap year.

8.3.4.10.1 Parameters:

1. dt : a datetime to be queried.

8.3.4.10.2 Returns:

An **integer**, of 1 if leap year, otherwise 0.

8.3.4.10.3 Example 1:

```
d = new(2008, 1, 1, 0, 0, 0)
? is_leap_year(d) -- prints 1
d = new(2005, 1, 1, 0, 0, 0)
? is_leap_year(d) -- prints 0
```

8.3.4.10.4 See Also:

[days_in_month](#)

8.3.4.11 days_in_month

```
include std/datetime.e
public function days_in_month(datetime dt)
```

Return the number of days in the month of dt.

This takes into account leap year.

8.3.4.11.1 Parameters:

1. dt : a datetime to be queried.

8.3.4.11.2 Example 1:

```
d = new(2008, 1, 1, 0, 0, 0)
? days_in_month(d) -- 31
d = new(2008, 2, 1, 0, 0, 0) -- Leap year
? days_in_month(d) -- 29
```

8.3.4.11.3 See Also:

[is_leap_year](#)

8.3.4.12 days_in_year

```
include std/datetime.e
public function days_in_year(datetime dt)
```

Return the number of days in the year of dt.

This takes into account leap year.

8.3.4.12.1 Parameters:

1. dt : a datetime to be queried.

8.3.4.12.2 Example 1:

```
d = new(2007, 1, 1, 0, 0, 0)
? days_in_year(d) -- 365
d = new(2008, 1, 1, 0, 0, 0) -- leap year
? days_in_year(d) -- 366
```

8.3.4.12.3 See Also:

[is_leap_year](#), [days_in_month](#)

8.3.4.13 to_unix

```
include std/datetime.e
public function to_unix(datetime dt)
```

Convert a datetime value to the unix numeric format (seconds since EPOCH_1970)

8.3.4.13.1 Parameters:

1. dt : a datetime to be queried.

8.3.4.13.2 Returns:

An **atom**, so this will not overflow during the winter 2038-2039.

8.3.4.13.3 Example 1:

```
secs_since_epoch = to_unix(now())
-- secs_since_epoch is equal to the current seconds since epoch
```

8.3.4.13.4 See Also:

[from_unix](#), [format](#)

8.3.4.14 from_unix

```
include std/datetime.e
public function from_unix(atom unix)
```

Create a datetime value from the unix numeric format (seconds since EPOCH)

8.3.4.14.1 Parameters:

1. `unix` : an atom, counting seconds elapsed since EPOCH.

8.3.4.14.2 Returns:

A **sequence**, more precisely a **datetime** representing the same moment in time.

8.3.4.14.3 Example 1:

```
d = from_unix(0)
-- d is 1970-01-01 00:00:00 (zero seconds since EPOCH)
```

8.3.4.14.4 See Also:

[to_unix](#), [from_date](#), [now](#), [new](#)

8.3.4.15 format

```
include std/datetime.e
public function format(datetime d, sequence pattern = "%Y-%m-%d %H:%M:%S")
```

Format the date according to the format pattern string

8.3.4.15.1 Parameters:

1. `d` : a datetime which is to be printed out
2. `pattern` : a format string, similar to the ones `sprintf()` uses, but with some Unicode encoding. The default is "%Y-%m-%d %H:%M:%S".

8.3.4.15.2 Returns:

A **string**, with the date `d` formatted according to the specification in `pattern`.

8.3.4.15.3 Comments:

Pattern string can include the following specifiers:

- %% -- a literal %
- %a -- locale's abbreviated weekday name (e.g., Sun)
- %A -- locale's full weekday name (e.g., Sunday)
- %b -- locale's abbreviated month name (e.g., Jan)
- %B -- locale's full month name (e.g., January)
- %C -- century; like %Y, except omit last two digits (e.g., 21)
- %d -- day of month (e.g., 01)
- %H -- hour (00..23)
- %I -- hour (01..12)
- %j -- day of year (001..366)
- %k -- hour (0..23)
- %l -- hour (1..12)
- %m -- month (01..12)
- %M -- minute (00..59)
- %p -- locale's equivalent of either AM or PM; blank if not known
- %P -- like %p, but lower case
- %s -- seconds since 1970-01-01 00:00:00 UTC
- %S -- second (00..60)
- %u -- day of week (1..7); 1 is Monday
- %w -- day of week (0..6); 0 is Sunday
- %y -- last two digits of year (00..99)
- %Y -- year

8.3.4.15.4 Example 1:

```
d = new(2008, 5, 2, 12, 58, 32)
s = format(d, "%Y-%m-%d %H:%M:%S")
-- s is "2008-05-02 12:58:32"
```

8.3.4.15.5 Example 2:

```
d = new(2008, 5, 2, 12, 58, 32)
s = format(d, "%A, %B %d '%y %H:%M%p")
-- s is "Friday, May 2 '08 12:58PM"
```

8.3.4.15.6 See Also:

[to_unix](#), [parse](#)

8.3.4.16 parse

```
include std/datetime.e
public function parse(sequence val, sequence fmt = "%Y-%m-%d %H:%M:%S")
```

Parse a datetime string according to the given format.

8.3.4.16.1 Parameters:

1. `val` : string datetime value
2. `fmt` : datetime format. Default is "%Y-%m-%d %H:%M:%S"

8.3.4.16.2 Returns:

A **datetime**, value.

8.3.4.16.3 Comments:

8.3.4.16.4 Only a subset of the format specification is currently supported:

- %d -- day of month (e.g, 01)
- %H -- hour (00..23)
- %m -- month (01..12)
- %M -- minute (00..59)
- %S -- second (00..60)
- %Y -- year

More format codes will be added in future versions.

All non-format characters in the format string are ignored and are not matched against the input string.

All non-digits in the input string are ignored.

8.3.4.16.5 Example 1:

```
datetime d = parse("05/01/2009 10:20:30", "%m/%d/%Y %H:%M:%S")
```

8.3.4.16.6 See Also:

[format](#)

8.3.4.17 add

```
include std/datetime.e
public function add(datetime dt, object qty, integer interval)
```

Add a number of *intervals* to a datetime.

8.3.4.17.1 Parameters:

1. dt : the base datetime
2. qty : the number of *intervals* to add. It should be positive.
3. interval : which kind of interval to add.

8.3.4.17.2 Returns:

A **sequence**, more precisely a **datetime** representing the new moment in time.

8.3.4.17.3 Comments:

Please see Constants for Date/Time for a reference of valid intervals.

Do not confuse the item access constants such as YEAR, MONTH, DAY, etc... with the interval constants YEARS, MONTHS, DAYS, etc...

When adding MONTHS, it is a calendar based addition. For instance, a date of 5/2/2008 with 5 MONTHS added will become 10/2/2008. MONTHS does not compute the number of days per each month and the average number of days per month.

When adding YEARS, leap year is taken into account. Adding 4 YEARS to a date may result in a different day of month number due to leap year.

8.3.4.17.4 Example 1:

```
d2 = add(d1, 35, SECONDS) -- add 35 seconds to d1
d2 = add(d1, 7, WEEKS)    -- add 7 weeks to d1
d2 = add(d1, 19, YEARS)   -- add 19 years to d1
```

8.3.4.17.5 See Also:

[subtract](#), [diff](#)

8.3.4.18 subtract

```
include std/datetime.e
public function subtract(datetime dt, atom qty, integer interval)
```

Subtract a number of *intervals* to a base datetime.

8.3.4.18.1 Parameters:

1. dt : the base datetime
2. qty : the number of *intervals* to subtract. It should be positive.
3. interval : which kind of interval to subtract.

8.3.4.18.2 Returns:

A **sequence**, more precisely a **datetime** representing the new moment in time.

8.3.4.18.3 Comments:

Please see Constants for Date/Time for a reference of valid intervals.

See the function `add()` for more information on adding and subtracting date intervals

8.3.4.18.4 Example 1:

```
dt2 = subtract(dt1, 18, MINUTES) -- subtract 18 minutes from dt1
dt2 = subtract(dt1, 7, MONTHS)   -- subtract 7 months from dt1
dt2 = subtract(dt1, 12, HOURS)   -- subtract 12 hours from dt1
```

8.3.4.18.5 See Also:

[add](#), [diff](#)

8.3.4.19 diff

```
include std/datetime.e
public function diff(datetime dt1, datetime dt2)
```

Compute the difference, in seconds, between two dates.

8.3.4.19.1 Parameters:

1. dt1 : the end datetime
2. dt2 : the start datetime

8.3.4.19.2 Returns:

An **atom**, the number of seconds elapsed from dt2 to dt1.

8.3.4.19.3 Comments:

dt2 is subtracted from dt1, therefore, you can come up with a negative value.

8.3.4.19.4 Example 1:

```
d1 = now()
sleep(15)  -- sleep for 15 seconds
d2 = now()

i = diff(d1, d2)  -- i is 15
```

8.3.4.19.5 See Also:

[add](#), [subtract](#)

8.4 File System

Cross platform file operations for EUPHORIA

Constants

SLASH
SLASHES
EOLSEP
EOL
PATHSEP
NULLDEVICE
SHARED_LIB_EXT

Directory Handling

D_NAME
D_ATTRIBUTES
D_SIZE
D_YEAR
D_MONTH
D_DAY
D_HOUR

D_MINUTE
D_SECOND
W_BAD_PATH
dir
current_dir
chdir
my_dir
walk_dir
create_directory
delete_file
curdir
init_curdir
clear_directory
remove_directory

File name parsing

PATH_DIR
PATH_FILENAME
PATH_BASENAME
PATH_FILEEXT
PATH_DRIVEID
pathinfo
dirname
filename
filebase
fileext
driveid
defaulttext
absolute_path
canonical_path

File Types

FILETYPE_UNDEFINED
FILETYPE_NOT_FOUND
FILETYPE_FILE
FILETYPE_DIRECTORY
file_type

File Handling

SECTORS_PER_CLUSTER
BYTES_PER_SECTOR
NUMBER_OF_FREE_CLUSTERS
TOTAL_NUMBER_OF_CLUSTERS
TOTAL_BYTES
FREE_BYTES
USED_BYTES
COUNT_DIRS
COUNT_FILES
COUNT_SIZE
COUNT_TYPES
EXT_NAME
EXT_COUNT

EXT_SIZE
file_exists
file_timestamp
copy_file
rename_file
move_file
file_length
locate_file
disk_metrics
disk_size
dir_size
temp_file

8.4.1 Constants

8.4.1.1 SLASH

```
public constant SLASH
```

Current platform's path separator character

8.4.1.1.1 Comments:

When on *Windows*, `'\'`. When on *Unix*, `'/'`.

8.4.1.2 SLASHES

```
public constant SLASHES
```

Current platform's possible path separators. This is slightly different in that on *Windows* the path separators variable contains `\\` as well as `:` and `/` as newer *Windows* versions support `/` as a path separator. On *Unix* systems, it only contains `/`.

8.4.1.3 EOLSEP

```
public constant EOLSEP
```

Current platform's newline string: `"\n"` on *Unix*, else `"\r\n"`.

8.4.1.4 EOL

```
public constant EOL
```

All platform's newline character: `'\n'`. When text lines are read the native platform's EOLSEP string is replaced by a single character EOL.

8.4.1.5 PATHSEP

```
public constant PATHSEP
```

Current platform's path separator character: `:` on *Unix*, else `;`.

8.4.1.6 NULLDEVICE

```
public constant NULLDEVICE
```

Current platform's null device path: `/dev/null` on *Unix*, else `NUL:`.

8.4.1.7 SHARED_LIB_EXT

```
public constant SHARED_LIB_EXT
```

Current platform's shared library extension. For instance it can be `dll`, `so` or `dylib` depending on the platform.

8.4.2 Directory Handling

8.4.2.1 D_NAME

```
include std/filesys.e
public enum D_NAME
```

8.4.2.2 D_ATTRIBUTES

```
include std/filesys.e
public enum D_ATTRIBUTES
```

8.4.2.3 D_SIZE

```
include std/filesys.e
public enum D_SIZE
```

8.4.2.4 D_YEAR

```
include std/filesys.e
public enum D_YEAR
```

8.4.2.5 D_MONTH

```
include std/filesys.e
public enum D_MONTH
```

8.4.2.6 D_DAY

```
include std/filesys.e
public enum D_DAY
```

8.4.2.7 D_HOUR

```
include std/filesys.e
public enum D_HOUR
```

8.4.2.8 D_MINUTE

```
include std/filesys.e
public enum D_MINUTE
```

8.4.2.9 D_SECOND

```
include std/filesys.e
public enum D_SECOND
```

8.4.2.10 W_BAD_PATH

```
include std/filesys.e
public constant W_BAD_PATH
```

Bad path error code. See [walk_dir](#)

8.4.2.11 dir

```
include std/filesys.e
public function dir(sequence name)
```

Return directory information for the specified file or directory.

8.4.2.11.1 Parameters:

1. `name` : a sequence, the name to be looked up in the file system.

8.4.2.11.2 Returns:

An **object**, -1 if no match found, else a sequence of sequence entries

8.4.2.11.3 Errors:

The length of `name` should not exceed 1,024 characters.

8.4.2.11.4 Comments:

`name` can also contain * and ? wildcards to select multiple files.

The returned information is similar to what you would get from the DIR command. A sequence is returned where each element is a sequence that describes one file or subdirectory.

If `name` refers to a **directory** you may have entries for "." and "..", just as with the DIR command. If it refers to an existing **file**, and has no wildcards, then the returned sequence will have just one entry, i.e. its length will be 1. If `name` contains wildcards you may have multiple entries.

Each entry contains the name, attributes and file size as well as the year, month, day, hour, minute and second of the last modification.

8.4.2.11.5 You can refer to the elements of an entry with the following constants:

```
public constant
  -- File Attributes
  D_NAME      = 1,
  D_ATTRIBUTES = 2,
  D_SIZE      = 3,
  D_YEAR      = 4,
  D_MONTH     = 5,
  D_DAY       = 6,
  D_HOUR      = 7,
  D_MINUTE    = 8,
  D_SECOND    = 9
```

8.4.2.11.6 The attributes element is a string sequence containing characters chosen from:

Attribute	Description
'd'	directory
'r'	read only file
'h'	hidden file
's'	system file
'v'	volume-id entry
'a'	archive file

A normal file without special attributes would just have an empty string, "", in this field.

The top level directory, e.g. c:\ does not have "." or ".." entries.

This function is often used just to test if a file or directory exists.

Under *WIN32*, st can have a long file or directory name anywhere in the path.

Under *Unix*, the only attribute currently available is 'd'.

WIN32: The file name returned in D_NAME will be a long file name.

8.4.2.11.7 Example 1:

```
d = dir(current_dir())

-- d might have:
-- {
--   {".", "d", 0 1994, 1, 18, 9, 30, 02},
--   {"..", "d", 0 1994, 1, 18, 9, 20, 14},
--   {"fred", "ra", 2350, 1994, 1, 22, 17, 22, 40},
--   {"sub", "d", 0, 1993, 9, 20, 8, 50, 12}
-- }

d[3][D_NAME] would be "fred"
```

8.4.2.11.8 See Also:

[walk_dir](#)

8.4.2.12 current_dir

```
include std/filesys.e
public function current_dir()
```

Return the name of the current working directory.

8.4.2.12.1 Returns:

A **sequence**, the name of the current working directory

8.4.2.12.2 Comments:

There will be no slash or backslash on the end of the current directory, except under *Windows*, at the top-level of a drive, e.g. C:\

8.4.2.12.3 Example 1:

```
sequence s
s = current_dir()
-- s would have "C:\EUPHORIA\DOC" if you were in that directory
```

8.4.2.12.4 See Also:

[dir](#), [chdir](#)

8.4.2.13 chdir

```
include std/filesys.e
public function chdir(sequence newdir)
```

Set a new value for the current directory

8.4.2.13.1 Parameters:

`newdir` : a sequence, the name for the new working directory.

8.4.2.13.2 Returns:

An **integer**, 0 on failure, 1 on success.

8.4.2.13.3 Comments:

By setting the current directory, you can refer to files in that directory using just the file name.

The `current_dir()` function will return the name of the current directory.

On *WIN32* the current directory is a public property shared by all the processes running under one shell. On *Unix* a subprocess can change the current directory for itself, but this won't affect the current directory of its parent process.

8.4.2.13.4 Example 1:

```
if chdir("c:\\euphoria") then
    f = open("readme.doc", "r")
else
    puts(STDERR, "Error: No euphoria directory?\n")
end if
```

8.4.2.13.5 See Also:

`current_dir`, `dir`

8.4.2.14 my_dir

```
include std/filesys.e
public integer my_dir
```

Deprecated, so therefore not documented.

8.4.2.15 walk_dir

```
include std/filesys.e
public function walk_dir(sequence path_name, object your_function, integer scan_subdirs = FALSE)
```

Generalized Directory Walker

8.4.2.15.1 Parameters:

1. `path_name` : a sequence, the name of the directory to walk through
2. `your_function` : the routine id of a function that will receive each path returned from the result of `dir_source`, one at a time.
3. `scan_subdirs` : an optional integer, 1 to also walk though subfolders, 0 (the default) to skip them all.
4. `dir_source` : an optional integer. A routine_id of a user-defined routine that returns the list of paths to pass to `your_function`. If omitted, the `dir()` function is used.

8.4.2.15.2 Returns:

An object,

- 0 on success
- `W_BAD_PATH`: an error occurred
- anything else: the custom function returned something to stop `walk_dir()`.

8.4.2.15.3 Comments:

This routine will "walk" through a directory named `path_name`. For each entry in the directory, it will call a function, whose routine_id is `your_function`. If `scan_subdirs` is non-zero (TRUE), then the subdirectories in `path_name` will be walked through recursively in the very same way.

The routine that you supply should accept two sequences, the path name and `dir()` entry for each file and subdirectory. It should return 0 to keep going, or non-zero to stop `walk_dir()`. Returning `W_BAD_PATH` is taken as denoting some error.

This mechanism allows you to write a simple function that handles one file at a time, while `walk_dir()` handles the process of walking through all the files and subdirectories.

By default, the files and subdirectories will be visited in alphabetical order. To use a different order, use the `dir_source` to pass the routine_id of your own modified `dir` function that sorts the directory entries differently.

The path that you supply to `walk_dir()` must not contain wildcards (* or ?). Only a single directory (and its subdirectories) can be searched at one time.

For non-unix systems, any '/' characters in `path_name` are replaced with '\\

All trailing slash and whitespace characters are removed from `path_name`.

8.4.2.15.4 Example 1:

```
function look_at(sequence path_name, sequence item)
-- this function accepts two sequences as arguments
-- it displays all C/C++ source files and their sizes
```



```

    if find('d', item[D_ATTRIBUTES]) then
        return 0 -- Ignore directories
    end if
    if not find(fileext(item[D_NAME]), {"c,h,cpp,hpp,cp"}) then
        return 0 -- ignore non-C/C++ files
    end if
    printf(STDOUT, "%s%s%s: %d\n",
        {path_name, SLASH, item[D_NAME], item[D_SIZE]})
    return 0 -- keep going
end function

function mysort(sequence path)
    dobject
        (path)dir
    atom(dif then
        d      return
    end if
    -- Sort in descending file size.
    return sort_columns(d, {-D_SIZE})
end function

exit_code = walk_dir("C:\\MYFILES\\", routine_id("look_at"), TRUE, routine_id("mysort"))

```

8.4.2.15.5 See Also:

[dir](#), [sort](#), [sort_columns](#)

8.4.2.16 create_directory

```

include std/filesys.e
public function create_directory(sequence name, integer mode = 448, integer mkparent = 1)

```

Create a new directory.

8.4.2.16.1 Parameters:

1. name : a sequence, the name of the new directory to create
2. mode : on *Unix* systems, permissions for the new directory. Default is 448 (all rights for owner, none for others).
3. mkparent : If true (default) the parent directories are also created if needed.

8.4.2.16.2 Returns:

An **integer**, 0 on failure, 1 on success.

8.4.2.16.3 Comments:

mode is ignored on non-Unix platforms.

8.4.2.16.4 Example 1:

```
if not create_directory("the_new_folder") then
    ("Fatal system problem - could not create the new folder")
end if

-- This example will also create "myapp/" and "myapp/interface/" if they don't exist.
if not create_directory("myapp/interface/letters") then
    ("Fatal system problem - could not create the new folder")
end if

-- This example will NOT create "myapp/" and "myapp/interface/" if they don't exist.
if not create_directory("myapp/interface/letters",,0) then
    ("Fatal system problem - could not create the new folder")
end if
```

8.4.2.16.5 See Also:

[remove_directory](#), [chdir](#)

8.4.2.17 delete_file

```
include std/filesys.e
public function delete_file(sequence name)
```

Delete a file.

8.4.2.17.1 Parameters:

1. name : a sequence, the name of the file to delete.

8.4.2.17.2 Returns:

An integer, 0 on failure, 1 on success.

8.4.2.18 curdir

```
include std/filesys.e
public function curdir(integer drive_id = 0)
```

Returns the current directory, with a trailing SLASH

8.4.2.18.1 Parameters:

1. `drive_id`: For non-Unix systems only. This is the Drive letter to to get the current directory of. If omitted, the current drive is used.

8.4.2.18.2 Returns:

A **sequence**, the current directory.

8.4.2.18.3 Comment:

Windows maintain a current directory for each disk drive. You would use this routine if you wanted the current directory for a drive that may not be the current drive.

For Unix systems, this is simply ignored because there is only one current directory at any time on Unix.

8.4.2.18.4 Note:

This always ensures that the returned value has a trailing SLASH character.

8.4.2.18.5 Example 1:

```
res = curdir('D') -- Find the current directory on the D: drive.  
-- res might be "D:\backup\music\  
res = curdir()   -- Find the current directory on the current drive.  
-- res might be "C:\myapp\work\"
```

8.4.2.19 `init_curdir`

```
include std/filesys.e  
public function init_curdir()
```

Returns the original current directory

8.4.2.19.1 Parameters:

1. None.

8.4.2.19.2 Returns:

A **sequence**, the current directory at the time the program started running.

8.4.2.19.3 Comment:

You would use this if the program might change the current directory during its processing and you wanted to return to the original directory.

8.4.2.19.4 Note:

This always ensures that the returned value has a trailing SLASH character.

8.4.2.19.5 Example 1:

```
res = init_curdir() -- Find the original current directory.
```

8.4.2.20 clear_directory

```
include std/filesys.e
public function clear_directory(sequence path, integer recurse = 1)
```

Clear (delete) a directory of all files, but retaining sub-directories.

8.4.2.20.1 Parameters:

1. **name** : a sequence, the name of the directory whose files you want to remove.
2. **recurse** : an integer, whether or not to remove files in the directory's sub-directories. If 0 then this function is identical to `remove_directory()`. If 1, then we recursively delete the directory and its contents. Defaults to 1.

8.4.2.20.2 Returns:

An **integer**, 0 on failure, otherwise the number of files plus 1.

8.4.2.20.3 Comment:

This never removes a directory. It only ever removes files. It is used to clear a directory structure of all existing files, leaving the structure intact.

8.4.2.20.4 Example 1:

```
integer cnt = clear_directory("the_old_folder")
if cnt = 0 then
    ("Flash! system problem - could not remove one or more of the files.")
end if
printf(1, "Number of files removed: %d\n", cnt - 1)
```

8.4.2.20.5 See Also:

[remove_directory](#), [delete_file](#)

8.4.2.21 remove_directory

```
include std/filesys.e
public function remove_directory(sequence dir_name, integer force = 0)
```

Remove a directory.

8.4.2.21.1 Parameters:

1. `name` : a sequence, the name of the directory to remove.
2. `force` : an integer, if 1 this will also remove files and sub-directories in the directory. The default is 0, which means that it will only remove the directory if it is already empty.

8.4.2.21.2 Returns:

An **integer**, 0 on failure, 1 on success.

8.4.2.21.3 Example 1:

```
if not remove_directory("the_old_folder") then
    ("Flash! system problem - could not remove the old folder")
end if
```

8.4.2.21.4 See Also:

[create_directory](#), [chdir](#), [clear_directory](#)

8.4.3 File name parsing

8.4.3.1 PATH_DIR

```
include std/filesys.e
public enum PATH_DIR
```

8.4.3.2 PATH_FILENAME

```
include std/filesys.e
public enum PATH_FILENAME
```

8.4.3.3 PATH_BASENAME

```
include std/filesys.e
public enum PATH_BASENAME
```

8.4.3.4 PATH_FILEEXT

```
include std/filesys.e
public enum PATH_FILEEXT
```

8.4.3.5 PATH_DRIVEID

```
include std/filesys.e
public enum PATH_DRIVEID
```

8.4.3.6 pathinfo

```
include std/filesys.e
public function pathinfo(sequence path, integer std_slash = 0)
```

Parse a fully qualified pathname.

8.4.3.6.1 Parameters:

1. `path` : a sequence, the path to parse

8.4.3.6.2 Returns:

A **sequence**, of length 5. Each of these elements is a string:

- The path name
- The full unqualified file name
- the file name, without extension
- the file extension
- the drive id

8.4.3.6.3 Comments:

The host operating system path separator is used in the parsing.

8.4.3.6.4 Example 1:

```
-- WIN32
info = pathinfo("C:\\euphoria\\docs\\readme.txt")
-- info is {"C:\\euphoria\\docs", "readme.txt", "readme", "txt", "C"}
```

8.4.3.6.5 Example 2:

```
-- Unix variants
info = pathinfo("/opt/euphoria/docs/readme.txt")
-- info is {"opt/euphoria/docs", "readme.txt", "readme", "txt", ""}
```

8.4.3.6.6 Example 3:

```
-- no extension
info = pathinfo("/opt/euphoria/docs/readme")
-- info is {"opt/euphoria/docs", "readme", "readme", "", ""}
```

8.4.3.6.7 See Also:

[driveid](#), [dirname](#), [filename](#), [fileext](#), [PATH_BASENAME](#), [PATH_DIR](#), [PATH_DRIVEID](#), [PATH_FILEEXT](#), [PATH_FILENAME](#)

8.4.3.7 dirname

```
include std/filesys.e
public function dirname(sequence path, integer pcd = 0)
```

Return the directory name of a fully qualified filename

8.4.3.7.1 Parameters:

1. `path` : the path from which to extract information
2. `pcd` : If not zero and there is no directory name in `path` then "." is returned. The default (0) will just return any directory name in `path`.

8.4.3.7.2 Returns:

A **sequence**, the full file name part of `path`.

8.4.3.7.3 Comments:

The host operating system path separator is used.

8.4.3.7.4 Example 1:

```
fname = dirname("/opt/euphoria/docs/readme.txt")
-- fname is "/opt/euphoria/docs"
```

8.4.3.7.5 See Also:

[driveid](#), [filename](#), [pathinfo](#)

8.4.3.8 filename

```
include std/filesys.e
public function filename(sequence path)
```

Return the file name portion of a fully qualified filename

8.4.3.8.1 Parameters:

1. `path` : the path from which to extract information

8.4.3.8.2 Returns:

A **sequence**, the file name part of `path`.

8.4.3.8.3 Comments:

The host operating system path separator is used.

8.4.3.8.4 Example 1:

```
fname = filename("/opt/euphoria/docs/readme.txt")
-- fname is "readme.txt"
```

8.4.3.8.5 See Also:

[pathinfo](#), [filebase](#), [fileext](#)

8.4.3.9 filebase

```
include std/filesys.e
public function filebase(sequence path)
```

Return the base filename of path.

8.4.3.9.1 Parameters:

1. `path` : the path from which to extract information

8.4.3.9.2 Returns:

A **sequence**, the base file name part of path.

TODO: Test

8.4.3.9.3 Example 1:

```
base = filebase("/opt/euphoria/readme.txt")
-- base is "readme"
```

8.4.3.9.4 See Also:

[pathinfo](#), [filename](#), [fileext](#)

8.4.3.10 fileext

```
include std/filesys.e
public function fileext(sequence path)
```

Return the file extension of a fully qualified filename

8.4.3.10.1 Parameters:

1. `path` : the path from which to extract information

8.4.3.10.2 Returns:

A **sequence**, the file extension part of `path`.

8.4.3.10.3 Comments:

The host operating system path separator is used.

8.4.3.10.4 Example 1:

```
fname = fileext("/opt/euphoria/docs/readme.txt")
-- fname is "txt"
```

8.4.3.10.5 See Also:

[pathinfo](#), [filename](#), [filebase](#)

8.4.3.11 driveid

```
include std/filesys.e
public function driveid(sequence path)
```

Return the drive letter of the path on *WIN32* platforms.

8.4.3.11.1 Parameters:

1. `path` : the path from which to extract information

8.4.3.11.2 Returns:

A **sequence**, the file extension part of `path`.

TODO: Test

8.4.3.11.3 Example:

```
letter = driveid("C:\\EUPHORIA\\Readme.txt")
-- letter is "C"
```

8.4.3.11.4 See Also:

[pathinfo](#), [dirname](#), [filename](#)

8.4.3.12 defaulttext

```
include std/filesys.e
public function defaulttext(sequence path, sequence defext)
```

Returns the supplied filepath with the supplied extension, if the filepath does not have an extension already.

8.4.3.12.1 Parameters:

1. `path` : the path to check for an extension.
2. `defext` : the extension to add if `path` does not have one.

8.4.3.12.2 Returns:

A **sequence**, the path with an extension.

8.4.3.12.3 Example:

```
-- ensure that the supplied path has an extension, but if it doesn't use "tmp".
theFile = defaulttext(UserFileName, "tmp")
```

8.4.3.12.4 See Also:

[pathinfo](#)

8.4.3.13 absolute_path

```
include std/filesys.e
public function absolute_path(sequence filename)
```

Determine if the supplied string is an absolute path or a relative path.

8.4.3.13.1 Parameters:

1. `filename` : a sequence, the name of the file path

8.4.3.13.2 Returns:

An **integer**, 0 if filename is a relative path or 1 otherwise.

8.4.3.13.3 Comment:

A *relative* path is one which is relative to the current directory and an *absolute* path is one that doesn't need to know the current directory to find the file.

8.4.3.13.4 Example 1:

```
? absolute_path("") -- returns 0
? absolute_path("/usr/bin/abc") -- returns 1
? absolute_path("\\temp\\somefile.doc") -- returns 1
? absolute_path("../abc") -- returns 0
? absolute_path("local/abc.txt") -- returns 0
? absolute_path("abc.txt") -- returns 0
? absolute_path("c:..\\abc") -- returns 0
-- The next two examples return 0 on Unix platforms and 1 on Microsoft platforms
? absolute_path("c:\\windows\\system32\\abc")
? absolute_path("c:/windows/system32/abc")
```

8.4.3.14 canonical_path

```
include std/filesys.e
public function canonical_path(sequence path_in, integer directory_given = 0)
```

Returns the full path and file name of the supplied file name.

8.4.3.14.1 Parameters:

1. `path_in`: A sequence. This is the file name whose full path you want.
2. `directory_given`: An integer. This is zero if `path_in` is to be interpreted as a file specification otherwise it is assumed to be a directory specification. The default is zero.

8.4.3.14.2 Returns:

A **sequence**, the full path and file name.

8.4.3.14.3 Comment:

- In non-Unix systems, the result is always in lowercase.
- The supplied file/directory does not have to actually exist.
- Does not (yet) handle UNC paths or unix links.

8.4.3.14.4 Example 1:

```
-- Assuming the current directory is "/usr/foo/bar"
res = canonical_path("../abc.def")
-- res is now "/usr/foo/abc.def"
```

8.4.4 File Types

8.4.4.1 FILETYPE_UNDEFINED

```
include std/filesys.e
public enum FILETYPE_UNDEFINED
```

8.4.4.2 FILETYPE_NOT_FOUND

```
include std/filesys.e
public enum FILETYPE_NOT_FOUND
```

8.4.4.3 FILETYPE_FILE

```
include std/filesys.e
public enum FILETYPE_FILE
```

8.4.4.4 FILETYPE_DIRECTORY

```
include std/filesys.e
public enum FILETYPE_DIRECTORY
```

8.4.4.5 file_type

```
include std/filesys.e
public function file_type(sequence filename)
```

Get the type of a file.

8.4.4.5.1 Parameters:

1. `filename` : the name of the file to query. It must not have wildcards.

8.4.4.5.2 Returns:

An **integer**,

- -1 if file could be multiply defined
- 0 if filename does not exist
- 1 if filename is a file
- 2 if filename is a directory

8.4.4.5.3 See Also:

[dir](#), [FILETYPE_DIRECTORY](#), [FILETYPE_FILE](#), [FILETYPE_NOT_FOUND](#), [FILETYPE_UNDEFINED](#)

8.4.5 File Handling

8.4.5.1 SECTORS_PER_CLUSTER

```
include std/filesys.e
public enum SECTORS_PER_CLUSTER
```

8.4.5.2 BYTES_PER_SECTOR

```
include std/filesys.e
public enum BYTES_PER_SECTOR
```

8.4.5.3 NUMBER_OF_FREE_CLUSTERS

```
include std/filesys.e
public enum NUMBER_OF_FREE_CLUSTERS
```

8.4.5.4 TOTAL_NUMBER_OF_CLUSTERS

```
include std/filesys.e
public enum TOTAL_NUMBER_OF_CLUSTERS
```

8.4.5.5 TOTAL_BYTES

```
include std/filesys.e
public enum TOTAL_BYTES
```

8.4.5.6 FREE_BYTES

```
include std/filesys.e
public enum FREE_BYTES
```

8.4.5.7 USED_BYTES

```
include std/filesys.e
public enum USED_BYTES
```

8.4.5.8 COUNT_DIRS

```
include std/filesys.e
public enum COUNT_DIRS
```

8.4.5.9 COUNT_FILES

```
include std/filesys.e
public enum COUNT_FILES
```

8.4.5.10 COUNT_SIZE

```
include std/filesys.e
public enum COUNT_SIZE
```

8.4.5.11 COUNT_TYPES

```
include std/filesys.e
public enum COUNT_TYPES
```

8.4.5.12 EXT_NAME

```
include std/filesys.e
public enum EXT_NAME
```

8.4.5.13 EXT_COUNT

```
include std/filesys.e
public enum EXT_COUNT
```

8.4.5.14 EXT_SIZE

```
include std/filesys.e
public enum EXT_SIZE
```

8.4.5.15 file_exists

```
include std/filesys.e
public function file_exists(object name)
```

Check to see if a file exists

8.4.5.15.1 Parameters:

1. name : filename to check existence of

8.4.5.15.2 Returns:

An **integer**, 1 on yes, 0 on no

8.4.5.15.3 Example 1:

```
if file_exists("abc.e") then
    puts(1, "abc.e exists already\n")
end if
```

8.4.5.16 file_timestamp

```
include std/filesys.e
public function file_timestamp(sequence fname)
```

Get the timestamp of the file

8.4.5.16.1 Parameters:

1. `name` : the filename to get the date of

8.4.5.16.2 Returns:

A valid **datetime type**, representing the files date and time or -1 if the file's date and time could not be read.

8.4.5.17 copy_file

```
include std/filesys.e
public function copy_file(sequence src, sequence dest, integer overwrite = 0)
```

Copy a file.

8.4.5.17.1 Parameters:

1. `src` : a sequence, the name of the file or directory to copy
2. `dest` : a sequence, the new name or location of the file
3. `overwrite` : an integer; 0 (the default) will prevent an existing destination file from being overwritten. Non-zero will overwrite the destination file.

8.4.5.17.2 Returns:

An **integer**, 0 on failure, 1 on success.

8.4.5.17.3 Comments:

If `overwrite` is true, and if `dest` file already exists, the function overwrites the existing file and succeeds.

8.4.5.17.4 See Also:

[move_file](#), [rename_file](#)

8.4.5.18 rename_file

```
include std/filesys.e
public function rename_file(sequence src, sequence dest, integer overwrite = 0)
```


Rename a file.

8.4.5.18.1 Parameters:

1. `src` : a sequence, the name of the file or directory to rename.
2. `dest` : a sequence, the new name for the renamed file
3. `overwrite` : an integer, 0 (the default) to prevent renaming if destination file exists, 1 to delete existing destination file first

8.4.5.18.2 Returns:

An **integer**, 0 on failure, 1 on success.

8.4.5.18.3 Comments:

- If `dest` contains a path specification, this is equivalent to moving the file, as well as possibly changing its name. However, the path must be on the same drive for this to work.
- If `overwrite` was requested but the rename fails, any existing destination file is preserved.

8.4.5.18.4 See Also:

[move_file](#), [copy_file](#)

8.4.5.19 move_file

```
include std/filesys.e
public function move_file(sequence src, sequence dest, integer overwrite = 0)
```

Move a file to another location.

8.4.5.19.1 Parameters:

1. `src` : a sequence, the name of the file or directory to move
2. `dest` : a sequence, the new location for the file
3. `overwrite` : an integer, 0 (the default) to prevent overwriting an existing destination file, 1 to overwrite existing destination file

8.4.5.19.2 Returns:

An **integer**, 0 on failure, 1 on success.

8.4.5.19.3 Comments:

- If `overwrite` was requested but the move fails, any existing destination file is preserved.

8.4.5.19.4 See Also:

[rename_file](#), [copy_file](#)

8.4.5.20 file_length

```
include std/filesys.e
public function file_length(sequence filename)
```

Return the size of a file.

8.4.5.20.1 Parameters:

1. `filename` : the name of the queried file

8.4.5.20.2 Returns:

An **atom**, the file size, or -1 if file is not found.

8.4.5.20.3 Comments:

This function does not compute the total size for a directory, and returns 0 instead.

8.4.5.20.4 See Also:

[dir](#)

8.4.5.21 locate_file

```
include std/filesys.e
public function locate_file(sequence filename, sequence search_list = {}, sequence subdir = {})
```

Locates a file by looking in a set of directories for it.

8.4.5.21.1 Parameters:

1. `filename` : a sequence, the name of the file to search for.
2. `search_list` : a sequence, the list of directories to look in. By default this is "", meaning that a predefined set of directories is scanned. See comments below.
3. `subdir` : a sequence, the sub directory within the search directories to check. This is optional.

8.4.5.21.2 Returns:

A **sequence**, the located file path if found, else the original file name.

8.4.5.21.3 Comments:

If `filename` is an absolute path, it is just returned and no searching takes place.

If `filename` is located, the full path of the file is returned.

If `search_list` is supplied, it can be either a sequence of directory names, of a string of directory names delimited by ':' in UNIX and ';' in Windows.

If the `search_list` is omitted or "", this will look in the following places...

- The current directory
- The directory that the program is run from.
- The directory in \$HOME (\$HOMEDRIVE & \$HOMEPATH in Windows)
- The parent directory of the current directory
- The directories returned by `include_paths()`
- \$EUDIR/bin
- \$EUDIR/docs
- \$EUDIST/
- \$EUDIST/etc
- \$EUDIST/data
- The directories listed in \$USERPATH
- The directories listed in \$PATH

If the `subdir` is supplied, the function looks in this sub directory for each of the directories in the search list.

8.4.5.21.4 Example 1:

```
res = locate_file("abc.def", {"/usr/bin", "/u2/someapp", "/etc"})
res = locate_file("abc.def", "/usr/bin:/u2/someapp:/etc")
res = locate_file("abc.def") -- Scan default locations.
res = locate_file("abc.def", , "app") -- Scan the 'app' sub directory in the default locations
```

8.4.5.22 disk_metrics

```
include std/filesys.e
public function disk_metrics(object disk_path)
```

Returns some information about a disk drive.

8.4.5.22.1 Parameters:

1. `disk_path` : A sequence. This is the path that identifies the disk to inquire upon.

8.4.5.22.2 Returns:

A **sequence**, containing SECTORS_PER_CLUSTER, BYTES_PER_SECTOR, NUMBER_OF_FREE_CLUSTERS, and TOTAL_NUMBER_OF_CLUSTERS

8.4.5.22.3 Example 1:

```
res = disk_metrics("C:\\")
min_file_size = res[SECTORS_PER_CLUSTER] * res[BYTES_PER_SECTOR]
```

8.4.5.23 disk_size

```
include std/filesys.e
public function disk_size(object disk_path)
```

Returns the amount of space for a disk drive.

8.4.5.23.1 Parameters:

1. `disk_path` : A sequence. This is the path that identifies the disk to inquire upon.

8.4.5.23.2 Returns:

A **sequence**, containing TOTAL_BYTES, USED_BYTES, FREE_BYTES, and a string which represents the filesystem name

8.4.5.23.3 Example 1:

```
res = disk_size("C:\\")
printf(1, "Drive %s has %3.2f%% free space\n", {"C:"}, res[FREE_BYTES] / res[TOTAL_BYTES])
```

8.4.5.24 dir_size

```
include std/filesys.e
public function dir_size(sequence dir_path, integer count_all = 0)
```

Returns the amount of space used by a directory.

8.4.5.24.1 Parameters:

1. `dir_path` : A sequence. This is the path that identifies the directory to inquire upon.
2. `count_all` : An integer. Used by Windows systems. If zero (the default) it will not include *system* or *hidden* files in the count, otherwise they are included.

8.4.5.24.2 Returns:

A **sequence**, containing four elements; the number of sub-directories [COUNT_DIRS], the number of files [COUNT_FILES], the total space used by the directory [COUNT_SIZE], and breakdown of the file contents by file extension [COUNT_TYPES].

8.4.5.24.3 Comments:

- The total space used by the directory does not include space used by any sub-directories.
- The file breakdown is a sequence of three-element sub-sequences. Each sub-sequence contains the extension [EXT_NAME], the number of files of this extension [EXT_COUNT], and the space used by these files [EXT_SIZE]. The sub-sequences are presented in extension name order. On Windows the extensions are all in lowercase.

8.4.5.24.4 Example 1:

```
res = dir_size("/usr/localbin")
printf(1, "Directory %s contains %d files\n", {"usr/localbin", res[COUNT_FILES]})
for i = 1 to length(res[COUNT_TYPES]) do
    printf(1, "  Type: %s (%d files %d bytes)\n", {res[COUNT_TYPES][i][EXT_NAME],
                                                    res[COUNT_TYPES][i][EXT_COUNT],
                                                    res[COUNT_TYPES][i][EXT_SIZE]})
end for
```

8.4.5.25 temp_file

```
include std/filesys.e
public function temp_file(sequence temp_location = "", sequence temp_prefix = "", sequence temp
```

Returns a file name that can be used as a temporary file.

8.4.5.25.1 Parameters:

1. `temp_location`: A sequence. A directory where the temporary file is expected to be created.
 - ◆ If omitted (the default) the 'temporary' directory will be used. The temporary directory is defined in the "TEMP" environment symbol, or failing that the "TMP" symbol and failing that "C:\TEMP\" is used in non-Unix systems and "/tmp/" is used in Unix systems.
 - ◆ If `temp_location` was supplied,
 - ◇ If it is an existing file, that file's directory is used.
 - ◇ If it is an existing directory, it is used.
 - ◇ If it doesn't exist, the directory name portion is used.
2. `temp_prefix`: A sequence: The is prepended to the start of the generated file name. The default is "".
3. `temp_extn`: A sequence: The is a file extension used in the generated file. The default is "_T_".
4. `reserve_temp`: An integer: If not zero an empty file is created using the generated name. The default is not to reserve (create) the file.

8.4.5.25.2 Returns:

A **sequence**, A generated file name.

8.4.5.25.3 Comments:

8.4.5.25.4 Example 1:

```
? temp_file("/usr/space", "myapp", "tmp") --> /usr/space/myapp736321.tmp
? temp_file() --> /tmp/277382._T_
? temp_file("/users/me/abc.exw") --> /users/me/992831._T_
```

8.5 I/O

Constants

STDIN
STDOUT
STDERR
SCREEN
EOF

Read/Write Routines

?
print
printf
puts
getc
gets
get_key
get_bytes



- get_integer32
- get_integer16
- put_integer32
- put_integer16
- get_dstring
- Low Level File/Device Handling
 - LOCK_SHARED
 - LOCK_EXCLUSIVE
 - file_number
 - file_position
 - lock_type
 - byte_range
 - open
 - close
 - seek
 - where
 - flush
 - lock_file
 - unlock_file
- File Reading/Writing
 - read_lines
 - process_lines
 - write_lines
 - append_lines
 - BINARY_MODE
 - TEXT_MODE
 - UNIX_TEXT
 - DOS_TEXT
 - ANSI
 - UTF
 - UTF_8
 - UTF_16
 - UTF_16BE
 - UTF_16LE
 - UTF_32
 - UTF_32BE
 - UTF_32LE
 - read_file
 - write_file
 - writef
 - writeln

8.5.1 Constants

8.5.1.1 STDIN

```
include std/io.e
public constant STDIN
```

Standard Input

8.5.1.2 STDOUT

```
include std/io.e
public constant STDOUT
```

Standard Output

8.5.1.3 STDERR

```
include std/io.e
public constant STDERR
```

Standard Error

8.5.1.4 SCREEN

```
include std/io.e
public constant SCREEN
```

Screen (Standard Out)

8.5.1.5 EOF

```
include std/io.e
public constant EOF
```

End of file

8.5.2 Read/Write Routines

q_print

8.5.2.1 ?

<built-in> `procedure ?` (no parentheses around the unique parameter)

Shorthand way of saying: `pretty_print(STDOUT, x, {})` -- i.e. printing the value of an expression to the standard output, with braces and indentation to show the structure.

8.5.2.1.1 Example 1:

```
? {1, 2} + {3, 4}  -- will display {4, 6}
```

8.5.2.1.2 See Also:

[print](#)

8.5.2.2 print

<built-in> `procedure print(integer fn, object x)`

Writes out a **text** representation of an object to a file or device. If the object `x` is a sequence, it uses braces `{ , , , }` to show the structure.

8.5.2.2.1 Parameters:

1. `fn` : an integer, the handle to a file or device to output to
2. `x` : the object to print

8.5.2.2.2 Errors:

The target file or device must be open.

8.5.2.2.3 Comments:

This is not used to write to "binary" files as it only outputs text.

8.5.2.2.4 Example 1:

```
print(STDOUT, "ABC")  -- output is: "{65,66,67}"
puts(STDOUT, "ABC")   -- output is: "ABC"
print(STDOUT, 65)     -- output is: "65"
puts(STDOUT, 65)      -- output is: "A"   (ASCII-65 ==> 'A')
puts(STDOUT, 65.1234) -- output is: "65.1234"
puts(STDOUT, 65.1234) -- output is: "A"   (Converts to integer first)
```

8.5.2.2.5 Example 2:

```
print(STDOUT, repeat({10,20}, 3)) -- output is: {{10,20},{10,20},{10,20}}
```

8.5.2.2.6 See Also:

?, puts

8.5.2.3 printf

```
<built-in> procedure printf(integer fn, sequence format, object values)
```

Print one or more values to a file or device, using a format string to embed them in and define how they should be represented.

8.5.2.3.1 Parameters:

1. `fn` : an integer, the handle to a file or device to output to
2. `format` : a sequence, the text to print. This text may contain format specifiers.
3. `values` : usually, a sequence of values. It should have as many elements as format specifiers in `format`, as these values will be substituted to the specifiers.

8.5.2.3.2 Errors:

If there are less values to show than format specifiers, a run time error will occur.

The target file or device must be open.

8.5.2.3.3 Comments:

A format specifier is a string of characters starting with a percent sign (%) and ending in a letter. Some extra information may come in the middle.

`format` will be scanned for format specifiers. Whenever one is found, the current value in `values` will be turned into a string according to the format specifier. The resulting string will be plugged in the result, as if replacing the modifier with the printed value. Then moving on to next value and carrying the process on.

This way, `printf()` always takes exactly 3 arguments, no matter how many values are to be printed. Only the length of the last argument, containing the values to be printed, will vary.

The basic format specifiers are...

- `%d` -- print an atom as a decimal integer

- `%x` -- print an atom as a hexadecimal integer. Negative numbers are printed in two's complement, so `-1` will print as `FFFFFFFF`
- `%o` -- print an atom as an octal integer
- `%s` -- print a sequence as a string of characters, or print an atom as a single character
- `%e` -- print an atom as a floating-point number with exponential notation
- `%f` -- print an atom as a floating-point number with a decimal point but no exponent
- `%g` -- print an atom as a floating-point number using whichever format seems appropriate, given the magnitude of the number
- `%%` -- print the `'%'` character itself. This is not an actual format specifier.

Field widths can be added to the basic formats, e.g. `%5d`, `%8.2f`, `%10.4s`. The number before the decimal point is the minimum field width to be used. The number after the decimal point is the precision to be used.

If the field width is negative, e.g. `%-5d` then the value will be left-justified within the field. Normally it will be right-justified. If the field width starts with a leading 0, e.g. `%08d` then leading zeros will be supplied to fill up the field. If the field width starts with a `'+'` e.g. `%+7d` then a plus sign will be printed for positive values.

8.5.2.3.4 Comments:

8.5.2.3.5 Watch out for the following common mistake:

```
name="John Smith"
printf(STDOUT, "%s", name)      -- error!
```

This will print only the first character, J, of name, as each element of name is taken to be a separate value to be formatted. You must say this instead:

```
name="John Smith"
printf(STDOUT, "%s", {name})    -- correct
```

Now, the third argument of `printf()` is a one-element sequence containing the item to be formatted.

If there is only one `%` format specifier, and if the value it stands for is an atom, then values may be simply that atom.

8.5.2.3.6 Example 1:

```
rate = 7.875
printf(my_file, "The interest rate is: %8.2f\n", rate)

--      The interest rate is:      7.88
```

8.5.2.3.7 Example 2:

```
name="John Smith"
score=97
printf(STDOUT, "%15s, %5d\n", {name, score})
```

```
--      John Smith,      97
```

8.5.2.3.8 Example 3:

```
printf(STDOUT, "%-10.4s $ %s", {"ABCDEFGHJKLMNOP", "XXX"})
--      ABCD      $ XXX
```

8.5.2.3.9 Example 4:

```
printf(STDOUT, "%d %e %f %g", 7.75) -- same value in different formats
--      7  7.750000e+000  7.750000  7.75
```

8.5.2.3.10 See Also:

[sprintf](#), [sprint](#), [print](#)

8.5.2.4 puts

<built-in> [procedure puts](#)(integer fn, object text)

Output, to a file or device, a single byte (atom) or sequence of bytes. The low order 8-bits of each value is actually sent out. If outputting to the screen you will see text characters displayed.

8.5.2.4.1 Parameters:

1. `fn` : an integer, the handle to an opened file or device
2. `text` : an object, either a single character or a sequence of characters.

8.5.2.4.2 Errors:

The target file or device must be open.

8.5.2.4.3 Comments:

When you output a sequence of bytes it must not have any (sub)sequences within it. It must be a sequence of atoms only. (Typically a string of ASCII codes).

Avoid outputting 0's to the screen or to standard output. Your output might get truncated.

Remember that if the output file was opened in text mode, *Windows* will change `\n` (10) to `\r\n` (13 10). Open the file in binary mode if this is not what you want.

8.5.2.4.4 Example 1:

```
puts(SCREEN, "Enter your first name: ")
```

8.5.2.4.5 Example 2:

```
puts(output, 'A') -- the single byte 65 will be sent to output
```

8.5.2.4.6 See Also:

[print](#)

8.5.2.5 getc

```
<built-in> function getc(integer fn)
```

Get the next character (byte) from a file or device fn.

8.5.2.5.1 Parameters:

1. `fn` : an integer, the handle of the file or device to read from.

8.5.2.5.2 Returns:

An **integer**, the character read from the file, in the 0..255 range. If no character is left to read, **EOF** is returned instead.

8.5.2.5.3 Errors:

The target file or device must be open.

8.5.2.5.4 Comments:

File input using `getc()` is buffered, i.e. `getc()` does not actually go out to the disk for each character. Instead, a large block of characters will be read in at one time and returned to you one by one from a memory buffer.

When `getc()` reads from the keyboard, it will not see any characters until the user presses Enter. Note that the user can type CTRL+Z, which the operating system treats as "end of file". **EOF** will be returned.

8.5.2.5.5 See Also:

[gets](#), [get_key](#)

8.5.2.6 gets

```
<built-in> function gets(integer fn)
```

Get the next sequence (one line, including '\n') of characters from a file or device.

8.5.2.6.1 Parameters:

1. `fn` : an integer, the handle of the file or device to read from.

8.5.2.6.2 Returns:

An **object**, either [EOF](#) on end of file, or the next line of text from the file.

8.5.2.6.3 Errors:

The file or device must be open.

8.5.2.6.4 Comments:

The characters will have values from 0 to 255.

If the line had an end of line marker, a '~\n' terminates the line. The last line of a file needs not have an end of line marker.

After reading a line of text from the keyboard, you should normally output a '\n' character, e.g. `puts(1, '\n')`, before printing something. Only on the last line of the screen does the operating system automatically scroll the screen and advance to the next line.

When your program reads from the keyboard, the user can type control-Z, which the operating system treats as "end of file". [EOF](#) will be returned.

8.5.2.6.5 Example 1:

```
sequence buffer
object line
integer fn

-- read a text file into a sequence
fn = open("my_file.txt", "r")
```

```
if fn = -1 then
    puts(1, "Couldn't open my_file.txt\n")
    abort(1)
end if

buffer = {}
while 1 do
    line = gets(fn)
    if atom(line) then
        exit -- EOF is returned at end of file
    end if
    buffer = append(buffer, line)
end while
```

8.5.2.6.6 Example 2:

```
object line

puts(1, "What is your name?\n")
line = gets(0) -- read standard input (keyboard)
line = line[1..$-1] -- get rid of \n character at end
puts(1, '\n') -- necessary
puts(1, line & " is a nice name.\n")
```

8.5.2.6.7 See Also:

[getc](#), [read_lines](#)

8.5.2.7 get_key

<built-in> `function get_key()`

Get the next keystroke without waiting for it or echoing it on the console.

8.5.2.7.1 Parameters:

1. None.

8.5.2.7.2 Returns:

An **integer**, the code number for the key pressed. If there is no key press waiting, then this returns -1.

8.5.2.7.3 See Also:

[gets](#), [getc](#)

8.5.2.8 get_bytes

```
include std/io.e
public function get_bytes(integer fn, integer n)
```

Read the next bytes from a file.

8.5.2.8.1 Parameters:

1. `fn` : an integer, the handle to an open file to read from.
2. `n` : a positive integer, the number of bytes to read.

8.5.2.8.2 Returns:

A **sequence**, of length at most `n`, made of the bytes that could be read from the file.

8.5.2.8.3 Comments:

When `n > 0` and the function returns a sequence of length less than `n` you know you've reached the end of file. Eventually, an empty sequence will be returned.

This function is normally used with files opened in binary mode, "rb". This avoids the confusing situation in text mode where *Windows* will convert CR LF pairs to LF.

8.5.2.8.4 Example 1:

```
integer fn
fn = open("temp", "rb")  -- an existing file

sequence whole_file
whole_file = {}

sequence chunk

while 1 do
    chunk = get_bytes(fn, 100) -- read 100 bytes at a time
    whole_file &= chunk        -- chunk might be empty, that's ok
    if length(chunk) < 100 then
        exit
    end if
end while

close(fn)
? length(whole_file)  -- should match DIR size of "temp"
```


8.5.2.8.5 See Also:

[getc](#), [gets](#), [get_integer32](#), [get_dstring](#)

8.5.2.9 get_integer32

```
include std/io.e
public function get_integer32(integer fh)
```

Read the next four bytes from a file and returns them as a single integer.

8.5.2.9.1 Parameters:

1. `fh` : an integer, the handle to an open file to read from.

8.5.2.9.2 Returns:

An **atom**, made of the bytes that could be read from the file.

8.5.2.9.3 Comments:

- This function is normally used with files opened in binary mode, "rb".
- Assumes that there at least four bytes available to be read.

8.5.2.9.4 Example 1:

```
integer fn
fn = open("temp", "rb")  -- an existing file

atom file_type_code
file_type_code = get_integer32(fn)
```

8.5.2.9.5 See Also:

[getc](#), [gets](#), [get_bytes](#), [get_dstring](#)

8.5.2.10 get_integer16

```
include std/io.e
public function get_integer16(integer fh)
```

Read the next two bytes from a file and returns them as a single integer.

8.5.2.10.1 Parameters:

1. `fh` : an integer, the handle to an open file to read from.

8.5.2.10.2 Returns:

An **atom**, made of the bytes that could be read from the file.

8.5.2.10.3 Comments:

- This function is normally used with files opened in binary mode, "rb".
- Assumes that there at least two bytes available to be read.

8.5.2.10.4 Example 1:

```
integer fn
fn = open("temp", "rb")  -- an existing file

atom file_type_code
file_type_code = get_integer16(fn)
```

8.5.2.10.5 See Also:

[getc](#), [gets](#), [get_bytes](#), [get_dstring](#)

8.5.2.11 put_integer32

```
include std/io.e
public procedure put_integer32(integer fh, integer val)
```

Write the supplied integer as four bytes to a file.

8.5.2.11.1 Parameters:

1. `fh` : an integer, the handle to an open file to write to.
2. `val` : an integer

8.5.2.11.2 Comments:

- This function is normally used with files opened in binary mode, "wb".

8.5.2.11.3 Example 1:

```
integer fn
fn = open("temp", "wb")

put_integer32(fn, 1234)
```

8.5.2.11.4 See Also:

[getc](#), [gets](#), [get_bytes](#), [get_dstring](#)

8.5.2.12 put_integer16

```
include std/io.e
public procedure put_integer16(integer fh, integer val)
```

Write the supplied integer as two bytes to a file.

8.5.2.12.1 Parameters:

1. `fh` : an integer, the handle to an open file to write to.
2. `val` : an integer

8.5.2.12.2 Comments:

- This function is normally used with files opened in binary mode, "wb".

8.5.2.12.3 Example 1:

```
integer fn
fn = open("temp", "wb")

put_integer16(fn, 1234)
```

8.5.2.12.4 See Also:

[getc](#), [gets](#), [get_bytes](#), [get_dstring](#)

8.5.2.13 get_dstring

```
include std/io.e
public function get_dstring(integer fh, integer delim = 0)
```

Read a delimited byte string from an opened file .

8.5.2.13.1 Parameters:

1. `fh` : an integer, the handle to an open file to read from.
2. `delim` : an integer, the delimiter that marks the end of a byte string. If omitted, a zero is assumed.

8.5.2.13.2 Returns:

An **sequence**, made of the bytes that could be read from the file.

8.5.2.13.3 Comments:

- If the end-of-file is found before the delimiter, the delimiter is appended to the returned string.

8.5.2.13.4 Example 1:

```
integer fn
fn = open("temp", "rb")  -- an existing file

sequence text
text = get_dstring(fn)- Get a zero-delimited string
text = get_dstring(fn, '$')- Get a '$'-delimited string
```

8.5.2.13.5 See Also:

[getc](#), [gets](#), [get_bytes](#), [get_integer32](#)

8.5.3 Low Level File/Device Handling

8.5.3.1 LOCK_SHARED

```
include std/io.e
public enum LOCK_SHARED
```

8.5.3.2 LOCK_EXCLUSIVE

```
include std/io.e
public enum LOCK_EXCLUSIVE
```

8.5.3.3 file_number

```
include std/io.e
public type file_number(integer f)
```

File number type

8.5.3.4 file_position

```
include std/io.e
public type file_position(atom p)
```

File position type

8.5.3.5 lock_type

```
include std/io.e
public type lock_type(integer t)
```

Lock Type

8.5.3.6 byte_range

```
include std/io.e
public type byte_range(sequence r)
```

Byte Range Type

8.5.3.7 open

```
<built-in> function open(sequence path, sequence mode, integer cleanup = 0)
```

Open a file or device, to get the file number.

8.5.3.7.1 Parameters:

1. `path` : a string, the path to the file or device to open.
2. `mode` : a string, the mode being used to open the file.
3. `cleanup` : an integer, if 0, then the file must be manually closed by the coder. If 1, then the file will be closed when either the file handle's references goes to 0, or if called as a parameter to `delete()`.

8.5.3.7.2 Returns:

A small **integer**, -1 on failure, else 0 or more.

8.5.3.7.3 Errors:

There is a limit on the number of files that can be simultaneously opened, currently 40. If this limit is reached, the next attempt to `open()` a file will error out.

The length of `path` should not exceed 1,024 characters.

8.5.3.7.4 Comments:

8.5.3.7.5 Possible modes are:

- "r" -- open text file for reading
- "rb" -- open binary file for reading
- "w" -- create text file for writing
- "wb" -- create binary file for writing
- "u" -- open text file for update (reading and writing)
- "ub" -- open binary file for update
- "a" -- open text file for appending
- "ab" -- open binary file for appending

Files opened for read or update must already exist. Files opened for write or append will be created if necessary. A file opened for write will be set to 0 bytes. Output to a file opened for append will start at the end of file.

On *Windows*, output to text files will have carriage-return characters automatically added before linefeed characters. On input, these carriage-return characters are removed. A control-Z character (ASCII 26) will signal an immediate end of file.

I/O to binary files is not modified in any way. Any byte values from 0 to 255 can be read or written. On *Unix*, all files are binary files, so "r" mode and "rb" mode are equivalent, as are "w" and "wb", "u" and "ub", and "a" and "ab".

8.5.3.7.6 Some typical devices that you can open on Windows are:

- "CON" -- the console (screen)
- "AUX" -- the serial auxiliary port
- "COM1" -- serial port 1
- "COM2" -- serial port 2
- "PRN" -- the printer on the parallel port
- "NUL" -- a non-existent device that accepts and discards output

Close a file or device when done with it, flushing out any still-buffered characters prior.

WIN32 and *Unix*: Long filenames are fully supported for reading and writing and creating.

WIN32: Be careful not to use the special device names in a file name, even if you add an extension. e.g. CON.TXT, CON.DAT, CON.JPG etc. all refer to the CON device, **not a file**.

8.5.3.7.7 Example 1:

```
integer file_num, file_num95
sequence first_line
constant ERROR = 2

file_num = open("my_file", "r")
if file_num = -1 then
    puts(ERROR, "couldn't open my_file\n")
else
    first_line = gets(file_num)
end if

file_num = open("PRN", "w") -- open printer for output

-- on Windows 95:
file_num95 = open("big_directory_name\\very_long_file_name.abcdefg",
                  "r")
if file_num95 != -1 then
    puts(STDOUT, "it worked!\n")
end if
```

8.5.3.8 close

```
<built-in> procedure close(atom fn)
```

Close a file or device and flush out any still-buffered characters.

8.5.3.8.1 Parameters:

1. *fn* : an integer, the handle to the file or device to query.

8.5.3.8.2 Errors:

The target file or device must be open.

8.5.3.8.3 Comments:

Any still-open files will be closed automatically when your program terminates.

8.5.3.9 seek

```
include std/io.e
public function seek(file_number fn, file_position pos)
```

Seek (move) to any byte position in a file.

8.5.3.9.1 Parameters:

1. `fn` : an integer, the handle to the file or device to seek()
2. `pos` : an atom, either an absolute 0-based position or -1 to seek to end of file.

8.5.3.9.2 Returns:

An **integer**, 0 on success, 1 on failure.

8.5.3.9.3 Errors:

The target file or device must be open.

8.5.3.9.4 Comments:

For each open file, there is a current byte position that is updated as a result of I/O operations on the file. The initial file position is 0 for files opened for read, write or update. The initial position is the end of file for files opened for append. It is possible to seek past the end of a file. If you seek past the end of the file, and write some data, undefined bytes will be inserted into the gap between the original end of file and your new data.

After seeking and reading (writing) a series of bytes, you may need to call seek() explicitly before you switch to writing (reading) bytes, even though the file position should already be what you want.

This function is normally used with files opened in binary mode. In text mode, Windows converts CR LF to LF on input, and LF to CR LF on output, which can cause great confusion when you are trying to count bytes because seek() counts the Windows end of line sequences as two bytes, even if the file has been opened in text mode.

8.5.3.9.5 Example 1:

```
include std/io.e

integer fn
fn = open("my.data", "rb")
-- read and display first line of file 3 times:
for i = 1 to 3 do
    puts(STDOUT, gets(fn))
    if seek(fn, 0) then
        puts(STDOUT, "rewind failed!\n")
    end if
end for
```



```
end for
```

8.5.3.9.6 See Also:

[get_bytes](#), [puts](#), [where](#)

8.5.3.10 where

```
include std/io.e
public function where(file_number fn)
```

Retrieves the current file position for an opened file or device.

8.5.3.10.1 Parameters:

1. `fn` : an integer, the handle to the file or device to query.

8.5.3.10.2 Returns:

An **atom**, the current byte position in the file.

8.5.3.10.3 Errors:

The target file or device must be open.

8.5.3.10.4 Comments:

The file position is the place in the file where the next byte will be read from, or written to. It is updated by reads, writes and seeks on the file. This procedure always counts Windows end of line sequences (CR LF) as two bytes even when the file number has been opened in text mode.

8.5.3.11 flush

```
include std/io.e
public procedure flush(file_number fn)
```

Force writing any buffered data to an open file or device.

8.5.3.11.1 Parameters:

1. `fn` : an integer, the handle to the file or device to close.

8.5.3.11.2 Errors:

The target file or device must be open.

8.5.3.11.3 Comments:

When you write data to a file, EUPHORIA normally stores the data in a memory buffer until a large enough chunk of data has accumulated. This large chunk can then be written to disk very efficiently. Sometimes you may want to force, or flush, all data out immediately, even if the memory buffer is not full. To do this you must call `flush(fn)`, where `fn` is the file number of a file open for writing or appending.

When a file is closed, (see `close()`), all buffered data is flushed out. When a program terminates, all open files are flushed and closed automatically. Use `flush()` when another process may need to see all of the data written so far, but you are not ready to close the file yet. `flush()` is also used in crash routines, where files may not be closed in the cleanest possible way.

8.5.3.11.4 Example 1:

```
f = open("file.log", "w")
puts(f, "Record#1\n")
puts(STDOUT, "Press Enter when ready\n")

flush(f)  -- This forces "Record #1" into "file.log" on disk.
          -- Without this, "file.log" will appear to have
          -- 0 characters when we stop for keyboard input.

s = gets(0) -- wait for keyboard input
```

8.5.3.11.5 See Also:

[close](#), [crash_routine](#)

8.5.3.12 lock_file

```
include std/io.e
public function lock_file(file_number fn, lock_type t, byte_range r = {})
```

When multiple processes can simultaneously access a file, some kind of locking mechanism may be needed to avoid mangling the contents of the file, or causing erroneous data to be read from the file.

8.5.3.12.1 Parameters:

1. `fn` : an integer, the handle to the file or device to (partially) lock.
2. `t` : an integer which defines the kind of lock to apply.
3. `r` : a sequence, defining a section of the file to be locked, or `{ }` for the whole file (the default).

8.5.3.12.2 Returns:

An **integer**, 0 on failure, 1 on success.

8.5.3.12.3 Errors:

The target file or device must be open.

8.5.3.12.4 Comments:

`lock_file()` attempts to place a lock on an open file, `fn`, to stop other processes from using the file while your program is reading it or writing it.

Under *Unix*, there are two types of locks that you can request using the `t` parameter. (Under *WIN32* the parameter `t` is ignored, but should be an integer.) Ask for a **shared** lock when you intend to read a file, and you want to temporarily block other processes from writing it. Ask for an **exclusive** lock when you intend to write to a file and you want to temporarily block other processes from reading or writing it. It's ok for many processes to simultaneously have shared locks on the same file, but only one process can have an exclusive lock, and that can happen only when no other process has any kind of lock on the file. *io.e* contains the following declarations:

```
public enum
    LOCK_SHARED,
    LOCK_EXCLUSIVE
```

On *WIN32* you can lock a specified portion of a file using the `r` parameter. `r` is a sequence of the form: `{first_byte, last_byte}`. It indicates the first byte and last byte in the file, that the lock applies to. Specify the empty sequence `{ }`, if you want to lock the whole file, or don't specify it at all, as this is the default. In the current release for *Unix*, locks always apply to the whole file, and you should use this default value.

`lock_file()` does not wait for other processes to relinquish their locks. You may have to call it repeatedly, before the lock request is granted.

On *Unix*, these locks are called advisory locks, which means they aren't enforced by the operating system. It is up to the processes that use a particular file to cooperate with each other. A process can access a file without first obtaining a lock on it. On *WIN32* locks are enforced by the operating system.

8.5.3.12.5 Example 1:

```
include std/io.e
integer v
atom t
v = open("visitor_log", "a") -- open for append
t = time()
while not lock_file(v, LOCK_EXCLUSIVE, {}) do
  if time() > t + 60 then
    puts(STDOUT, "One minute already ... I can't wait forever!\n")
    abort(1)
  end if
  sleep(5) -- let other processes run
end while
puts(v, "Yet another visitor\n")
unlock_file(v, {})
close(v)
```

8.5.3.12.6 See Also:

[unlock_file](#)

8.5.3.13 unlock_file

```
include std/io.e
public procedure unlock_file(file_number fn, byte_range r = {})
```

Unlock (a portion of) an open file.

8.5.3.13.1 Parameters:

1. *fn* : an integer, the handle to the file or device to (partially) lock.
2. *r* : a sequence, defining a section of the file to be locked, or { } for the whole file (the default).

8.5.3.13.2 Errors:

The target file or device must be open.

8.5.3.13.3 Comments:

You must have previously locked the file using `lock_file()`. On *WIN32* you can unlock a range of bytes within a file by specifying the *r* as {first_byte, last_byte}. The same range of bytes must have been locked by a previous call to [lock_file](#)(). On *Unix* you can currently only lock or unlock an entire file. *r* should be { } when you want to unlock an entire file. On *Unix*, *r* must always be { }, which is the default.

You should unlock a file as soon as possible so other processes can use it.

Any files that you have locked, will automatically be unlocked when your program terminates.

8.5.3.13.4 See Also:

[lock_file](#)

8.5.4 File Reading/Writing

8.5.4.1 read_lines

```
include std/io.e
public function read_lines(object file)
```

Read the contents of a file as a sequence of lines.

8.5.4.1.1 Parameters:

`file` : an object, either a file path or the handle to an open file. If this is an empty string, STDIN (the console) is used.

8.5.4.1.2 Returns:

A **sequence**, made of lines from the file, as [gets](#) could read them.

8.5.4.1.3 Comments:

If `file` was a sequence, the file will be closed on completion. Otherwise, it will remain open, but at end of file.

8.5.4.1.4 Example 1:

```
data = read_lines("my_file.txt")
-- data contains the entire contents of ##my_file.txt##, 1 sequence per line:
-- {"Line 1", "Line 2", "Line 3"}
```

8.5.4.1.5 Example 2:

```
fh = open("my_file.txt", "r")
data = read_lines(fh)
close(fh)

-- data contains the entire contents of ##my_file.txt##, 1 sequence per line:
```

```
-- {"Line 1", "Line 2", "Line 3"}
```

8.5.4.1.6 See Also:

[gets](#), [write_lines](#), [read_file](#)

8.5.4.2 process_lines

```
include std/io.e
public function process_lines(object file, integer proc, object user_data = 0)
```

Process the contents of a file, one line at a time.

8.5.4.2.1 Parameters:

1. `file` : an object. Either a file path or the handle to an open file. An empty string signifies STDIN - the console keyboard.
2. `proc` : an integer. The routine_id of a function that will process the line.
3. `user_data` : on object. This is passed untouched to `proc` for each line.

8.5.4.2.2 Returns:

An object. If 0 then all the file was processed successfully. Anything else means that something went wrong and this is whatever value was returned by `proc`.

8.5.4.2.3 Comments:

- The function `proc` must accept three parameters ...
 - ◆ A sequence: The line to process. It will **not** contain an end-of-line character.
 - ◆ An integer: The line number.
 - ◆ An object : This is the `user_data` that was passed to `process_lines`.
- If `file` was a sequence, the file will be closed on completion. Otherwise, it will remain open, and be positioned where ever reading stopped.

8.5.4.2.4 Example:

```
-- Format each supplied line according to the format pattern supplied as well.
function show(sequence aLine, integer line_no, object data)
  writefln( data[1], {line_no, aLine})
  if data[2] > 0 and line_no = data[2] then
    return 1
  else
    return 0
  end if
end function
```

```
-- Show the first 20 lines.  
process_lines("sample.txt", routine_id("show"), {"[1z:4] : [2]", 20})
```

8.5.4.2.5 See Also:

[gets](#), [read_lines](#), [read_file](#)

8.5.4.3 write_lines

```
include std/io.e  
public function write_lines(object file, sequence lines)
```

Write a sequence of lines to a file.

8.5.4.3.1 Parameters:

1. `file` : an object, either a file path or the handle to an open file.
2. `lines` : the sequence of lines to write

8.5.4.3.2 Returns:

An **integer**, 1 on success, -1 on failure.

8.5.4.3.3 Errors:

If [puts\(\)](#) cannot write some line of text, a runtime error will occur.

8.5.4.3.4 Comments:

If `file` was a sequence, the file will be closed on completion. Otherwise, it will remain open, but at end of file.

Whatever integer the lines in `lines` holds will be truncated to its 8 lowest bits so as to fall in the 0.255 range.

8.5.4.3.5 Example 1:

```
if write_lines("data.txt", {"This is important data", "Goodbye"}) != -1 then  
    puts(STDERR, "Failed to write data\n")  
end if
```

8.5.4.3.6 See Also:

[read_lines](#), [write_file](#), [puts](#)

8.5.4.4 `append_lines`

```
include std/io.e
public function append_lines(sequence file, sequence lines)
```

Append a sequence of lines to a file.

8.5.4.4.1 Parameters:

1. `file` : an object, either a file path or the handle to an open file.
2. `lines` : the sequence of lines to write

8.5.4.4.2 Returns:

An **integer**, 1 on success, -1 on failure.

8.5.4.4.3 Errors:

If [puts](#)() cannot write some line of text, a runtime error will occur.

8.5.4.4.4 Comments:

`file` is opened, written to and then closed.

8.5.4.4.5 Example 1:

```
if append_lines("data.txt", {"This is important data", "Goodbye"}) != -1 then
    puts(STDERR, "Failed to append data\n")
end if
```

8.5.4.4.6 See Also:

[write_lines](#), [puts](#)

8.5.4.5 BINARY_MODE

```
include std/io.e
public enum BINARY_MODE
```

8.5.4.6 TEXT_MODE

```
include std/io.e
public enum TEXT_MODE
```

8.5.4.7 UNIX_TEXT

```
include std/io.e
public enum UNIX_TEXT
```

8.5.4.8 DOS_TEXT

```
include std/io.e
public enum DOS_TEXT
```

8.5.4.9 ANSI

```
include std/io.e
public enum ANSI
```

8.5.4.10 UTF

```
include std/io.e
public enum UTF
```

8.5.4.11 UTF_8

```
include std/io.e
public enum UTF_8
```

8.5.4.12 UTF_16

```
include std/io.e
public enum UTF_16
```

8.5.4.13 UTF_16BE

```
include std/io.e
public enum UTF_16BE
```

8.5.4.14 UTF_16LE

```
include std/io.e
public enum UTF_16LE
```

8.5.4.15 UTF_32

```
include std/io.e
public enum UTF_32
```

8.5.4.16 UTF_32BE

```
include std/io.e
public enum UTF_32BE
```

8.5.4.17 UTF_32LE

```
include std/io.e
public enum UTF_32LE
```

8.5.4.18 read_file

```
include std/io.e
public function read_file(object file, integer as_text = BINARY_MODE, integer encoding = ANSI)
```

Read the contents of a file as a single sequence of bytes.

8.5.4.18.1 Parameters:

1. `file`: an object, either a file path or the handle to an open file.
2. `as_text`: integer, **BINARY_MODE** (the default) assumes *binary mode* that causes every byte to be read in, and **TEXT_MODE** assumes *text mode* that ensures that lines end with just a Ctrl-J (NewLine) character, and the first byte value of 26 (Ctrl-Z) is interpreted as End-Of-File.
3. `encoding`: An integer. One of ANSI, UTF, UTF_8, UTF_16, UTF_16BE, UTF_16LE, UTF_32, UTF_32BE, UTF_32LE. The default is ANSI.

8.5.4.18.2 Returns:

A **sequence**, holding the entire file.

Comments

- When using **BINARY_MODE**, each byte in the file is returned as an element in the return sequence.
- When not using **BINARY_MODE**, the file will be interpreted as a text file. This means that all line endings will be transformed to a single 0x0A character and the first 0x1A character (Ctrl-Z) will indicate the end of file (all data after this will not be returned to the caller.)
- Text files are always returned as UTF_32 encoded files.
- Encoding ...
 - ◆ ANSI: no interpretation of the file data is done. All bytes are simply returned as characters.
 - ◆ UTF: The file data is examined to work out which UTF encoding method was used to create the file. If the file starts with a valid Byte Order Marker (BOM) it can quickly decide between UTF_8, UTF_16 and UTF_32. For files without a BOM, if the file is completely valid UTF_8 encoding then that is what is used. Failing that, if there are no null bytes, the ANSI is assumed. Failing that, it is tested for being a valid UTF_16 or UTF_32 format. As a last resort, it will be assumed to be an ANSI file.
 - ◆ UTF_8: Any valid UTF_8 BOM is removed and the data is converted to UTF_32 format before returning. This means that if it contains any invalidly encoded Unicode characters, they will be ignored.
 - ◆ UTF_16: Any valid UTF_16 BOM is removed and the data is converted to UTF_32 format before returning. This means that if it contains any invalidly encoded Unicode characters, they will be ignored.
 - ◆ UTF_16LE: Any valid little-endian UTF_16 BOM is removed and the data is converted to UTF_32 format before returning. This means that if it contains any invalidly encoded Unicode characters, they will be ignored.
 - ◆ UTF_16BE: Any valid big-endian UTF_16 BOM is removed and the data is converted to UTF_32 format before returning. This means that if it contains any invalidly encoded Unicode characters, they will be ignored.
 - ◆ UTF_32: Any valid UTF_32 BOM is removed.
 - ◆ UTF_32LE: Any valid little-endian UTF_32 BOM is removed.
 - ◆ UTF_32BE: Any valid big-endian UTF_32 BOM is removed.
- If one of the UTF_32 encodings is supplied, invalid Unicode characters are not stripped out but are returned in the file data.

8.5.4.18.3 Example 1:

```
data = read_file("my_file.txt")
-- data contains the entire contents of ##my_file.txt##
```

8.5.4.18.4 Example 2:

```
fh = open("my_file.txt", "r")
data = read_file(fh)
close(fh)

-- data contains the entire contents of ##my_file.txt##
```

8.5.4.18.5 Example 3:

```
data = read_file("my_file.txt", TEXT_MODE, UTF_8)
-- The UTF encoded contents of ##my_file.txt## is stored in 'data' as UTF_32
```

8.5.4.18.6 See Also:

[write_file](#), [read_lines](#)

8.5.4.19 write_file

```
include std/io.e
public function write_file(object file, sequence data, integer as_text = BINARY_MODE, integer encoding = ANSI, integer with_bom = 0)
```

Write a sequence of bytes to a file.

8.5.4.19.1 Parameters:

1. `file`: an object, either a file path or the handle to an open file.
2. `data`: the sequence of bytes to write
3. `as_text`: integer
 - ◆ **BINARY_MODE** (the default) assumes *binary mode* that causes every byte to be written out as is,
 - ◆ **TEXT_MODE** assumes *text mode* that causes a NewLine to be written out according to the operating system's end of line convention. In Unix this is Ctrl-J and in Windows this is the pair {Ctrl-L, Ctrl-J}.
 - ◆ **UNIX_TEXT** ensures that lines are written out with unix style line endings (Ctrl-J).
 - ◆ **DOS_TEXT** ensures that lines are written out with Windows style line endings {Ctrl-L, Ctrl-J}.
4. `encoding`: an integer. One of ANSI, UTF_8, UTF_16LE, UTF_16BE, UTF_32LE, UTF_32BE. The default is ANSI.
5. `with_bom`: an integer. Either 0 or 1. If 1 then when encoding as a UTF file, this will prepend a Byte

Order Marker (BOM) to the file output.

8.5.4.19.2 Returns:

An **integer**, 1 on success, -1 on failure.

8.5.4.19.3 Comments:

- UTF_16LE, and UTF_32LE create little-endian files, which are the normal ones for Intel based CPUs. Big-endian files are more commonly found on Motorola CPUs.

8.5.4.19.4 Errors:

If [puts](#) cannot write data, a runtime error will occur.

8.5.4.19.5 Comments:

- When `file` is a file handle, the file is not closed after writing is finished. When `file` is a file name, it is opened, written to and then closed.
- Note that when writing the file in any of the text modes, the file is truncated at the first Ctrl-Z character in the input data.

8.5.4.19.6 Example 1:

```
if write_file("data.txt", "This is important data\nGoodbye") = -1 then
    puts(STDERR, "Failed to write data\n")
end if
```

8.5.4.19.7 See Also:

[read_file](#), [write_lines](#)

8.5.4.20 `wprintf`

```
include std/io.e
public procedure wprintf(object fm, object data = {}, object fn = 1, object data_not_string = 0)
```

Write formatted text to a file..

8.5.4.20.1 Parameters:

There are two ways to pass arguments to this function,

1. Traditional way with first arg being a file handle.
 1. : integer, The file handle.
 2. : sequence, The format pattern.
 3. : object, The data that will be formatted.
 4. `data_not_string`: object, If not 0 then the data is not a string. By default this is 0 meaning that data could be a single string.
2. Alternative way with first argument being the format pattern.
 1. : sequence, Format pattern.
 2. : sequence, The data that will be formatted,
 3. : object, The file to receive the formatted output. Default is to the STDOUT device (console).
 4. `data_not_string`: object, If not 0 then the data is not a string. By default this is 0 meaning that data could be a single string.

8.5.4.20.2 Comments:

- With the traditional arguments, the first argument must be an integer file handle.
- With the alternative arguments, the third argument can be a file name string, in which case it is opened for output, written to and then closed.
- With the alternative arguments, the third argument can be a two-element sequence containing a file name string and an output type ("a" for append, "w" for write), in which case it is opened accordingly, written to and then closed.
- With the alternative arguments, the third argument can a file handle, in which case it is written to only
- The format pattern uses the formatting codes defined in [text.e:format](#).
- When the data to be formatted is a single text string, it does not have to be enclosed in braces,

8.5.4.20.3 Example 1:

```
-- To console
writef("Today is [4], [u2:3] [3:02], [1:4].", {Year, MonthName, Day, DayName})
-- To "sample.txt"
writef("Today is [4], [u2:3] [3:02], [1:4].", {Year, MonthName, Day, DayName}, "sample.txt")
-- To "sample.dat"
integer dat = open("sample.dat", "w")
writef("Today is [4], [u2:3] [3:02], [1:4].", {Year, MonthName, Day, DayName}, dat)
-- Appended to "sample.log"
writef("Today is [4], [u2:3] [3:02], [1:4].", {Year, MonthName, Day, DayName}, {"sample.log", "a"})
-- Simple message to console
writef("A message")
-- Another console message
writef(STDERR, "This is a []", "message")
-- Outputs two numbers
writef(STDERR, "First [], second []", {65, 100},, 1) -- Note that {65, 100} is also "Ad"
```

8.5.4.20.4 See Also:

[text.e:format](#), [writeln](#), [write_lines](#)

8.5.4.21 writeln

```
include std/io.e
public procedure writeln(object fm, object data = {}, object fn = 1, object data_not_string =
```

Write formatted text to a file, ensuring that a new line is also output.

8.5.4.21.1 Parameters:

1. `fm` : sequence, Format pattern.
2. `data` : sequence, The data that will be formatted,
3. `fn` : object, The file to receive the formatted output. Default is to the STDOUT device (console).
4. `data_not_string`: object, If not 0 then the data is not a string. By default this is 0 meaning that data could be a single string.

8.5.4.21.2 Comments:

- This is the same as [writef](#), except that it always adds a New Line to the output.
- When `fn` is a file name string, it is opened for output, written to and then closed.
- When `fn` is a two-element sequence containing a file name string and an output type ("a" for append, "w" for write), it is opened accordingly, written to and then closed.
- When `fn` is a file handle, it is written to only
- The `fm` uses the formatting codes defined in [text.e:format](#).

8.5.4.21.3 Example 1:

```
-- To console
writeln("Today is [4], [u2:3] [3:02], [1:4].", {Year, MonthName, Day, DayName})
-- To "sample.txt"
writeln("Today is [4], [u2:3] [3:02], [1:4].", {Year, MonthName, Day, DayName}, "sample.txt")
-- Appended to "sample.log"
writeln("Today is [4], [u2:3] [3:02], [1:4].", {Year, MonthName, Day, DayName}, {"sample.log",
```

8.5.4.21.4 See Also:

[text.e:format](#), [writef](#), [write_lines](#)

8.6 Math

Sign and comparisons

- abs
- sign
- max
- min
- ensure_in_range
- ensure_in_list

Roundings and remainders

- remainder
- mod
- trunc
- frac
- intdiv
- floor
- ceil
- round

Trigonometry

- arctan
- tan
- cos
- sin
- arccos
- arcsin
- atan2
- rad2deg
- deg2rad

Logarithms and powers.

- log
- log10
- exp
- power
- sqrt

Hyperbolic trigonometry

- cosh
- sinh
- tanh
- arsinh
- arcosh
- artanh

Accumulation

- sum
- product
- or_all

Bitwise operations

- and_bits
- xor_bits

or_bits
not_bits
shift_bits
rotate_bits
gcd
approx
powof2
is_even
is_even_obj

8.6.1 Sign and comparisons

8.6.1.1 abs

```
include std/math.e  
public function abs(object a)
```

Returns the absolute value of numbers.

8.6.1.1.1 Parameters:

1. `value` : an object, each atom is processed, no matter how deeply nested.

8.6.1.1.2 Returns:

An **object**, the same shape as `value`. When `value` is an atom, the result is the same if not less than zero, and the opposite value otherwise.

8.6.1.1.3 Comments:

This function may be applied to an atom or to all elements of a sequence

8.6.1.1.4 Example 1:

```
x = abs({10.5, -12, 3})  
-- x is {10.5, 12, 3}  
  
i = abs(-4)  
-- i is 4
```

8.6.1.1.5 See Also:[sign](#)**8.6.1.2 sign**

```
include std/math.e
public function sign(object a)
```

Return -1, 0 or 1 for each element according to it being negative, zero or positive

8.6.1.2.1 Parameters:

1. `value` : an object, each atom of which will be acted upon, no matter how deeply nested.

8.6.1.2.2 Returns:

An **object**, the same shape as `value`. When `value` is an atom, the result is -1 if `value` is less than zero, 1 if greater and 0 if equal.

8.6.1.2.3 Comments:

This function may be applied to an atom or to all elements of a sequence.

For an atom, `sign(x)` is the same as [compare\(x,0\)](#).

8.6.1.2.4 Example 1:

```
i = sign(5)
i is 1

i = sign(0)
-- i is 0

i = sign(-2)
-- i is -1
```

8.6.1.2.5 See Also:[compare](#)

8.6.1.3 max

```
include std/math.e
public function max(object a)
```

Computes the maximum value among all the argument's elements

8.6.1.3.1 Parameters:

1. `values` : an object, all atoms of which will be inspected, no matter how deeply nested.

8.6.1.3.2 Returns:

An **atom**, the maximum of all atoms in `flatten(values)`.

8.6.1.3.3 Comments:

This function may be applied to an atom or to a sequence of any shape.

8.6.1.3.4 Example 1:

```
a = max({10, 15.4, 3})
-- a is 15.4
```

8.6.1.3.5 See Also:

[min](#), [compare](#), [flatten](#)

8.6.1.4 min

```
include std/math.e
public function min(object a)
```

Computes the minimum value among all the argument's elements

8.6.1.4.1 Parameters:

1. `values` : an object, all atoms of which will be inspected, no matter how deeply nested.

8.6.1.4.2 Returns:

An **atom**, the minimum of all atoms in `flatten(values)`.

8.6.1.4.3 Comments:

This function may be applied to an atom or to a sequence of any shape.

8.6.1.4.4 Example 1:

```
a = min({10,15.4,3})  
-- a is 3
```

8.6.1.5 ensure_in_range

```
include std/math.e  
public function ensure_in_range(object item, sequence range_limits)
```

Ensures that the `item` is in a range of values supplied by inclusive `range_limits`

8.6.1.5.1 Parameters:

1. `item`: The object to test for.
2. `range_limits`: A sequence of two or more elements. The first is assumed to be the smallest value and the last is assumed to be the highest value.

8.6.1.5.2 Returns:

A **object**. If `item` is lower than the first item in the `range_limits` it returns the first item. If `item` is higher than the last element in the `range_limits` it returns the last item. Otherwise it returns `item`.

8.6.1.5.3 Example 1:

```
object valid_data = ensure_in_range(user_data, {2, 75})  
if not equal(valid_data, user_data) then  
    errmsg("Invalid input supplied. Using %d instead.", valid_data)  
end if  
procA(valid_data)
```

8.6.1.6 ensure_in_list

```
include std/math.e
public function ensure_in_list(object item, sequence list, integer default = 1)
```

Ensures that the `item` is in a list of values supplied by `list`

8.6.1.6.1 Parameters:

1. `item` : The object to test for.
2. `list` : A sequence of elements that `item` should be a member of.
3. `default` : an integer, the index of the list item to return if `item` is not found. Defaults to 1.

8.6.1.6.2 Returns:

An **object**, if `item` is not in the list, it returns the list item of index `default`, otherwise it returns `item`.

8.6.1.6.3 Comments:

If `default` is set to an invalid index, the first item on the list is returned instead when `item` is not on the list.

8.6.1.6.4 Example 1:

```
object valid_data = ensure_in_list(user_data, {100, 45, 2, 75, 121})
if not equal(valid_data, user_data) then
    errmsg("Invalid input supplied. Using %d instead.", valid_data)
end if
procA(valid_data)
```

8.6.2 Roundings and remainders

8.6.2.1 remainder

```
<built-in> function remainder(object dividend, object divisor)
```

Compute the remainder of the division of two objects using truncated division.

8.6.2.1.1 Parameters:

1. `dividend` : any EUPHORIA object.
2. `divisor` : any EUPHORIA object.

8.6.2.1.2 Returns:

An **object**, the shape of which depends on `dividend`'s and `divisor`'s. For two atoms, this is the remainder of dividing `dividend` by `divisor`, with `dividend`'s sign.

8.6.2.1.3 Errors:

1. If any atom in `divisor` is 0, this is an error condition as it amounts to an attempt to divide by zero.
2. If both `dividend` and `divisor` are sequences, they must be the same length as each other.

8.6.2.1.4 Comments:

- There is a integer `N` such that `dividend = N * divisor + result`.
- The result has the sign of `dividend` and lesser magnitude than `divisor`.
- The result has the same sign as the `dividend`.
- This differs from `mod()` in that when the operands' signs are different this function rounds `dividend/divisor` towards zero whereas `mod()` rounds away from zero.

The arguments to this function may be atoms or sequences. The rules for [operations on sequences](#) apply, and determine the shape of the returned object.

8.6.2.1.5 Example 1:

```
a = remainder(9, 4)
-- a is 1
```

8.6.2.1.6 Example 2:

```
s = remainder({81, -3.5, -9, 5.5}, {8, -1.7, 2, -4})
-- s is {1, -0.1, -1, 1.5}
```

8.6.2.1.7 Example 3:

```
s = remainder({17, 12, 34}, 16)
-- s is {1, 12, 2}
```

8.6.2.1.8 Example 4:

```
s = remainder(16, {2, 3, 5})
-- s is {0, 1, 1}
```

8.6.2.1.9 See Also:

[mod](#), [Relational operators](#), [Operations on sequences](#)

8.6.2.2 mod

```
include std/math.e
public function mod(object x, object y)
```

Compute the remainder of the division of two objects using floored division.

8.6.2.2.1 Parameters:

1. `dividend`: any EUPHORIA object.
2. `divisor`: any EUPHORIA object.

8.6.2.2.2 Returns:

An **object**, the shape of which depends on `dividend`'s and `divisor`'s. For two atoms, this is the remainder of dividing `dividend` by `divisor`, with `divisor`'s sign.

8.6.2.2.3 Comments:

- There is a integer `N` such that `dividend = N * divisor + result`.
- The result is non-negative and has lesser magnitude than `divisor`. `n` needs not fit in an EUPHORIA integer.
- The result has the same sign as the dividend.
- The arguments to this function may be atoms or sequences. The rules for [operations on sequences](#) apply, and determine the shape of the returned object.
- When both arguments have the same sign, `mod()` and [remainder\(\)](#) return the same result.
- This differs from [remainder\(\)](#) in that when the operands' signs are different this function rounds `dividend/divisor` away from zero whereas `remainder()` rounds towards zero.

8.6.2.2.4 Example 1:

```
a = mod(9, 4)
-- a is 1
```

8.6.2.2.5 Example 2:

```
s = mod({81, -3.5, -9, 5.5}, {8, -1.7, 2, -4})
-- s is {1,-0.1,1,-2.5}
```

8.6.2.2.6 Example 3:

```
s = mod({17, 12, 34}, 16)
-- s is {1, 12, 2}
```

8.6.2.2.7 Example 4:

```
s = mod(16, {2, 3, 5})
-- s is {0, 1, 1}
```

8.6.2.2.8 See Also:

[remainder](#), [Relational operators](#), [Operations on sequences](#)

8.6.2.3 trunc

```
include std/math.e
public function trunc(object x)
```

Return the integer portion of a number.

8.6.2.3.1 Parameters:

1. `value` : any EUPHORIA object.

8.6.2.3.2 Returns:

An **object**, the shape of which depends on `value`'s. Each item in the returned object will be an integer. These are the same corresponding items in `value` except with any fractional portion removed.

8.6.2.3.3 Comments:

- This is essentially done by always rounding towards zero. The [floor\(\)](#) function rounds towards negative infinity, which means it rounds towards zero for positive values and away from zero for negative values.
- Note that `trunc(x) + frac(x) = x`

8.6.2.3.4 Example 1:

```
a = trunc(9.4)
-- a is 9
```


8.6.2.3.5 Example 2:

```
s = trunc({81, -3.5, -9.999, 5.5})
-- s is {81,-3, -9, 5}
```

8.6.2.3.6 See Also:

[floor](#) [frac](#)

8.6.2.4 frac

```
include std/math.e
public function frac(object x)
```

Return the fractional portion of a number.

8.6.2.4.1 Parameters:

1. `value` : any EUPHORIA object.

8.6.2.4.2 Returns:

An **object**, the shape of which depends on `value`'s. Each item in the returned object will be the same corresponding items in `value` except with the integer portion removed.

8.6.2.4.3 Comments:

Note that `trunc(x) + frac(x) = x`

8.6.2.4.4 Example 1:

```
a = frac(9.4)
-- a is 0.4
```

8.6.2.4.5 Example 2:

```
s = frac({81, -3.5, -9.999, 5.5})
-- s is {0, -0.5, -0.999, 0.5}
```

8.6.2.4.6 See Also:

[trunc](#)

8.6.2.5 intdiv

```
include std/math.e
public function intdiv(object a, object b)
```

Return an integral division of two objects.

8.6.2.5.1 Parameters:

1. `divided`: any EUPHORIA object.
2. `divisor`: any EUPHORIA object.

8.6.2.5.2 Returns:

An **object**, which will be a sequence if either `dividend` or `divisor` is a sequence.

8.6.2.5.3 Comments:

- This calculates how many non-empty sets when `dividend` is divided by `divisor`.
- The result's sign is the same as the dividend's sign.

8.6.2.5.4 Example 1:

```
object Tokens = 101
object MaxPerEnvelope = 5
integer Envelopes = intdiv( Tokens, MaxPerEnvelope) --> 21
```

8.6.2.6 floor

```
<built-in> function floor(object value)
```

Rounds `value` down to the next integer less than or equal to `value`. It does not simply truncate the fractional part, but actually rounds towards negative infinity.

8.6.2.6.1 Parameters:

1. `value`: any EUPHORIA object; each atom in `value` will be acted upon.

8.6.2.6.2 Returns:

An **object**, the same shape as `value` but with each item guaranteed to be an integer less than or equal to the corresponding item in `value`.

8.6.2.6.3 Example 1:

```
y = floor({0.5, -1.6, 9.99, 100})
-- y is {0, -2, 9, 100}
```

8.6.2.6.4 See Also:

[ceil](#), [round](#)

8.6.2.7 ceil

```
include std/math.e
public function ceil(object a)
```

Computes the next integer equal or greater than the argument.

8.6.2.7.1 Parameters:

1. `value` : an object, each atom of which processed, no matter how deeply nested.

8.6.2.7.2 Returns:

An **object**, the same shape as `value`. Each atom in `value` is returned as an integer that is the smallest integer equal to or greater than the corresponding atom in `value`.

8.6.2.7.3 Comments:

This function may be applied to an atom or to all elements of a sequence.

`ceil(X)` is 1 more than `floor(X)` for non-integers. For integers, `X = floor(X) = ceil(X)`.

8.6.2.7.4 Example 1:

```
sequence nums
nums = {8, -5, 3.14, 4.89, -7.62, -4.3}
nums = ceil(nums) -- {8, -5, 4, 5, -7, -4}
```

8.6.2.7.5 See Also:

[floor](#), [round](#)

8.6.2.8 round

```
include std/math.e
public function round(object a, object precision = 1)
```

Return the argument's elements rounded to some precision

8.6.2.8.1 Parameters:

1. `value` : an object, each atom of which will be acted upon, no matter how deeply nested.
2. `precision` : an object, the rounding precision(s). If not passed, this defaults to 1.

8.6.2.8.2 Returns:

An **object**, the same shape as `value`. When `value` is an atom, the result is that atom rounded to the nearest integer multiple of `1/precision`.

8.6.2.8.3 Comments:

This function may be applied to an atom or to all elements of a sequence.

8.6.2.8.4 Example 1:

```
round(5.2) -- 5
round({4.12, 4.67, -5.8, -5.21}, 10) -- {4.1, 4.7, -5.8, -5.2}
round(12.2512, 100) -- 12.25
```

8.6.2.8.5 See Also:

[floor](#), [ceil](#)

8.6.3 Trigonometry

8.6.3.1 arctan

```
<built-in> function arctan(object tangent)
```

Return an angle with given tangent.

8.6.3.1.1 Parameters:

1. `tangent` : an object, each atom of which will be converted, no matter how deeply nested.

8.6.3.1.2 Returns:

An **object**, of the same shape as `tangent`. For each atom in `flatten(tangent)`, the angle with smallest magnitude that has this atom as tangent is computed.

8.6.3.1.3 Comments:

All atoms in the returned value lie between $-\pi/2$ and $\pi/2$, exclusive.

This function may be applied to an atom or to all elements of a sequence (of sequence (...)).

`arctan()` is faster than `arcsin()` or `arccos()`.

8.6.3.1.4 Example 1:

```
s = arctan({1,2,3})
-- s is {0.785398, 1.10715, 1.24905}
```

8.6.3.1.5 See Also:

[arcsin](#), [arccos](#), [tan](#), [flatten](#)

8.6.3.2 tan

```
<built-in> function tan(object angle)
```

Return the tangent of an angle, or a sequence of angles.

8.6.3.2.1 Parameters:

1. `angle` : an object, each atom of which will be converted, no matter how deeply nested.

8.6.3.2.2 Returns:

An **object**, of the same shape as `angle`. Each atom in the flattened `angle` is replaced by its tangent.

8.6.3.2.3 Errors:

If any atom in `angle` is an odd multiple of $\pi/2$, an error occurs, as its tangent would be infinite.

8.6.3.2.4 Comments:

This function may be applied to an atom or to all elements of a sequence of arbitrary shape, recursively.

8.6.3.2.5 Example 1:

```
t = tan(1.0)
-- t is 1.55741
```

8.6.3.2.6 See Also:

[sin](#), [cos](#), [arctan](#)

8.6.3.3 cos

```
<built-in> function cos(object angle)
```

Return the cosine of an angle expressed in radians

8.6.3.3.1 Parameters:

1. `angle` : an object, each atom of which will be converted, no matter how deeply nested.

8.6.3.3.2 Returns:

An **object**, the same shape as `angle`. Each atom in `angle` is turned into its cosine.

8.6.3.3.3 Comments:

This function may be applied to an atom or to all elements of a sequence.

The cosine of an angle is an atom between -1 and 1 inclusive. 0.0 is hit by odd multiples of $\pi/2$ only.

8.6.3.3.4 Example 1:

```
x = cos({.5, .6, .7})  
-- x is {0.8775826, 0.8253356, 0.7648422}
```

8.6.3.3.5 See Also:

[sin](#), [tan](#), [arccos](#), [PI](#), [deg2rad](#)

8.6.3.4 sin

<built-in> `function sin(object angle)`

Return the sine of an angle expressed in radians

8.6.3.4.1 Parameters:

1. `angle` : an object, each atom in which will be acted upon.

8.6.3.4.2 Returns:

An **object**, the same shape as `angle`. When `angle` is an atom, the result is the sine of `angle`.

8.6.3.4.3 Comments:

This function may be applied to an atom or to all elements of a sequence.

The sine of an angle is an atom between -1 and 1 inclusive. 0.0 is hit by integer multiples of π only.

8.6.3.4.4 Example 1:

```
sin_x = sin({.5, .9, .11})  
-- sin_x is {.479, .783, .110}
```

8.6.3.4.5 See Also:

[cos](#), [arcsin](#), [PI](#), [deg2rad](#)

8.6.3.5 arccos

```
include std/math.e
public function arccos(trig_range x)
```

Return an angle given its cosine.

8.6.3.5.1 Parameters:

1. `value` : an object, each atom in which will be acted upon.

8.6.3.5.2 Returns:

An **object**, the same shape as `value`. When `value` is an atom, the result is an atom, an angle whose cosine is `value`.

8.6.3.5.3 Errors:

If any atom in `value` is not in the `-1..1` range, it cannot be the cosine of a real number, and an error occurs.

8.6.3.5.4 Comments:

A value between 0 and [PI](#) radians will be returned.

This function may be applied to an atom or to all elements of a sequence.

`arccos()` is not as fast as [arctan\(\)](#).

8.6.3.5.5 Example 1:

```
s = arccos({-1,0,1})
-- s is {3.141592654, 1.570796327, 0}
```

8.6.3.5.6 See Also:

[cos](#), [PI](#), [arctan](#)

8.6.3.6 arcsin

```
include std/math.e
public function arcsin(trig_range x)
```

Return an angle given its sine.

8.6.3.6.1 Parameters:

1. `value` : an object, each atom in which will be acted upon.

8.6.3.6.2 Returns:

An **object**, the same shape as `value`. When `value` is an atom, the result is an atom, an angle whose sine is `value`.

8.6.3.6.3 Errors:

If any atom in `value` is not in the $-1..1$ range, it cannot be the sine of a real number, and an error occurs.

8.6.3.6.4 Comments:

A value between $-\pi/2$ and $+\pi/2$ (radians) inclusive will be returned.

This function may be applied to an atom or to all elements of a sequence.

`arcsin()` is not as fast as `arctan()`.

8.6.3.6.5 Example 1:

```
s = arcsin({-1,0,1})
s is {-1.570796327, 0, 1.570796327}
```

8.6.3.6.6 See Also:

[arccos](#), [arccos](#), [sin](#)

8.6.3.7 atan2

```
include std/math.e
public function atan2(atom y, atom x)
```

Calculate the arctangent of a ratio.

8.6.3.7.1 Parameters:

1. `y` : an atom, the numerator of the ratio
2. `x` : an atom, the denominator of the ratio

8.6.3.7.2 Returns:

An **atom**, which is equal to `arctan(y/x)`, except that it can handle zero denominator and is more accurate.

8.6.3.7.3 Example 1:

```
a = atan2(10.5, 3.1)
-- a is 1.283713958
```

8.6.3.7.4 See Also:

[arctan](#)

8.6.3.8 rad2deg

```
include std/math.e
public function rad2deg(object x)
```

Convert an angle measured in radians to an angle measured in degrees

8.6.3.8.1 Parameters:

1. `angle` : an object, all atoms of which will be converted, no matter how deeply nested.

8.6.3.8.2 Returns:

An **object**, the same shape as `angle`, all atoms of which were multiplied by $180/\pi$.

8.6.3.8.3 Comments:

This function may be applied to an atom or sequence. A flat angle is π radians and 180 degrees.

[arcsin\(\)](#), [arccos\(\)](#) and [arctan\(\)](#) return angles in radians.

8.6.3.8.4 Example 1:

```
x = rad2deg(3.385938749)
-- x is 194
```

8.6.3.8.5 See Also:

[deg2rad](#)

8.6.3.9 deg2rad

```
include std/math.e
public function deg2rad(object x)
```

Convert an angle measured in degrees to an angle measured in radians

8.6.3.9.1 Parameters:

1. `angle` : an object, all atoms of which will be converted, no matter how deeply nested.

8.6.3.9.2 Returns:

An **object**, the same shape as `angle`, all atoms of which were multiplied by $\text{PI}/180$.

8.6.3.9.3 Comments:

This function may be applied to an atom or sequence. A flat angle is PI radians and 180 degrees. [sin\(\)](#), [cos\(\)](#) and [tan\(\)](#) expect angles in radians.

8.6.3.9.4 Example 1:

```
x = deg2rad(194)
-- x is 3.385938749
```

8.6.3.9.5 See Also:

[rad2deg](#)

8.6.4 Logarithms and powers.

8.6.4.1 log

```
<built-in> function log(object value)
```

Return the natural logarithm of a positive number.

8.6.4.1.1 Parameters:

1. `value` : an object, any atom of which `log()` acts upon.

8.6.4.1.2 Returns:

An **object**, the same shape as `value`. For an atom, the returned atom is its logarithm of base E.

8.6.4.1.3 Errors:

If any atom in `value` is not greater than zero, an error occurs as its logarithm is not defined.

8.6.4.1.4 Comments:

This function may be applied to an atom or to all elements of a sequence.

To compute the inverse, you can use `power(E, x)` where E is 2.7182818284590452, or equivalently `exp(x)`. Beware that the logarithm grows very slowly with x, so that `exp()` grows very fast.

8.6.4.1.5 Example 1:

```
a = log(100)
-- a is 4.60517
```

8.6.4.1.6 See Also:

[E](#), [exp](#), [log10](#)

8.6.4.2 log10

```
include std/math.e
public function log10(object x1)
```

Return the base 10 logarithm of a number.

8.6.4.2.1 Parameters:

1. `value` : an object, each atom of which will be converted, no matter how deeply nested.

8.6.4.2.2 Returns:

An **object**, the same shape as `value`. When `value` is an atom, raising 10 to the returned atom yields `value` back.

8.6.4.2.3 Errors:

If any atom in `value` is not greater than zero, its logarithm is not a real number and an error occurs.

8.6.4.2.4 Comments:

This function may be applied to an atom or to all elements of a sequence.

`log10()` is proportional to `log()` by a factor of $1/\log(10)$, which is about 0.435.

8.6.4.2.5 Example 1:

```
a = log10(12)
-- a is 2.48490665
```

8.6.4.2.6 See Also:

[log](#)

8.6.4.3 exp

```
include std/math.e
public function exp(atom x)
```

Computes some power of E.

8.6.4.3.1 Parameters:

1. `value` : an object, all atoms of which will be acted upon, no matter how deeply nested.

8.6.4.3.2 Returns:

An **object**, the same shape as `value`. When `value` is an atom, its exponential is being returned.

8.6.4.3.3 Comments:

This function can be applied to a single atom or to a sequence of any shape.

Due to its rapid growth, the returned values start losing accuracy as soon as values are greater than 10. Values above 710 will cause an overflow in hardware.

8.6.4.3.4 Example 1:

```
x = exp(5.4)
-- x is 221.4064162
```

8.6.4.3.5 See Also:

[log](#)

8.6.4.4 power

```
<built-in> function power(object base, object exponent)
```

Raise a base value to some power.

8.6.4.4.1 Parameters:

1. `base` : an object, the value(s) to raise to some power.
2. `exponent` : an object, the exponent(s) to apply to base.

8.6.4.4.2 Returns:

An **object**, the shape of which depends on `base`'s and `exponent`'s. For two atoms, this will be `base` raised to the power `exponent`.

8.6.4.4.3 Errors:

If some atom in `base` is negative and is raised to a non integer exponent, an error will occur, as the result is undefined.

If 0 is raised to any negative power, this is the same as a zero divide and causes an error.

`power(0, 0)` is illegal, because there is not an unique value that can be assigned to that quantity.

8.6.4.4.4 Comments:

The arguments to this function may be atoms or sequences. The rules for [operations on sequences](#) apply.

Powers of 2 are calculated very efficiently.

Other languages have a `**` or `^` operator to perform the same action. But they don't have sequences.

8.6.4.4.5 Example 1:

```
? power(5, 2)
-- 25 is printed
```

8.6.4.4.6 Example 2:

```
? power({5, 4, 3.5}, {2, 1, -0.5})
-- {25, 4, 0.534522} is printed
```

8.6.4.4.7 Example 3:

```
? power(2, {1, 2, 3, 4})
-- {2, 4, 8, 16}
```

8.6.4.4.8 Example 4:

```
? power({1, 2, 3, 4}, 2)
-- {1, 4, 9, 16}
```

8.6.4.4.9 See Also:

[log](#), [Operations on sequences](#)

8.6.4.5 sqrt

```
<built-in> function sqrt(object value)
```

Calculate the square root of a number.

8.6.4.5.1 Parameters:

1. `value` : an object, each atom in which will be acted upon.

8.6.4.5.2 Returns:

An **object**, the same shape as `value`. When `value` is an atom, the result is the positive atom whose square is `value`.

8.6.4.5.3 Errors:

If any atom in `value` is less than zero, an error will occur, as no squared real can be less than zero.

8.6.4.5.4 Comments:

This function may be applied to an atom or to all elements of a sequence.

8.6.4.5.5 Example 1:

```
r = sqrt(16)
-- r is 4
```

8.6.4.5.6 See Also:

[power](#), [Operations on sequences](#)

8.6.5 Hyperbolic trigonometry

8.6.5.1 cosh

```
include std/math.e
public function cosh(object a)
```

Computes the hyperbolic cosine of an object.

8.6.5.1.1 Parameters:

1. `x` : the object to process.

8.6.5.1.2 Returns:

An **object**, the same shape as `x`, each atom of which was acted upon.

8.6.5.1.3 Comments:

The hyperbolic cosine grows like the exponential function.

For all reals, $\text{power}(\cosh(x), 2) - \text{power}(\sinh(x), 2) = 1$. Compare with ordinary trigonometry.

8.6.5.1.4 Example 1:

```
? cosh(LN2) -- prints out 1.25
```

8.6.5.1.5 See Also:

[cos](#), [sinh](#), [arccosh](#)

8.6.5.2 sinh

```
include std/math.e
public function sinh(object a)
```

Computes the hyperbolic sine of an object.

8.6.5.2.1 Parameters:

1. x : the object to process.

8.6.5.2.2 Returns:

An **object**, the same shape as x , each atom of which was acted upon.

8.6.5.2.3 Comments:

The hyperbolic sine grows like the exponential function.

For all reals, $\text{power}(\cosh(x), 2) - \text{power}(\sinh(x), 2) = 1$. Compare with ordinary trigonometry.

8.6.5.2.4 Example 1:

```
? sinh(LN2) -- prints out 0.75
```

8.6.5.2.5 See Also:

[cosh](#), [sin](#), [arcsinh](#)

8.6.5.3 tanh

```
include std/math.e
public function tanh(object a)
```

Computes the hyperbolic tangent of an object.

8.6.5.3.1 Parameters:

1. x : the object to process.

8.6.5.3.2 Returns:

An **object**, the same shape as x , each atom of which was acted upon.

8.6.5.3.3 Comments:

The hyperbolic tangent takes values from -1 to +1.

$\tanh()$ is the ratio $\sinh() / \cosh()$. Compare with ordinary trigonometry.

8.6.5.3.4 Example 1:

```
? tanh(LN2) -- prints out 0.6
```

8.6.5.3.5 See Also:

[cosh](#), [sinh](#), [tan](#), [arctanh](#)

8.6.5.4 arcsinh

```
include std/math.e
public function arcsinh(object a)
```

Computes the reverse hyperbolic sine of an object.

8.6.5.4.1 Parameters:

1. x : the object to process.

8.6.5.4.2 Returns:

An **object**, the same shape as x , each atom of which was acted upon.

8.6.5.4.3 Comments:

The hyperbolic sine grows like the logarithm function.

8.6.5.4.4 Example 1:

```
? arcsinh(1) -- prints out 0,4812118250596034
```

8.6.5.4.5 See Also:

[arccosh](#), [arcsin](#), [sinh](#)

8.6.5.5 arccosh

```
include std/math.e
public function arccosh(not_below_1 a)
```

Computes the reverse hyperbolic cosine of an object.

8.6.5.5.1 Parameters:

1. x : the object to process.

8.6.5.5.2 Returns:

An **object**, the same shape as x , each atom of which was acted upon.

8.6.5.5.3 Errors:

Since [cosh](#) only takes values not below 1, an argument below 1 causes an error.

8.6.5.5.4 Comments:

The hyperbolic cosine grows like the logarithm function.

8.6.5.5.5 Example 1:

```
? arccosh(1) -- prints out 0
```

8.6.5.5.6 See Also:

[arccos](#), [arcsinh](#), [cosh](#)

8.6.5.6 arctanh

```
include std/math.e
public function arctanh(abs_below_1 a)
```

Computes the reverse hyperbolic tangent of an object.

8.6.5.6.1 Parameters:

1. x : the object to process.

8.6.5.6.2 Returns:

An **object**, the same shape as x , each atom of which was acted upon.

8.6.5.6.3 Errors:

Since [tanh](#) only takes values between -1 and +1 excluded, an out of range argument causes an error.

8.6.5.6.4 Comments:

The hyperbolic cosine grows like the logarithm function.

8.6.5.6.5 Example 1:

```
? arctanh(1/2) -- prints out 0,5493061443340548456976
```

8.6.5.6.6 See Also:

[arccos](#), [arcsinh](#), [cosh](#)

8.6.6 Accumulation

8.6.6.1 sum

```
include std/math.e
public function sum(object a)
```

Compute the sum of all atoms in the argument, no matter how deeply nested

8.6.6.1.1 Parameters:

1. `values` : an object, all atoms of which will be added up, no matter how nested.

8.6.6.1.2 Returns:

An **atom**, the sum of all atoms in [flatten](#)(`values`).

8.6.6.1.3 Comments:

This function may be applied to an atom or to all elements of a sequence

8.6.6.1.4 Example 1:

```
a = sum({10, 20, 30})
-- a is 60

a = sum({10.5, {11.2} , 8.1})
-- a is 29.8
```

8.6.6.1.5 See Also:

[can_add](#), [product](#), [or_all](#)

8.6.6.2 product

```
include std/math.e
public function product(object a)
```

Compute the product of all the atom in the argument, no matter how deeply nested.

8.6.6.2.1 Parameters:

1. `values` : an object, all atoms of which will be multiplied up, no matter how nested.

8.6.6.2.2 Returns:

An **atom**, the product of all atoms in `flatten(values)`.

8.6.6.2.3 Comments:

This function may be applied to an atom or to all elements of a sequence

8.6.6.2.4 Example 1:

```
a = product({10, 20, 30})
-- a is 6000

a = product({10.5, {11.2} , 8.1})
-- a is 952.56
```

8.6.6.2.5 See Also:

[can_add](#), [sum](#), [or_all](#)

8.6.6.3 or_all

```
include std/math.e
public function or_all(object a)
```

Or's together all atoms in the argument, no matter how deeply nested.

8.6.6.3.1 Parameters:

1. `values` : an object, all atoms of which will be added up, no matter how nested.

8.6.6.3.2 Returns:

An **atom**, the result of or'ing all atoms in `flatten(values)`.

8.6.6.3.3 Comments:

This function may be applied to an atom or to all elements of a sequence. It performs `or_bits()` operations repeatedly.

8.6.6.3.4 Example 1:

```
a = sum({10, 7, 35})  
-- a is 47
```

8.6.6.3.5 See Also:

[can_add](#), [sum](#), [product](#), [or_bits](#)

8.6.7 Bitwise operations

8.6.7.1 and_bits

```
<built-in> function and_bits(object a, object b)
```

Perform the logical AND operation on corresponding bits in two objects. A bit in the result will be 1 only if the corresponding bits in both arguments are 1.

8.6.7.1.1 Parameters:

1. a : one of the objects involved
2. b : the second object

8.6.7.1.2 Returns:

An **object**, whose shape depends on the shape of both arguments. Each atom in this object is obtained by logical AND between atoms on both objects.

8.6.7.1.3 Comments:

The arguments to this function may be atoms or sequences. The rules for operations on sequences apply. The atoms in the arguments must be representable as 32-bit numbers, either signed or unsigned.

If you intend to manipulate full 32-bit values, you should declare your variables as atom, rather than integer. EUPHORIA's integer type is limited to 31-bits.

Results are treated as signed numbers. They will be negative when the highest-order bit is 1.

To understand the binary representation of a number you should display it in hexadecimal notation. Use the `%x` format of `printf()`. Using `int_to_bits()` is an even more direct approach.

8.6.7.1.4 Example 1:

```
a = and_bits(#0F0F0000, #12345678)
-- a is #02040000
```

8.6.7.1.5 Example 2:

```
a = and_bits(#FF, {#123456, #876543, #2211})
-- a is {#56, #43, #11}
```

8.6.7.1.6 Example 3:

```
a = and_bits(#FFFFFFFF, #FFFFFFFF)
-- a is -1
-- Note that #FFFFFFFF is a positive number,
-- but the result of a bitwise logical operation is interpreted
-- as a signed 32-bit number, so it's negative.
```

8.6.7.1.7 See Also:

[or_bits](#), [xor_bits](#), [not_bits](#), [int_to_bits](#)

8.6.7.2 xor_bits

```
<built-in> function xor_bits(object a, object b)
```

Perform the logical XOR operation on corresponding bits in two objects. A bit in the result will be 1 only if the corresponding bits in both arguments are different.

8.6.7.2.1 Parameters:

1. `a` : one of the objects involved
2. `b` : the second object

8.6.7.2.2 Returns:

An **object**, whose shape depends on the shape of both arguments. Each atom in this object is obtained by logical XOR between atoms on both objects.

8.6.7.2.3 Comments:

The arguments must be representable as 32-bit numbers, either signed or unsigned.

If you intend to manipulate full 32-bit values, you should declare your variables as atom, rather than integer. EUPHORIA's integer type is limited to 31-bits.

Results are treated as signed numbers. They will be negative when the highest-order bit is 1.

8.6.7.2.4 Example 1:

```
a = xor_bits(#0110, #1010)
-- a is #1100
```

8.6.7.2.5 See Also:

[and_bits](#), [or_bits](#), [not_bits](#), [int_to_bits](#)

8.6.7.3 or_bits

```
<built-in> function or_bits(object a, object b)
```

Perform the logical OR operation on corresponding bits in two objects. A bit in the result will be 1 only if the corresponding bits in both arguments are both 0.

8.6.7.3.1 Parameters:

1. a : one of the objects involved
2. b : the second object

8.6.7.3.2 Returns:

An **object**, whose shape depends on the shape of both arguments. Each atom in this object is obtained by logical XOR between atoms on both objects.

8.6.7.3.3 Comments:

The arguments must be representable as 32-bit numbers, either signed or unsigned.

If you intend to manipulate full 32-bit values, you should declare your variables as atom, rather than integer. EUPHORIA's integer type is limited to 31-bits.

Results are treated as signed numbers. They will be negative when the highest-order bit is 1.

8.6.7.3.4 Example 1:

```
a = or_bits(#0F0F0000, #12345678)
-- a is #1F3F5678
```

8.6.7.3.5 Example 2:

```
a = or_bits(#FF, {#123456, #876543, #2211})
-- a is {#1234FF, #8765FF, #22FF}
```

8.6.7.3.6 See Also:

[and_bits](#), [xor_bits](#), [not_bits](#), [int_to_bits](#)

8.6.7.4 not_bits

```
<built-in> function not_bits(object a)
```

Perform the logical NOT operation on each bit in an object. A bit in the result will be 1 when the corresponding bit in x1 is 0, and will be 0 when the corresponding bit in x1 is 1.

8.6.7.4.1 Parameters:

1. a : the object to invert the bits of.

8.6.7.4.2 Returns:

An **object**, the same shape as a. Each bit in an atom of the result is the reverse of the corresponding bit inside a.

8.6.7.4.3 Comments:

The argument to this function may be an atom or a sequence.

The argument must be representable as a 32-bit number, either signed or unsigned.

If you intend to manipulate full 32-bit values, you should declare your variables as atom, rather than integer. EUPHORIA's integer type is limited to 31-bits.

Results are treated as signed numbers. They will be negative when the highest-order bit is 1.

A simple equality holds for an atom a: $a + \text{not_bits}(a) = -1$.

8.6.7.4.4 Example 1:

```
a = not_bits(#000000F7)
-- a is -248 (i.e. FFFFFFF08 interpreted as a negative number)
```

8.6.7.4.5 See Also:

[and_bits](#), [or_bits](#), [xor_bits](#), [int_to_bits](#)

8.6.7.5 shift_bits

```
include std/math.e
public function shift_bits(object source_number, integer shift_distance)
```

Moves the bits in the input value by the specified distance.

8.6.7.5.1 Parameters:

1. `source_number` : object: The value(s) whose bits will be moved.
2. `shift_distance` : integer: number of bits to be moved by.

8.6.7.5.2 Comments:

- If `source_number` is a sequence, each element is shifted.
- The value(s) in `source_number` are first truncated to a 32-bit integer.
- The output is truncated to a 32-bit integer.
- Vacated bits are replaced with zero.
- If `shift_distance` is negative, the bits in `source_number` are moved left.
- If `shift_distance` is positive, the bits in `source_number` are moved right.
- If `shift_distance` is zero, the bits in `source_number` are not moved.

8.6.7.5.3 Returns:

Atom(s) containing a 32-bit integer. A single atom in `source_number` is an atom, or a sequence in the same form as `source_number` containing 32-bit integers.

8.6.7.5.4 Example 1:

```
? shift_bits((7, -3) --> 56
? shift_bits((0, -9) --> 0
? shift_bits((4, -7) --> 512
? shift_bits((8, -4) --> 128
? shift_bits((0xFE427AAC, -7) --> 0x213D5600
? shift_bits((-7, -3) --> -56 which is 0xFFFFFC8
? shift_bits((131, 0) --> 131
```

```
? shift_bits((184.464, 0) --> 184
? shift_bits((999_999_999_999_999, 0) --> -1530494977 which is 0xA4C67FFF
? shift_bits((184, 3) -- 23
? shift_bits((48, 2) --> 12
? shift_bits((121, 3) --> 15
? shift_bits((0xFE427AAC, 7) --> 0x01FC84F5
? shift_bits((-7, 3) --> 0x1FFFFFFF
? shift_bits({48, 121}, 2) --> {12, 30}
```

8.6.7.5.5 See Also:

[rotate_bits](#)

8.6.7.6 rotate_bits

```
include std/math.e
public function rotate_bits(object source_number, integer shift_distance)
```

Rotates the bits in the input value by the specified distance.

8.6.7.6.1 Parameters:

1. `source_number` : object: value(s) whose bits will be rotated.
2. `shift_distance` : integer: number of bits to be moved by.

8.6.7.6.2 Comments:

- If `source_number` is a sequence, each element is rotated.
- The value(s) in `source_number` are first truncated to a 32-bit integer.
- The output is truncated to a 32-bit integer.
- If `shift_distance` is negative, the bits in `source_number` are rotated left.
- If `shift_distance` is positive, the bits in `source_number` are rotated right.
- If `shift_distance` is zero, the bits in `source_number` are not rotated.

8.6.7.6.3 Returns:

Atom(s) containing a 32-bit integer. A single atom in `source_number` is an atom, or a sequence in the same form as `source_number` containing 32-bit integers.

8.6.7.6.4 Example 1:

```
? rotate_bits(7, -3) --> 56
? rotate_bits(0, -9) --> 0
? rotate_bits(4, -7) --> 512
? rotate_bits(8, -4) --> 128
```

```
? rotate_bits(0xFE427AAC, -7) --> 0x213D567F
? rotate_bits(-7, -3) --> -49 which is 0xFFFFF7CF
? rotate_bits(131, 0) --> 131
? rotate_bits(184.464, 0) --> 184
? rotate_bits(999_999_999_999_999, 0) --> -1530494977 which is 0xA4C67FFF
? rotate_bits(184, 3) -- 23
? rotate_bits(48, 2) --> 12
? rotate_bits(121, 3) --> 536870927
? rotate_bits(0xFE427AAC, 7) --> 0x59FC84F5
? rotate_bits(-7, 3) --> 0x3FFFFFFF
? rotate_bits({48, 121}, 2) --> {12, 1073741854}
```

8.6.7.6.5 See Also:

[shift_bits](#)

Arithmetics

8.6.7.7 gcd

```
include std/math.e
public function gcd(atom p, atom q)
```

Returns the greater common divisor of to atoms

8.6.7.7.1 Parameters:

1. *p* : one of the atoms to consider
2. *q* : the other atom.

8.6.7.7.2 Returns:

A positive **atom**, without a fractional part, evenly dividing both parameters, and is the greatest value with those properties.

8.6.7.7.3 Comments:

Signs are ignored. Atoms are rounded down to integers.

Any zero parameter causes 0 to be returned.

Parameters and return value are atoms so as to take mathematical integers up to power (2, 53).

8.6.7.7.4 Example 1:

```
? gcd(76.3, -114) -- prints out gcd(76,114), which is 38
```

Floating Point

8.6.7.8 approx

```
include std/math.e
public function approx(object p, object q, atom epsilon = 0.005)
```

Compares two (sets of) numbers based on approximate equality.

8.6.7.8.1 Parameters:

1. *p* : an object, one of the sets to consider
2. *q* : an object, the other set.
3. *epsilon* : an atom used to define the amount of inequality allowed. This must be a positive value.
Default is 0.005

8.6.7.8.2 Returns:

An **integer**,

- 1 when $p > (q + \text{epsilon})$: *P* is definitely greater than *q*.
- -1 when $p < (q - \text{epsilon})$: *P* is definitely less than *q*.
- 0 when $p \geq (q - \text{epsilon})$ and $p \leq (q + \text{epsilon})$: *p* and *q* are approximately equal.

8.6.7.8.3 Comments:

This can be used to see if two numbers are near enough to each other.

Also, because of the way floating point numbers are stored, it not always possible express every real number exactly, especially after a series of arithmetic operations. You can use `approx()` to see if two floating point numbers are almost the same value.

If *p* and *q* are both sequences, they must be the same length as each other.

If *p* or *q* is a sequence, but the other is not, then the result is a sequence of results whose length is the same as the sequence argument.

8.6.7.8.4 Example 1:

```
? approx(10, 33.33 * 30.01 / 100) --> 0 because 10 and 10.002333 are within 0.005 of each other
? approx(10, 10.001) -> 0 because 10 and 10.001 are within 0.005 of each other
? approx(10, {10.001, 9.999, 9.98, 10.04}) --> {0,0,1,-1}
? approx({10.001, 9.999, 9.98, 10.04}, 10) --> {0,0,-1,1}
? approx({10.001, {9.999, 10.01}, 9.98, 10.04}, {10.01, 9.99, 9.8, 10.4}) --> {-1, {1,1}, 1, -1}
? approx(23, 32, 10) -> 0 because 23 and 32 are within 10 of each other.
```

8.6.7.9 powof2

```
include std/math.e
public function powof2(object p)
```

Tests for power of 2

8.6.7.9.1 Parameters:

1. p : an object. The item to test. This can be an integer, atom or sequence.

8.6.7.9.2 Returns:

An integer,

- 1 for each item in p that is a power of two, eg. 2,4,8,16,32, ...
- 0 for each item in p that is **not** a power of two, eg. 3, 54,322, -2

8.6.7.9.3 Example 1:

```
for i = 1 to 10 do
    ? {i, powof2(i)}
end for
-- output ...
-- {1,1}
-- {2,1}
-- {3,0}
-- {4,1}
-- {5,0}
-- {6,0}
-- {7,0}
-- {8,1}
-- {9,0}
-- {10,0}
```

8.6.7.10 is_even

```
include std/math.e
public function is_even(integer test_integer)
```

Test if the supplied integer is a even or odd number.

8.6.7.10.1 Parameters:

1. `test_integer` : an integer. The item to test.

8.6.7.10.2 Returns:

An **integer**,

- 1 if its even.
- 0 if its odd.

8.6.7.10.3 Example 1:

```
for i = 1 to 10 do
  ? {i, is_even(i)}
end for
-- output ...
-- {1,0}
-- {2,1}
-- {3,0}
-- {4,1}
-- {5,0}
-- {6,1}
-- {7,0}
-- {8,1}
-- {9,0}
-- {10,1}
```

8.6.7.11 is_even_obj

```
include std/math.e
public function is_even_obj(object test_object)
```

Test if the supplied EUPHORIA object is even or odd.

8.6.7.11.1 Parameters:

1. `test_object` : any EUPHORIA object. The item to test.

8.6.7.11.2 Returns:

An **object**,

- If `test_object` is an integer...
 - ◆ 1 if its even.
 - ◆ 0 if its odd.
- Otherwise if `test_object` is an atom this always returns 0
- otherwise if `test_object` is a sequence it tests each element recursively, returning a sequence of the same structure containing ones and zeros for each element. A 1 means that the element at this position was even otherwise it was odd.

8.6.7.11.3 Example 1:

```
for i = 1 to 5 do
  ? {i, is_even_obj(i)}
end for
-- output ...
-- {1,0}
-- {2,1}
-- {3,0}
-- {4,1}
-- {5,0}
```

8.6.7.11.4 Example 2:

```
? is_even_obj(3.4) --> 0
```

8.6.7.11.5 Example 3:

```
? is_even_obj({{1,2,3}, {{4,5},6,{7,8}},9}) --> {{0,1,0},{1,0},1,{0,1}},0}
```

8.7 Math Constants

Constants

PI
QUARTPI
HALFPI
TWOPI
PISQR
INVSQ2PI
PHI
E
LN2
INVLN2
LN10
INVLN10

SQRT2
HALFSQRT2
SQRT3
DEGREES_TO_RADIANS
RADIANS_TO_DEGREES
EULER_GAMMA
SQRTE
PINF
MINF

8.7.1 Constants

8.7.1.1 PI

```
include std/mathcons.e  
public constant PI
```

PI is the ratio of a circle's circumference to it's diameter.

$PI = C / D :: C = PI * D :: C = PI * 2 * R(\text{radius})$

8.7.1.2 QUARTPI

```
include std/mathcons.e  
public constant QUARTPI
```

Quarter of PI

8.7.1.3 HALFPI

```
include std/mathcons.e  
public constant HALFPI
```

Half of PI

8.7.1.4 TWOPI

```
include std/mathcons.e  
public constant TWOPI
```

Two times PI

8.7.1.5 PISQR

```
include std/mathcons.e
public constant PISQR
```

π^2

8.7.1.6 INVSQ2PI

```
include std/mathcons.e
public constant INVSQ2PI
```

$1 / (\sqrt{2\pi})$

8.7.1.7 PHI

```
include std/mathcons.e
public constant PHI
```

$\phi \Rightarrow \text{Golden Ratio} = (1 + \sqrt{5}) / 2$

8.7.1.8 E

```
include std/mathcons.e
public constant E
```

Euler (e) The base of the natural logarithm.

8.7.1.9 LN2

```
include std/mathcons.e
public constant LN2
```

$\ln(2) :: 2 = \text{power}(E, \text{LN2})$

8.7.1.10 INVLN2

```
include std/mathcons.e
public constant INVLN2
```

$1 / (\ln(2))$

8.7.1.11 LN10

```
include std/mathcons.e
public constant LN10
```

$\ln(10) :: 10 = \text{power}(E, \text{LN10})$

8.7.1.12 INVLN10

```
include std/mathcons.e
public constant INVLN10
```

$1 / \ln(10)$

8.7.1.13 SQRT2

```
include std/mathcons.e
public constant SQRT2
```

$\sqrt{2}$

8.7.1.14 HALFSQRT2

```
include std/mathcons.e
public constant HALFSQRT2
```

$\sqrt{2} / 2$

8.7.1.15 SQRT3

```
include std/mathcons.e
public constant SQRT3
```

Square root of 3

8.7.1.16 DEGREES_TO_RADIANS

```
include std/mathcons.e
public constant DEGREES_TO_RADIANS
```

Conversion factor: Degrees to Radians = $\text{PI} / 180$

8.7.1.17 RADIANS_TO_DEGREES

```
include std/mathcons.e
public constant RADIANS_TO_DEGREES
```

Conversion factor: Radians to Degrees = 180 / PI

8.7.1.18 EULER_GAMMA

```
include std/mathcons.e
public constant EULER_GAMMA
```

Gamma (Euler Gamma)

8.7.1.19 SQ RTE

```
include std/mathcons.e
public constant SQ RTE
```

sqrt(e)

8.7.1.20 PINF

```
include std/mathcons.e
public constant PINF
```

Positive Infinity

8.7.1.21 MINF

```
include std/mathcons.e
public constant MINF
```

Negative Infinity

8.8 Random Numbers

```
rand
rand_range
rnd
rnd_1
```

[set_rand](#)
[chance](#)
[roll](#)
[sample](#)

8.8.1 rand

<built-in> [function rand\(object maximum\)](#)

Return a random positive integer.

8.8.1.1 Parameters:

1. `maximum`: an atom, a cap on the value to return.

8.8.1.1.1 Returns:

An **integer**, from 1 to `maximum`.

8.8.1.1.2 Errors:

If [ceil](#)(`maximum`) is not a positive integer ≤ 1073741823 , an error will occur. It must also be at least 1.

8.8.1.1.3 Comments:

This function may be applied to an atom or to all elements of a sequence. In order to get reproducible results from this function, you should call [set_rand\(\)](#) with a reproducible value prior.

8.8.1.1.4 Example 1:

```
s = rand({10, 20, 30})  
-- s might be: {5, 17, 23} or {9, 3, 12} etc.
```

8.8.1.1.5 See Also:

[set_rand](#), [ceil](#)

8.8.1.2 rand_range

```
include std/rand.e
public function rand_range(integer lo, integer hi)
```

Return a random integer from a specified inclusive integer range.

8.8.1.2.1 Parameters:

1. `lo` : an integer, the lower bound of the range
2. `hi` : an integer, the upper bound of the range.

8.8.1.2.2 Returns:

An **integer**, randomly drawn between `lo` and `hi` inclusive.

8.8.1.2.3 Errors:

If `lo` is not less than `hi`, an error will occur.

8.8.1.2.4 Comments:

This function may be applied to an atom or to all elements of a sequence. In order to get reproducible results from this function, you should call `set_rand()` with a reproducible value prior.

8.8.1.2.5 Example 1:

```
s = rand_range(18, 24)
-- s could be any of: 18, 19, 20, 21, 22, 23 or 24
```

8.8.1.2.6 See Also:

[rand](#), [set_rand](#), [rnd](#)

8.8.1.3 rnd

```
include std/rand.e
public function rnd()
```

Return a random floating point number in the range 0 to 1.

8.8.1.3.1 Parameters:

None.

8.8.1.3.2 Returns:

An **atom**, randomly drawn between 0.0 and 1.0 inclusive.

8.8.1.3.3 Comments:

In order to get reproducible results from this function, you should call `set_rand()` with a reproducible value prior to calling this.

8.8.1.3.4 Example 1:

```
set_rand(1001)
s = rnd()
-- s is 0.2634879318
```

8.8.1.3.5 See Also:

[rand](#), [set_rand](#), [rand_range](#)

8.8.1.4 rnd_1

```
include std/rand.e
public function rnd_1()
```

Return a random floating point number in the range 0 to less than 1.

8.8.1.4.1 Parameters:

None.

8.8.1.4.2 Returns:

An **atom**, randomly drawn between 0.0 and a number less than 1.0

8.8.1.4.3 Comments:

In order to get reproducible results from this function, you should call `set_rand()` with a reproducible value prior to calling this.

8.8.1.4.4 Example 1:

```
set_rand(1001)
s = rnd_1()
-- s is 0.2634879318
```

8.8.1.4.5 See Also:

[rand](#), [set_rand](#), [rand_range](#)

8.8.1.5 set_rand

```
include std/rand.e
public procedure set_rand(object seed)
```

Reset the random number generator.

8.8.1.5.1 Parameters:

1. `seed` : an object. The generator uses this initialize itself for the next random number generated. This can be a single integer or atom, or a sequence of two integers, or an empty sequence or any other sort of sequence.

8.8.1.5.2 Comments:

- Starting from a `seed`, the values returned by `rand()` are reproducible. This is useful for demos and stress tests based on random data. Normally the numbers returned by the `rand()` function are totally unpredictable, and will be different each time you run your program. Sometimes however you may wish to repeat the same series of numbers, perhaps because you are trying to debug your program, or maybe you want the ability to generate the same output (e.g. a random picture) for your user upon request.
- Internally there are actually two seed values.
 - ◆ When `set_rand()` is called with a single integer or atom, the two internal seeds are derived from the parameter.
 - ◆ When `set_rand()` is called with a sequence of exactly two integers/atoms the internal seeds are set to the parameter values.
 - ◆ When `set_rand()` is called with an empty sequence, the internal seeds are set to random values and are unpredictable. This is how to reset the generator.
 - ◆ When `set_rand()` is called with any other sequence, the internal seeds are set based on the length of the sequence and the hashed value of the sequence.
- Aside from an empty `seed` parameter, this sets the generator to a known state and the random numbers generated after come in a predicable order, though they still appear to be random.

8.8.1.5.3 Example 1:

```
sequence s, t
s = repeat(0, 3)
t = s

set_rand(12345)
s[1] = rand(10)
s[2] = rand(100)
s[3] = rand(1000)

set_rand(12345)  -- same value for set_rand()
t[1] = rand(10)  -- same arguments to rand() as before
t[2] = rand(100)
t[3] = rand(1000)
-- at this point s and t will be identical
set_rand("") -- Reset the generator to an unknown seed.
t[1] = rand(10)  -- Could be anything now, no way to predict it.
```

8.8.1.5.4 See Also:

[rand](#)

8.8.1.6 chance

```
include std/rand.e
public function chance(atom my_limit, atom top_limit = 100)
```

Simulates the probability of a desired outcome.

8.8.1.6.1 Parameters:

1. `my_limit`: an atom. The desired chance of something happening.
2. `top_limit`: an atom. The maximum chance of something happening. The default is 100.

8.8.1.6.2 Returns:

an integer. 1 if the desired chance happened otherwise 0.

8.8.1.6.3 Comments:

This simulates the chance of something happening. For example, if you want something to happen with a probability of 25 times out of 100 times then you code `chance(25)` and if you want something to (most likely) occur 345 times out of 999 times, you code `chance(345, 999)` #.

8.8.1.6.4 Example 1:

```
-- 65% of the days are sunny, so ...
if chance(65) then
    puts(1, "Today will be a sunny day")
elseif chance(40) then
    -- And 40% of non-sunny days it will rain.
    puts(1, "It will rain today")
else
    puts(1, "Today will be a overcast day")
end if
```

8.8.1.6.5 See Also:

[rnd](#), [roll](#)

8.8.1.7 roll

```
include std/rand.e
public function roll(object desired, integer sides = 6)
```

Simulates the probability of a dice throw.

8.8.1.7.1 Parameters:

1. `desired`: an object. One or more desired outcomes.
2. `sides`: an integer. The number of sides on the dice. Default is 6.

8.8.1.7.2 Returns:

an integer. 0 if none of the desired outcomes occurred, otherwise the face number that was rolled.

8.8.1.7.3 Comments:

The minimum number of sides is 2 and there is no maximum.

8.8.1.7.4 Example 1:

```
res = roll(1, 2) -- Simulate a coin toss.
res = roll({1,6}) -- Try for a 1 or a 6 from a standard die toss.
res = roll({1,2,3,4}, 20) -- Looking for any number under 5 from a 20-sided die.
```

8.8.1.7.5 See Also:

[rnd](#), [chance](#)

8.8.1.8 sample

```
include std/rand.e
public function sample(sequence full_set, integer sample_size, integer return_remaining = 0)
```

Selects a random sample sub-set of items from a population set.

8.8.1.8.1 Parameters:

1. `full_set` : a sequence. The set of items from which to take a sample.
2. `sample_size`: an integer. The number of samples to take.
3. `return_remaining`: an integer. If non-zero, the sub-set not selected is also returned. If zero, the default, only the sampled set is returned.

8.8.1.8.2 Returns:

a sequence. When `return_remaining = 0` then this is the set of samples, otherwise it returns a two-element sequence; the first is the samples, and the second is the remainder of the population (in the original order).

8.8.1.8.3 Comments:

- If `sample_size` is less than 1, an empty set is returned.
- If `sample_size` is greater than or equal to the population count, the entire population set is returned, but in a random order.

8.8.1.8.4 Example 1:

```
set_rand("example")
printf(1, "%s\n", { sample("abcdefghijklmnopqrstuvwxyz", 1) }) --> "t"
printf(1, "%s\n", { sample("abcdefghijklmnopqrstuvwxyz", 5) }) --> "flukq"
printf(1, "%s\n", { sample("abcdefghijklmnopqrstuvwxyz", -1) }) --> ""
printf(1, "%s\n", { sample("abcdefghijklmnopqrstuvwxyz", 26) }) --> "kghrsxmjoeubaywlzftcpivqnd"
printf(1, "%s\n", { sample("abcdefghijklmnopqrstuvwxyz", 25) }) --> "omntrqsbjguaikzywvxflpedc"
```

8.8.1.8.5 Example 2:

```
-- Deal 4 hands of 5 cards from a standard deck of cards.
sequence theDeck
sequence hands = {}
sequence rt
```

```
function new_deck()  
    sequence  
    i = 1 to 4 do  
        j = 1 to 13 do  
            append(nd, {i,j})    nd =  
            end for  
        end for  
    end return  
end function  
theDeck = new_deck()  
for i = 1 to 4 do  
    rt = theDeck, 5, 1)  
    theDeck = rt  
    append(hands, rt[1])  
end for
```

8.9 Mouse

Requirements

Constants

MOVE
LEFT_DOWN
LEFT_UP
RIGHT_DOWN
RIGHT_UP
MIDDLE_DOWN
MIDDLE_UP
ANY_UP

Routines

get_mouse
mouse_events
mouse_pointer

8.9.1 Requirements

- *Linux* -- you need GPM server to be running
- *Windows* -- not implemented yet for the text console
- *FreeBSD* -- not implemented
- *OS X* -- not implemented

8.9.2 Constants

The following constants can be used to identify and specify mouse events.

8.9.2.1 MOVE

```
include std/mouse.e
public integer MOVE
```

8.9.2.2 LEFT_DOWN

```
include std/mouse.e
public integer LEFT_DOWN
```

8.9.2.3 LEFT_UP

```
include std/mouse.e
public integer LEFT_UP
```

8.9.2.4 RIGHT_DOWN

```
include std/mouse.e
public integer RIGHT_DOWN
```

8.9.2.5 RIGHT_UP

```
include std/mouse.e
public integer RIGHT_UP
```

8.9.2.6 MIDDLE_DOWN

```
include std/mouse.e
public integer MIDDLE_DOWN
```

8.9.2.7 MIDDLE_UP

```
include std/mouse.e
public integer MIDDLE_UP
```

8.9.2.8 ANY_UP

```
include std/mouse.e
public integer ANY_UP
```

8.9.3 Routines

8.9.3.1 get_mouse

```
include std/mouse.e
public function get_mouse()
```

Queries the last mouse event.

8.9.3.1.1 Returns:

An **object**, either -1 if there has not been a mouse event since the last time `get_mouse()` was called. Otherwise, returns a triple {event, x, y}.

Constants have been defined in `mouse.e` for the possible mouse events (the values for `event`):

```
public constant
    MOVE = 1,
    LEFT_DOWN = 2,
    LEFT_UP = 4,
    RIGHT_DOWN = 8,
    RIGHT_UP = 16,
    MIDDLE_DOWN = 32,
    MIDDLE_UP = 64
```

x and y are the coordinates of the mouse pointer at the time that the event occurred.

8.9.3.1.2 Comments:

`get_mouse()` returns immediately with either a -1 or a mouse event, without waiting for an event to occur. So, you must check it frequently enough to avoid missing an event: when the next event occurs, the current event will be lost, if you haven't read it. In practice it is not hard to catch almost all events. Losing a MOVE event is generally not too serious, as the next MOVE will tell you where the mouse pointer is.

Sometimes multiple events will be reported. For example, if the mouse is moving when the left button is clicked, `get_mouse()` will report an event value of `LEFT_DOWN+MOVE`, i.e. 2+1 or 3. For this reason you should test for a particular event using [and_bits\(\)](#). See examples below. Further, you can determine which events will be reported using [mouse_events](#).

In *Linux*, no scaling is required - x and y correspond to the line and column on the screen, with (1,1) at the top left.

In *Linux*, mouse movement events are not reported in an xterm window, only in the text console.

In *Linux*, LEFT_UP, RIGHT_UP and MIDDLE_UP are not distinguishable from one another.

The first call that you make to `get_mouse()` will turn on a mouse pointer, or a highlighted character.

The x,y coordinate returned could be that of the very tip of the mouse pointer or might refer to the pixel pointed-to by the mouse pointer.

8.9.3.1.3 Example 1:

8.9.3.1.4 a return value of:

{2, 100, 50} would indicate that the left button was pressed down when the mouse pointer was at location x=100, y=50 on the screen.

8.9.3.1.5 Example 2:

To test for LEFT_DOWN, write something like the following:

```
while 1 do
  object event = get_mouse()
  if sequence(event) then
    if and_bits(event[1], LEFT_DOWN) then
      -- left button was pressed
      exit
    end if
  end if
end while
```

8.9.3.1.6 See Also:

[mouse_events](#), [mouse_pointer](#)

8.9.3.2 mouse_events

```
include std/mouse.e
public procedure mouse_events(integer events)
```

Select the mouse events [get_mouse\(\)](#) is to report.

8.9.3.2.1 Parameters:

1. `events`: an integer, all requested event codes or'ed together.

8.9.3.2.2 Comments:

By default, `get_mouse()` will report all events. `mouse_events()` can be called at various stages of the execution of your program, as the need to detect events changes. Under *Unix*, `mouse_events()` currently has no effect.

It is good practice to ignore events that you are not interested in, particularly the very frequent MOVE event, in order to reduce the chance that you will miss a significant event.

The first call that you make to `mouse_events()` will turn on a mouse pointer, or a highlighted character.

8.9.3.2.3 Example 1:

```
mouse_events(LEFT_DOWN + LEFT_UP + RIGHT_DOWN)
```

will restrict `get_mouse()` to reporting the left button being pressed down or released, and the right button being pressed down. All other events will be ignored.

8.9.3.2.4 See Also:

[get_mouse](#), [mouse_pointer](#)

8.9.3.3 mouse_pointer

```
include std/mouse.e  
public procedure mouse_pointer(integer show_it)
```

Turn mouse pointer on or off.

8.9.3.3.1 Parameters:

1. `show_it`: an integer, 0 to hide and 1 to show.

8.9.3.3.2 Comments:

Multiple calls to hide the pointer will require multiple calls to turn it back on. The first call to either `get_mouse()` or `mouse_events()` will also turn the pointer on (once).

Under *Linux*, `mouse_pointer()` currently has no effect

It may be necessary to hide the mouse pointer temporarily when you update the screen.

After a call to [text_rows\(\)](#) you may have to call [mouse_pointer\(1\)](#) to see the mouse pointer again.

8.9.3.3.3 See Also:

[get_mouse](#), [mouse_pointer](#)

8.10 Operating System Helpers

[CMD_SWITCHES](#)

[Operating System Constants](#)

[WIN32](#)

[LINUX](#)

[OSX](#)

[SUNOS](#)

[OPENBSD](#)

[NETBSD](#)

[FREEBSD](#)

[Environment.](#)

[instance](#)

[get_pid](#)

[uname](#)

[is_win_nt](#)

[getenv](#)

[setenv](#)

[unsetenv](#)

[platform](#)

[Interacting with the OS](#)

[system](#)

[system_exec](#)

[Miscellaneous](#)

[sleep](#)

[include_paths](#)

[Pipe Input/Output](#)

[Notes](#)

[Accessor Constants](#)

[STDIN](#)

[STDOUT](#)

[STDERR](#)

[PID](#)

[PARENT](#)

[CHILD](#)

[Opening/Closing](#)

[process](#)

[close](#)

kill
Read/Write Process
read
write
error_no
create
exec

8.10.1 CMD_SWITCHES

```
include std/os.e  
public constant CMD_SWITCHES
```

8.10.2 Operating System Constants

8.10.2.1 WIN32

```
include std/os.e  
public enum WIN32
```

8.10.2.2 LINUX

```
include std/os.e  
public enum LINUX
```

8.10.2.3 OSX

```
include std/os.e  
public enum OSX
```

8.10.2.4 SUNOS

```
include std/os.e  
public enum SUNOS
```

8.10.2.5 OPENBSD

```
include std/os.e
public enum OPENBSD
```

8.10.2.6 NETBSD

```
include std/os.e
public enum NETBSD
```

8.10.2.7 FREEBSD

```
include std/os.e
public enum FREEBSD
```

These constants are returned by the [platform](#) function.

- WIN32 -- Host operating system is Windows
- LINUX -- Host operating system is Linux
- FREEBSD -- Host operating system is FreeBSD
- OSX -- Host operating system is Mac OS X
- SUNOS -- Host operating system is Sun's OpenSolaris
- OPENBSD -- Host operating system is OpenBSD
- NETBSD -- Host operating system is NetBSD

8.10.2.7.1 Note:

Via the [platform](#) call, there is no way to determine if you are on Linux or FreeBSD. This was done to provide a generic UNIX return value for [platform](#).

In most situations you are better off to test the host platform by using the [ifdef statement](#). It is both more precise and faster.

8.10.3 Environment.

8.10.3.1 instance

```
include std/os.e
public function instance()
```

Return `hInstance` on *Windows* and `Process ID (pid)` on *Unix*.

8.10.3.1.1 Comments:

On *Windows* the `hInstance` can be passed around to various *Windows* routines.

8.10.3.2 get_pid

```
include std/os.e
public function get_pid()
```

Return the ID of the current Process (`pid`)

8.10.3.2.1 Returns:

An atom: The current process' id.

8.10.3.2.2 Example:

```
mypid = get_pid()
```

8.10.3.3 uname

```
include std/os.e
public function uname()
```

Retrieves the name of the host OS.

8.10.3.3.1 Returns:

A **sequence**, starting with the OS name. If identification fails, returns an OS name of UNKNOWN. Extra information depends on the OS.

On Unix, returns the same information as the `uname()` syscall in the same order as the struct `utsname`. This information is: OS Name/Kernel Name Local Hostname Kernel Version/Kernel Release Kernel Specific Version information (This is usually the date that the kernel was compiled on and the name of the host that performed the compiling.) Architecture Name (Usually a string of `i386` vs `x86_64` vs `ARM` vs etc)

On Windows, returns the following in order: Windows Platform (out of WinCE, Win9x, WinNT, Win32s, or Unknown Windows) Name of Windows OS (Windows 3.1, Win95, WinXP, etc) Platform Number Build Number Minor OS version number Major OS version number

On UNKNOWN, returns an OS name of "UNKNOWN". No other information is returned.

Returns a string of "" if an internal error has occurred.

8.10.3.3.2 Comments:

On Unix, M_UNAME is defined as a machine_func() and this is passed to the C backend. If the M_UNAME call fails, the raw machine_func() returns -1. On non Unix platforms, calling the machine_func() directly returns 0.

8.10.3.4 is_win_nt

```
include std/os.e
public function is_win_nt()
```

Decides whether the host system is a newer Windows version (NT/2K/XP/Vista).

8.10.3.4.1 Returns:

An **integer**, 1 if host system is a newer Windows (NT/2K/XP/Vista), else 0.

8.10.3.5 getenv

```
<built-in> function getenv(sequence var_name)
```

Return the value of an environment variable.

8.10.3.5.1 Parameters:

1. `var_name` : a string, the name of the variable being queried.

8.10.3.5.2 Returns:

An **object**, -1 if the variable does not exist, else a sequence holding its value.

8.10.3.5.3 Comments:

Both the argument and the return value, may, or may not be, case sensitive. You might need to test this on your own system.

8.10.3.5.4 Example:

```
e = getenv("EUDIR")
-- e will be "C:\EUPHORIA" -- or perhaps D:, E: etc.
```

8.10.3.5.5 See Also:

[setenv](#), [command_line](#)

8.10.3.6 setenv

```
include std/os.e
public function setenv(sequence name, sequence val, integer overwrite = 1)
```

Set an environment variable.

8.10.3.6.1 Parameters:

1. `name` : a string, the environment variable name
2. `val` : a string, the value to set to
3. `overwrite` : an integer, nonzero to overwrite an existing variable, 0 to disallow this.

8.10.3.6.2 Example 1:

```
? setenv("NAME", "John Doe")
? setenv("NAME", "Jane Doe")
? setenv("NAME", "Jim Doe", 0)
```

8.10.3.6.3 See Also:

[getenv](#), [unsetenv](#)

8.10.3.7 unsetenv

```
include std/os.e
public function unsetenv(sequence env)
```

Unset an environment variable

8.10.3.7.1 Parameters:

1. `name` : name of environment variable to unset

8.10.3.7.2 Example 1:

```
? unsetenv("NAME")
```

8.10.3.7.3 See Also:

[setenv](#), [getenv](#)

8.10.3.8 platform

<built-in> `function platform()`

Indicates the platform that the program is being executed on.

8.10.3.8.1 Returns:

An **integer**,

```
public constant
    WIN32,
    LINUX,
    FREEBSD,
    OSX,
    SUNOS,
    OPENBSD,
    NETBSD,
    FREEBSD
```

8.10.3.8.2 Comments:

The [ifdef statement](#) is much more versatile and in most cases supersedes `platform()`.

`platform()` used to be the way to execute different code depending on which platform the program is running on. Additional platforms will be added as EUPHORIA is ported to new machines and operating environments.

8.10.3.8.3 Example 1:

```
ifdef WIN32 then
    -- call system Beep routine
    err = c_func(Beep, {0,0})
elseif
    -- do nothing (Linux/FreeBSD)
end if
```

8.10.3.8.4 See Also:

[Platform-Specific Issues](#), [ifdef statement](#)

8.10.4 Interacting with the OS

8.10.4.1 system

<built-in> `procedure system(sequence command, integer mode=0)`

Pass a command string to the operating system command interpreter.

8.10.4.1.1 Parameters:

1. `command` : a string to be passed to the shell
2. `mode` : an integer, indicating the manner in which to return from the call.

8.10.4.1.2 Errors:

`command` should not exceed 1,024 characters.

8.10.4.1.3 Comments:

Allowable values for `mode` are:

- 0: the previous graphics mode is restored and the screen is cleared.
- 1: a beep sound will be made and the program will wait for the user to press a key before the previous graphics mode is restored.
- 2: the graphics mode is not restored and the screen is not cleared.

`mode = 2` should only be used when it is known that the command executed by `system()` will not change the graphics mode.

You can use EUPHORIA as a sophisticated "batch" (.bat) language by making calls to `system()` and `system_exec()`.

`system()` will start a new command shell.

`system()` allows you to use command-line redirection of standard input and output in `command`.

8.10.4.1.4 Example 1:

```
system("copy temp.txt a:\\temp.bak", 2)
-- note use of double backslash in literal string to get
-- single backslash
```

8.10.4.1.5 Example 2:

```
system("eui \\test\\myprog.ex < indata > outdata", 2)
-- executes myprog by redirecting standard input and
-- standard output
```

8.10.4.1.6 See Also:

[system_exec](#), [command_line](#), [current_dir](#), [getenv](#)

8.10.4.2 system_exec

<built-in> `function system_exec(sequence command, integer mode=0)`

Try to run the a shell executable command

8.10.4.2.1 Parameters:

1. `command` : a string to be passed to the shell, representing an executable command
2. `mode` : an integer, indicating the manner in which to return from the call.

8.10.4.2.2 Returns:

An **integer**, basically the exit/return code from the called process.

8.10.4.2.3 Errors:

`command` should not exceed 1,024 characters.

8.10.4.2.4 Comments:

Allowable values for `mode` are:

- 0 -- the previous graphics mode is restored and the screen is cleared.
- 1 -- a beep sound will be made and the program will wait for the user to press a key before the previous graphics mode is restored.
- 2 -- the graphics mode is not restored and the screen is not cleared.

If it is not possible to run the program, `system_exec()` will return -1.

On *WIN32*, `system_exec()` will only run `.exe` and `.com` programs. To run `.bat` files, or built-in shell commands, you need [system\(\)](#). Some commands, such as `DEL`, are not programs, they are actually built-in to the command interpreter.

On WIN32, `system_exec()` does not allow the use of command-line redirection in `command`. Nor does it allow you to quote strings that contain blanks, such as file names.

exit codes from Windows programs are normally in the range 0 to 255, with 0 indicating "success".

You can run a EUPHORIA program using `system_exec()`. A EUPHORIA program can return an exit code using `abort()`.

`system_exec()` does not start a new command shell.

8.10.4.2.5 Example 1:

```
integer exit_code
exit_code = system_exec("xcopy templ.dat temp2.dat", 2)

if exit_code = -1 then
    puts(2, "\n couldn't run xcopy.exe\n")
elseif exit_code = 0 then
    puts(2, "\n xcopy succeeded\n")
else
    printf(2, "\n xcopy failed with code %d\n", exit_code)
end if
```

8.10.4.2.6 Example 2:

```
-- executes myprog with two file names as arguments
if system_exec("eui \\test\\myprog.ex indata outdata", 2) then
    puts(2, "failure!\n")
end if
```

8.10.4.2.7 See Also:

[system](#), [abort](#)

8.10.5 Miscellaneous

8.10.5.1 sleep

```
include std/os.e
public procedure sleep(atom t)
```

Suspend thread execution. for `t` seconds.

8.10.5.1.1 Parameters:

1. `t` : an atom, the number of seconds for which to sleep.

8.10.5.1.2 Comments:

The operating system will suspend your process and schedule other processes.

With multiple tasks, the whole program sleeps, not just the current task. To make just the current task sleep, you can call `task_schedule(task_self(), {i, i})` and then execute `task_yield()`. Another option is to call `task_delay()`.

8.10.5.1.3 Example:

```
puts(1, "Waiting 15 seconds and a quarter...\n")
sleep(15.25)
puts(1, "Done.\n")
```

8.10.5.1.4 See Also:

[task_schedule](#), [task_yield](#), [task_delay](#)

8.10.5.2 include_paths

```
<built-in> function include_paths(integer convert)
```

Returns the list of include paths, in the order in which they are searched

8.10.5.2.1 Parameters:

1. `convert` : an integer, nonzero to include converted path entries that were not validated yet.

8.10.5.2.2 Returns:

A **sequence**, of strings, each holding a fully qualified include path.

8.10.5.2.3 Comments:

`convert` is checked only under *Windows*. If a path has accented characters in it, then it may or may not be valid to convert those to the OEM code page. Setting `convert` to a nonzero value will force conversion for path entries that have accents and which have not been checked to be valid yet. The extra entries, if any, are returned at the end of the returned sequence.

8.10.5.2.4 The paths are ordered in the order they are searched:

1. current directory
2. configuration file,
3. command line switches,
4. EUINC
5. a default based on EUDIR.

8.10.5.2.5 Example 1:

```
sequence s = include_paths(0)
-- s might contain
{
  "/usr/euphoria/tests",
  "/usr/euphoria/include",
  "./include",
  "../include"
}
```

8.10.5.2.6 See Also:

[eu.cfg](#), [include](#), [option_switches](#)

8.10.6 Pipe Input/Output

8.10.6.1 Notes

Due to a bug, EUPHORIA does not handle STDERR properly STDERR cannot be captured for EUPHORIA programs (other programs will work fully) The IO functions currently work with file handles, a future version might wrap them in streams so that they can be used directly alongside other file/socket/other-streams with a `stream_select()` function.

8.10.7 Accessor Constants

8.10.7.1 STDIN

```
include std/pipeio.e
public enum STDIN
```

Child processes standard input

8.10.7.2 STDOUT

```
include std/pipeio.e  
public enum STDOUT
```

Child processes standard output

8.10.7.3 STDERR

```
include std/pipeio.e  
public enum STDERR
```

Child processes standard error

8.10.7.4 PID

```
include std/pipeio.e  
public enum PID
```

Process ID

8.10.7.5 PARENT

```
include std/pipeio.e  
public enum PARENT
```

Set of pipes that are for the use of the parent

8.10.7.6 CHILD

```
include std/pipeio.e  
public enum CHILD
```

Set of pipes that are given to the child - should not be used by the parent

8.10.8 Opening/Closing

8.10.8.1 process

```
include std/pipeio.e
public type process(object o)
```

Process Type

8.10.8.2 close

```
include std/pipeio.e
public function close(atom fd)
```

Close handle fd

8.10.8.2.1 Returns:

An **integer**, 0 on success, -1 on failure

8.10.8.2.2 Example 1:

```
integer status = pipeio:close(p[STDIN])
```

8.10.8.3 kill

```
include std/pipeio.e
public procedure kill(process p, atom signal = 15)
```

Close pipes and kill process p with signal signal (default 15)

8.10.8.3.1 Comments:

Signal is ignored on Windows.

8.10.8.3.2 Example 1:

```
kill(p)
```

8.10.9 Read/Write Process

8.10.9.1 read

```
include std/pipeio.e
public function read(atom fd, integer bytes)
```

Read `bytes` bytes from handle `fd`

8.10.9.1.1 Returns:

A **sequence**, containing data, an empty sequence on EOF or an error code. Similar to [get_bytes](#).

8.10.9.1.2 Example 1:

```
sequence data=read(p[STDOUT], 256)
```

Write `bytes` to handle `fd`

8.10.9.1.3 Returns:

A **integer**, number of bytes written, or -1 on error

8.10.9.1.4 Example 1:

```
integer bytes_written = write(p[STDIN], "Hello World!")
```

8.10.9.2 write

```
include std/pipeio.e
public function write(atom fd, sequence str)
```

8.10.9.3 error_no

```
include std/pipeio.e
public function error_no()
```

Get error no from last call to a pipe function

8.10.9.3.1 Comments:

Value returned will be OS-specific, and is not always set on Windows at least

8.10.9.3.2 Example 1:

```
integer error = error_no()
```

8.10.9.4 create

```
include std/pipeio.e
public function create()
```

Create pipes for inter-process communication

8.10.9.4.1 Returns:

A **handle**, process handles { {parent side pipes},{child side pipes} }

8.10.9.4.2 Example 1:

```
object p = exec("dir", create())
```

8.10.9.5 exec

```
include std/pipeio.e
public function exec(sequence cmd, sequence pipe)
```

Open process with command line cmd

8.10.9.5.1 Returns:

A **handle**, process handles { [PID](#), [STDIN](#), [STDOUT](#), [STDERR](#) }

8.10.9.5.2 Example 1:

```
object p = exec("dir", create())
```

8.11 Pretty Printing

Page Contents

[PRETTY_DEFAULT](#)
[DISPLAY_ASCII](#)
[INDENT](#)



START_COLUMN
WRAP
INT_FORMAT
FP_FORMAT
MIN_ASCII
MAX_ASCII
MAX_LINES
LINE_BREAKS

Routines

pretty_print
pretty_sprint

8.11.1 PRETTY_DEFAULT

```
include std/pretty.e  
public constant PRETTY_DEFAULT
```

8.11.1.1 DISPLAY_ASCII

```
include std/pretty.e  
public enum DISPLAY_ASCII
```

8.11.1.2 INDENT

```
include std/pretty.e  
public enum INDENT
```

8.11.1.3 START_COLUMN

```
include std/pretty.e  
public enum START_COLUMN
```

8.11.1.4 WRAP

```
include std/pretty.e  
public enum WRAP
```



8.11.1.5 INT_FORMAT

```
include std/pretty.e
public enum INT_FORMAT
```

8.11.1.6 FP_FORMAT

```
include std/pretty.e
public enum FP_FORMAT
```

8.11.1.7 MIN_ASCII

```
include std/pretty.e
public enum MIN_ASCII
```

8.11.1.8 MAX_ASCII

```
include std/pretty.e
public enum MAX_ASCII
```

8.11.1.9 MAX_LINES

```
include std/pretty.e
public enum MAX_LINES
```

8.11.1.10 LINE_BREAKS

```
include std/pretty.e
public enum LINE_BREAKS
```

8.11.2 Routines

8.11.2.1 pretty_print

```
include std/pretty.e
public procedure pretty_print(integer fn, object x, sequence options = PRETTY_DEFAULT)
```

Print an object to a file or device, using braces { , , , }, indentation, and multiple lines to show the structure.

8.11.2.1.1 Parameters:

1. `fn` : an integer, the file/device number to write to
2. `x` : the object to display/convert to printable form
3. `options` : is an (up to) 10-element options sequence.

8.11.2.1.2 Comments:

Pass `{}` in `options` to select the defaults, or set options as below:

1. display ASCII characters:
 - ◆ 0 -- never
 - ◆ 1 -- alongside any integers in printable ASCII range (default)
 - ◆ 2 -- display as "string" when all integers of a sequence are in ASCII range
 - ◆ 3 -- show strings, and quoted characters (only) for any integers in ASCII range as well as the characters: `\t \r \n`
2. amount to indent for each level of sequence nesting -- default: 2
3. column we are starting at -- default: 1
4. approximate column to wrap at -- default: 78
5. format to use for integers -- default: `"%d"`
6. format to use for floating-point numbers -- default: `"%.10g"`
7. minimum value for printable ASCII -- default 32
8. maximum value for printable ASCII -- default 127
9. maximum number of lines to output
10. line breaks between elements -- default 1 (0 = no line breaks, -1 = line breaks to wrap only)

If the length is less than 10, unspecified options at the end of the sequence will keep the default values. e.g. `{0, 5}` will choose "never display ASCII", plus 5-character indentation, with defaults for everything else.

The default options can be applied using the public constant `PRETTY_DEFAULT`, and the elements may be accessed using the following public enum:

1. `DISPLAY_ASCII`
2. `INDENT`
3. `START_COLUMN`
4. `WRAP`
5. `INT_FORMAT`
6. `FP_FORMAT`
7. `MIN_ASCII`
8. `MAX_ASCII`
9. `MAX_LINES`
10. `LINE_BREAKS`

The display will start at the current cursor position. Normally you will want to call `pretty_print()` when the cursor is in column 1 (after printing a `<code>\n</code>` character). If you want to start in a different column, you should call `position()` and specify a value for option [3]. This will ensure that the first and last braces in a sequence line up vertically.

When specifying the format to use for integers and floating-point numbers, you can add some decoration, e.g. "%d)" or "\$ %.2f"

8.11.2.1.3 Example 1:

```
pretty_print(1, "ABC", {})  
  
{65'A', 66'B', 67'C'}
```

8.11.2.1.4 Example 2:

```
pretty_print(1, {{1,2,3}, {4,5,6}}, {})  
  
{  
  {1,2,3},  
  {4,5,6}  
}
```

8.11.2.1.5 Example 3:

```
pretty_print(1, {"EUPHORIA", "Programming", "Language"}, {2})  
  
{  
  "EUPHORIA",  
  "Programming",  
  "Language"  
}
```

8.11.2.1.6 Example 4:

```
puts(1, "word_list = ") -- moves cursor to column 13  
pretty_print(1,  
  {"EUPHORIA", 8, 5.3},  
  {"Programming", 11, -2.9},  
  {"Language", 8, 9.8}},  
  {2, 4, 13, 78, "%03d", "%.3f"}) -- first 6 of 8 options  
  
word_list = {  
  {  
    "EUPHORIA",  
    008,  
    5.300  
  },  
  {  
    "Programming",  
    011,  
    -2.900  
  },  
  {  
    "Language",  
    008,  
    9.800  
  }  
}
```

```
}  
}
```

8.11.2.1.7 See Also:

[print](#), [sprint](#), [printf](#), [sprintf](#), [pretty_sprint](#)

8.11.2.2 pretty_sprint

```
include std/pretty.e  
public function pretty_sprint(object x, sequence options = PRETTY_DEFAULT)
```

Format an object using braces { , , , }, indentation, and multiple lines to show the structure.

8.11.2.2.1 Parameters:

1. `x` : the object to display
2. `options` : is an (up to) 10-element options sequence: Pass { } to select the defaults, or set options

8.11.2.2.2 Returns:

A **sequence**, of printable characters, representing `x` in an human-readable form.

8.11.2.2.3 Comments:

This function formats objects the same as [pretty_print\(\)](#), but returns the sequence obtained instead of sending it to some file..

8.11.2.2.4 See Also:

[pretty_print](#), [sprint](#)

8.12 Statistics

Page Contents

[Routines](#)
 [small](#)
 [largest](#)
 [smallest](#)
 [range](#)
 [ST_FULLPOP](#)

ST_SAMPLE
ST_ALLNUM
ST_IGNSTR
ST_ZEROSTR
stdev
avedev
sum
count
average
geomean
harmean
movavg
emovavg
median
raw_frequency
mode
central_moment
sum_central_moments
skewness
kurtosis

8.12.1 Routines

8.12.1.1 small

```
include std/stats.e  
public function small(sequence data_set, integer ordinal_idx)
```

Determines the k-th smallest value from the supplied set of numbers.

8.12.1.1.1 Parameters:

1. `data_set` : The list of values from which the smallest value is chosen.
2. `ordinal_idx` : The relative index of the desired smallest value.

8.12.1.1.2 Returns:

A **sequence**, {The k-th smallest value, its index in the set}

8.12.1.1.3 Comments:

`small()` is used to return a value based on its size relative to all the other elements in the sequence. When `index` is 1, the smallest index is returned. Use `index = length(data_set)` to return the highest.

If `ordinal_idx` is less than one, or greater then length of `data_set`, an empty sequence is returned.

The set of values does not have to be in any particular order. The values may be any EUPHORIA object.

8.12.1.1.4 Example 1:

```
? small( {4,5,6,8,5,4,3,"text"}, 3 ) -- Ans: {4,1} (The 3rd smallest value)
? small( {4,5,6,8,5,4,3,"text"}, 1 ) -- Ans: {3,7} (The 1st smallest value)
? small( {4,5,6,8,5,4,3,"text"}, 7 ) -- Ans: {8,4} (The 7th smallest value)
? small( {"def", "qwe", "abc", "try"}, 2 ) -- Ans: {"def", 1} (The 2nd smallest value)
? small( {1,2,3,4}, -1) -- Ans: {} -- no-value
? small( {1,2,3,4}, 10) -- Ans: {} -- no-value
```

8.12.1.2 largest

```
include std/stats.e
public function largest(object data_set)
```

Returns the largest of the data points that are atoms.

8.12.1.2.1 Parameters:

1. `data_set` : a list of 1 or more numbers among which you want the largest.

8.12.1.2.2 Returns:

An **object**, either of:

- an atom (the largest value) if there is at least one atom item in the set
- {} if there *is* no largest value.

8.12.1.2.3 Comments:

Any `data_set` element which is not an atom is ignored.

8.12.1.2.4 Example 1:

```
? largest( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,"text"} ) -- Ans: 8
? largest( {"just","text"} ) -- Ans: {}
```


8.12.1.2.5 See also:

[range](#)

8.12.1.3 smallest

```
include std/stats.e
public function smallest(object data_set)
```

Returns the smallest of the data points.

8.12.1.3.1 Parameters:

1. `data_set` : A list of 1 or more numbers for which you want the smallest. **Note:** only atom elements are included and any sub-sequences elements are ignored.

8.12.1.3.2 Returns:

An **object**, either of:

- an atom (the smallest value) if there is at least one atom item in the set
- `{}` if there *is* no largest value.

8.12.1.3.3 Comments:

Any `data_set` element which is not an atom is ignored.

8.12.1.3.4 Example 1:

```
? smallest( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,"text"} ) -- Ans: 1
? smallest( {"just","text"} ) -- Ans: {}
```

8.12.1.3.5 See also:

[range](#)

8.12.1.4 range

```
include std/stats.e
public function range(object data_set)
```

Determines a number of *range* statistics for the data set.

8.12.1.4.1 Parameters:

1. `data_set` : a list of 1 or more numbers for which you want the range data.

8.12.1.4.2 Returns:

A **sequence**, empty if no atoms were found, else like {Lowest, Highest, Range, Mid-range}

8.12.1.4.3 Comments:

Any sequence element in `data_set` is ignored.

8.12.1.4.4 Example 1:

```
? range( {7,2,8,5,6,6,4,8,6,16,3,3,4,1,8,"text"} ) -- Ans: {1, 16, 15, 8.5}
```

8.12.1.4.5 See also:

[smallest](#) [largest](#)

Enums used to influence the results of some of these functions.

8.12.1.5 ST_FULLPOP

```
include std/stats.e
public enum ST_FULLPOP
```

The supplied data is the entire population.

8.12.1.6 ST_SAMPLE

```
include std/stats.e
public enum ST_SAMPLE
```

The supplied data is only a random sample of the population.

8.12.1.7 ST_ALLNUM

```
include std/stats.e
public enum ST_ALLNUM
```

The supplied data consists of only atoms.

8.12.1.8 ST_IGNSTR

```
include std/stats.e
public enum ST_IGNSTR
```

Any sub-sequences (eg. strings) in the supplied data are ignored.

8.12.1.9 ST_ZEROSTR

```
include std/stats.e
public enum ST_ZEROSTR
```

Any sub-sequences (eg. strings) in the supplied data are assumed to have the value zero.

8.12.1.10 stdev

```
include std/stats.e
public function stdev(sequence data_set, object subseq_opt = ST_ALLNUM, integer population_type)
```

Returns the standard deviation based of the population.

8.12.1.10.1 Parameters:

1. `data_set` : a list of 1 or more numbers for which you want the estimated standard deviation.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.
3. `population_type` : an integer. `ST_SAMPLE` (the default) assumes that `data_set` is a random sample of the total population. `ST_FULLPOP` means that `data_set` is the entire population.

8.12.1.10.2 Returns:

An **atom**, the estimated standard deviation. An empty **sequence** means that there is no meaningful data to calculate from.

8.12.1.10.3 Comments:

`stdev()` is a measure of how values are different from the average.

The numbers in `data_set` can either be the entire population of values or just a random subset. You indicate which in the `population_type` parameter. By default `data_set` represents a sample and not the entire population. When using this function with sample data, the result is an *estimated* standard deviation.

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.10.4 The equation for standard deviation is:

```
stdev(X) ==> SQRT(SUM(SQ(X{1..N} - MEAN)) / (N))
```

8.12.1.10.5 Example 1:

```
? stdev( {4,5,6,7,5,4,3,7} )           -- Ans: 1.457737974
? stdev( {4,5,6,7,5,4,3,7} , , ST_FULLPOP ) -- Ans: 1.363589014
? stdev( {4,5,6,7,5,4,3,"text"} , ST_IGNSTR ) -- Ans: 1.345185418
? stdev( {4,5,6,7,5,4,3,"text"} , ST_IGNSTR, ST_FULLPOP ) -- Ans: 1.245399698
? stdev( {4,5,6,7,5,4,3,"text"} , 0 ) -- Ans: 2.121320344
? stdev( {4,5,6,7,5,4,3,"text"} , 0, ST_FULLPOP ) -- Ans: 1.984313483
```

8.12.1.10.6 See also:

[average](#), [avedev](#)

8.12.1.11 adev

```
include std/stats.e
public function adev(sequence data_set, object subseq_opt = ST_ALLNUM, integer population_type)
```

Returns the average of the absolute deviations of data points from their mean.

8.12.1.11.1 Parameters:

1. `data_set` : a list of 1 or more numbers for which you want the mean of the absolute deviations.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.
3. `population_type` : an integer. `ST_SAMPLE` (the default) assumes that `data_set` is a random sample of the total population. `ST_FULLPOP` means that `data_set` is the entire population.

8.12.1.11.2 Returns:

An **atom** , the deviation from the mean.

An empty **sequence**, means that there is no meaningful data to calculate from.

8.12.1.11.3 Comments:

avedev() is a measure of the variability in a data set. Its statistical properties are less well behaved than those of the standard deviation, which is why it is used less.

The numbers in `data_set` can either be the entire population of values or just a random subset. You indicate which in the `population_type` parameter. By default `data_set` represents a sample and not the entire population. When using this function with sample data, the result is an *estimated* deviation.

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

The equation for absolute average deviation is:

$$\text{avedev}(X) ==> \text{SUM}(\text{ABS}(X\{1..N\} - \text{MEAN}(X))) / N$$

8.12.1.11.4 Example 1:

```
? avedev( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,7} ) -- Ans: 1.966666667
? avedev( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,7},, ST_FULLPOP ) -- Ans: 1.84375
? avedev( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,"text"}, ST_IGNSTR ) -- Ans: 1.99047619
? avedev( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,"text"}, ST_IGNSTR,ST_FULLPOP ) -- Ans: 1.857777778
? avedev( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,"text"}, 0 ) -- Ans: 2.225
? avedev( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,"text"}, 0, ST_FULLPOP ) -- Ans: 2.0859375
```

8.12.1.11.5 See also:

[average](#), [stdev](#)

8.12.1.12 sum

```
include std/stats.e
public function sum(object data_set, object subseq_opt = ST_ALLNUM)
```

Returns the sum of all the atoms in an object.

8.12.1.12.1 Parameters:

1. `data_set` : Either an atom or a list of numbers to sum.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.12.2 Returns:

An **atom**, the sum of the set.

8.12.1.12.3 Comments:

`sum()` is used as a measure of the magnitude of a sequence of positive values.

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

The equation is:

$$\text{sum}(X) \implies \text{SUM}(X\{1..N\})$$

8.12.1.12.4 Example 1:

```
? sum( {7,2,8.5,6,6,-4.8,6,6,3.341,-8,"text"}, 0 ) -- Ans: 32.041
```

8.12.1.12.5 See also:

[average](#)

8.12.1.13 count

```
include std/stats.e
public function count(object data_set, object subseq_opt = ST_ALLNUM)
```

Returns the count of all the atoms in an object.

8.12.1.13.1 Parameters:

1. `data_set` : either an atom or a list.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.13.2 Comments:

This returns the number of numbers in `data_set`

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.13.3 Returns:

An **integer**, the number of atoms in the set. When `data_set` is an atom, 1 is returned.

8.12.1.13.4 Example 1:

```
? count( {7,2,8.5,6,6,-4.8,6,6,3.341,-8,"text"} ) -- Ans: 10
? count( {"cat", "dog", "lamb", "cow", "rabbit"} ) -- Ans: 0 (no atoms)
? count( 5 ) -- Ans: 1
```

8.12.1.13.5 See also:

[average](#), [sum](#)

8.12.1.14 average

```
include std/stats.e
public function average(object data_set, object subseq_opt = ST_ALLNUM)
```

Returns the average (mean) of the data points.

8.12.1.14.1 Parameters:

1. `data_set` : A list of 1 or more numbers for which you want the mean.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.14.2 Returns:

An **object**,

- `{ }` (the empty sequence) if there are no atoms in the set.
- an atom (the mean) if there are one or more atoms in the set.

8.12.1.14.3 Comments:

`average()` is the theoretical probable value of a randomly selected item from the set.

8.12.1.14.4 The equation for average is:

$$\text{average}(X) ==> \text{SUM}(X\{1..N\}) / N$$

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.14.5 Example 1:

```
? average( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,"text"}, ST_IGNSTR ) -- Ans: 5.13333333
```

8.12.1.14.6 See also:

[geomean](#), [harmean](#), [movavg](#), [emovavg](#)

8.12.1.15 geomean

```
include std/stats.e
public function geomean(object data_set, object subseq_opt = ST_ALLNUM)
```


Returns the geometric mean of the atoms in a sequence.

8.12.1.15.1 Parameters:

1. `data_set` : the values to take the geometric mean of.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.15.2 Returns:

An **atom**, the geometric mean of the atoms in `data_set`. If there is no atom to take the mean of, 1 is returned.

8.12.1.15.3 Comments:

The geometric mean of N atoms is the N -th root of their product. Signs are ignored.

This is useful to compute average growth rates.

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.15.4 Example 1:

```
? geomean({3, "abc", -2, 6}, ST_IGNSTR) -- prints out power(36,1/3) = 3,30192724889462669
? geomean({1,2,3,4,5,6,7,8,9,10}) -- = 4.528728688
```

8.12.1.15.5 See Also:

[average](#)

8.12.1.16 harmean

```
include std/stats.e
public function harmean(sequence data_set, object subseq_opt = ST_ALLNUM)
```

Returns the harmonic mean of the atoms in a sequence.

8.12.1.16.1 Parameters:

1. `data_set` : the values to take the harmonic mean of.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.16.2 Returns:

An **atom**, the harmonic mean of the atoms in `data_set`.

8.12.1.16.3 Comments:

The harmonic mean is the inverse of the average of their inverses.

This is useful in engineering to compute equivalent capacities and resistances.

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.16.4 Example 1:

```
? harmean({3, "abc", -2, 6}, ST_IGNSTR) -- = 0.  
? harmean({{2, 3, 4}}) -- 3 / (1/2 + 1/3 + 1/4) = 2.769230769
```

8.12.1.16.5 See Also:

[average](#)

8.12.1.17 movavg

```
include std/stats.e  
public function movavg(object data_set, object period_delta)
```

Returns the average (mean) of the data points for overlapping periods. This can be either a simple or weighted moving average.

8.12.1.17.1 Parameters:

1. `data_set` : a list of 1 or more numbers for which you want a moving average.
2. `period_delta` : an object, either
 - an integer representing the size of the period, or
 - a list of weightings to apply to the respective period positions.

8.12.1.17.2 Returns:

A **sequence**, either the requested averages or `{ }` if the Data sequence is empty or the supplied period is less than one.

If a list of weights was supplied, the result is a weighted average; otherwise, it is a simple average.

8.12.1.17.3 Comments:

A moving average is used to smooth out a set of data points over a period.
For example, given a period of 5:

1. the first returned element is the average of the first five data points [1..5],
 2. the second returned element is the average of the second five data points [2..6],
- and so on
until the last returned value is the average of the last 5 data points [\$-4 .. \$].

When `period_delta` is an atom, it is rounded down to the width of the average. When it is a sequence, the width is its length. If there are not enough data points, zeroes are inserted.

Note that only atom elements are included and any sub-sequence elements are ignored.

8.12.1.17.4 Example 1:

```
? movavg( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8}, 10 )
-- Ans: {5.8, 5.4, 5.5, 5.1, 4.7, 4.9}
? movavg( {7,2,8,5,6}, 2 )
-- Ans: {4.5, 5, 6.5, 5.5}
? movavg( {7,2,8,5,6}, {0.5, 1.5} )
-- Ans: {3.25, 6.5, 5.75, 5.75}
```

8.12.1.17.5 See also:

[average](#)

8.12.1.18 emovavg

```
include std/stats.e
public function emovavg(object data_set, atom smoothing_factor)
```

Returns the exponential moving average of a set of data points.

8.12.1.18.1 Parameters:

1. `data_set` : a list of 1 or more numbers for which you want a moving average.
2. `smoothing_factor` : an atom, the smoothing factor, typically between 0 and 1.

8.12.1.18.2 Returns:

A **sequence**, made of the requested averages, or `{ }` if `data_set` is empty or the supplied period is less than one.

8.12.1.18.3 Comments:

A moving average is used to smooth out a set of data points over a period.

The formula used is:

$$Y_i = Y_{i-1} + F * (X_i - Y_{i-1})$$

Note that only atom elements are included and any sub-sequences elements are ignored.

The smoothing factor controls how data is smoothed. 0 smooths everything to 0, and 1 means no smoothing at all.

Any value for `smoothing_factor` outside the 0.0..1.0 range causes `smoothing_factor` to be set to the periodic factor $(2 / (N+1))$.

8.12.1.18.4 Example 1:

```
? emovavg( {7,2,8,5,6}, 0.75 )
-- Ans: {5.25,2.8125,6.703125,5.42578125,5.856445313}
? emovavg( {7,2,8,5,6}, 0.25 )
-- Ans: {1.75,1.8125,3.359375,3.76953125,4.327148438}
? emovavg( {7,2,8,5,6}, -1 )
-- Ans: {2.333333333,2.22222222,4.148148148,4.432098765,4.95473251}
```

8.12.1.18.5 See also:

[average](#)

8.12.1.19 median

```
include std/stats.e
public function median(object data_set, object subseq_opt = ST_ALLNUM)
```

Returns the mid point of the data points.

8.12.1.19.1 Parameters:

1. `data_set` : a list of 1 or more numbers for which you want the mean.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.19.2 Returns:

An **object**, either `{ }` if there are no items in the set, or an **atom** (the median) otherwise.

8.12.1.19.3 Comments:

`median()` is the item for which half the items are below it and half are above it.

All elements are included; any sequence elements are assumed to have the value zero.

8.12.1.19.4 The equation for average is:

```
median(X) ==> sort(X)[N/2]
```

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.19.5 Example 1:

```
? median( {7,2,8,5,6,6,4,8,6,6,3,3,4,1,8,4} ) -- Ans: 5
```

8.12.1.19.6 See also:

[average](#), [geomean](#), [harmean](#), [movavg](#), [emovavg](#)

8.12.1.20 raw_frequency

```
include std/stats.e
public function raw_frequency(object data_set, object subseq_opt = ST_ALLNUM)
```

Returns the frequency of each unique item in the data set.

8.12.1.20.1 Parameters:

1. `data_set` : a list of 1 or more numbers for which you want the frequencies.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.20.2 Returns:

A **sequence**. This will contain zero or more 2-element sub-sequences. The first element is the frequency count and the second element is the data item that was counted. The returned values are in descending order, meaning that the highest frequencies are at the beginning of the returned list.

8.12.1.20.3 Comments:

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.20.4 Example 1:

```
? raw_frequency("the cat is the hatter")
```

This returns

```
{
{5, 116},
{4, 32},
{3, 104},
{3, 101},
{2, 97},
{1, 115},
{1, 114},
{1, 105},
{1, 99}
}
```

8.12.1.21 mode

```
include std/stats.e
public function mode(object data_set, object subseq_opt = ST_ALLNUM)
```

Returns the most frequent point(s) of the data set.

8.12.1.21.1 Parameters:

1. `data_set` : a list of 1 or more numbers for which you want the mode.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.21.2 Returns:

A **sequence**. The list of modal items in the data set.

8.12.1.21.3 Comments:

It is possible for the `mode()` to return more than one item when more than one item in the set has the same highest frequency count.

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.21.4 Example 1:

```
mode( { 7, 2, 8, 5, 6, 6, 4, 8, 6, 6, 3, 3, 4, 1, 8, 4 } ) -- Ans: { 6 }
mode( { 8, 2, 8, 5, 6, 6, 4, 8, 6, 6, 3, 3, 4, 1, 8, 4 } ) -- Ans: { 8, 6 }
```

8.12.1.21.5 See also:

[average](#), [geomean](#), [harmean](#), [movavg](#), [emovavg](#)

8.12.1.22 central_moment

```
include std/stats.e
public function central_moment(object data_set, object datum, integer order_mag = 1, object sub
```

Returns the distance between a supplied value and the mean, to some supplied order of magnitude. This is used to get a measure of the *shape* of a data set.

8.12.1.22.1 Parameters:

1. `data_set` : a list of 1 or more numbers whose mean is used.
2. `datum`: either a single value or a list of values for which you require the central moments.
3. `order_mag`: An integer. This is the order of magnitude required. Usually a number from 1 to 4, but can be anything.
4. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.22.2 Returns:

An **object**. The same data type as `datum`. This is the set of calculated central moments.

8.12.1.22.3 Comments:

For each of the items in `#datum`, its central moment is calculated as ...

```
CM = power( ITEM - AVG, MAGNITUDE)
```

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.22.4 Example 1:

```
central_moment("the cat is the hatter", "the",1) --> {23.14285714, 11.14285714, 8.142857143}
central_moment("the cat is the hatter", 't',2) --> 535.5918367
central_moment("the cat is the hatter", 't',3) --> 12395.12536
```


8.12.1.22.5 See also:

[average](#)

8.12.1.23 sum_central_moments

```
include std/stats.e
public function sum_central_moments(object data_set, integer order_mag = 1, object subseq_opt =
```

Returns sum of the central moments of each item in a data set.

8.12.1.23.1 Parameters:

1. `data_set` : a list of 1 or more numbers whose mean is used.
2. `order_mag`: An integer. This is the order of magnitude required. Usually a number from 1 to 4, but can be anything.
3. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.23.2 Returns:

An **atom**. The total of the central moments calculated for each of the items in `data_set`.

8.12.1.23.3 Comments:

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.23.4 Example 1:

```
sum_central_moments("the cat is the hatter", 1) --> -8.526512829e-14
sum_central_moments("the cat is the hatter", 2) --> 19220.57143
sum_central_moments("the cat is the hatter", 3) --> -811341.551
sum_central_moments("the cat is the hatter", 4) --> 56824083.71
```

8.12.1.23.5 See also:

[central_moment](#), [average](#)

8.12.1.24 skewness

```
include std/stats.e
public function skewness(object data_set, object subseq_opt = ST_ALLNUM)
```

Returns a measure of the asymmetry of a data set. Usually the `data_set` is a probability distribution but it can be anything. This value is used to assess how suitable the data set is in representing the required analysis. It can help detect if there are too many extreme values in the data set.

8.12.1.24.1 Parameters:

1. `data_set` : a list of 1 or more numbers whose mean is used.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.24.2 Returns:

An **atom**. The skewness measure of the data set.

8.12.1.24.3 Comments:

Generally speaking, a negative return indicates that most of the values are lower than the mean, while positive values indicate that most values are greater than the mean. However this might not be the case when there are a few extreme values on one side of the mean.

The larger the magnitude of the returned value, the more the data is skewed in that direction.

A returned value of zero indicates that the mean and median values are identical and that the data is symmetrical.

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.24.4 Example 1:

```
skewness("the cat is the hatter") --> -1.296820819
skewness("thecatisthehatter")    --> 0.1029393238
```

8.12.1.24.5 See also:

[kurtosis](#)

8.12.1.25 kurtosis

```
include std/stats.e
public function kurtosis(object data_set, object subseq_opt = ST_ALLNUM)
```

Returns a measure of the spread of values in a dataset when compared to a *normal* probability curve.

8.12.1.25.1 Parameters:

1. `data_set` : a list of 1 or more numbers whose kurtosis is required.
2. `subseq_opt` : an object. When this is `ST_ALLNUM` (the default) it means that `data_set` is assumed to contain no sub-sequences otherwise this gives instructions about how to treat sub-sequences. See comments for details.

8.12.1.25.2 Returns:

An **object**. If this is an atom it is the kurtosis measure of the data set. Otherwise it is a sequence containing an error integer. The return value `{0}` indicates that an empty dataset was passed, `{1}` indicates that the standard deviation is zero (all values are the same).

8.12.1.25.3 Comments:

Generally speaking, a negative return indicates that most of the values are further from the mean, while positive values indicate that most values are nearer to the mean.

The larger the magnitude of the returned value, the more the data is 'peaked' or 'flatter' in that direction.

If the data can contain sub-sequences, such as strings, you need to let the the function know about this otherwise it assumes every value in `data_set` is an number. If that is not the case then the function will crash. So it is important that if it can possibly contain sub-sequences that you tell this function what to do with them. Your choices are to ignore them or assume they have the value zero. To ignore them, use `ST_IGNSTR` as the `subseq_opt` parameter value otherwise use `ST_ZEROSTR`. However, if you know that `data_set` only contains numbers use the default `subseq_opt` value, `ST_ALLNUM`. **Note** It is faster if the data only contains numbers.

8.12.1.25.4 Example 1:

```
kurtosis("thecatisthehatter")    --> -1.737889192
```

8.12.1.25.5 See also:

[skewness](#)

8.13 Multi-tasking

Page Contents

[General Notes](#)

[Warning](#)

[Routines](#)

[task_delay](#)

[task_clock_start](#)

[task_clock_stop](#)

[task_create](#)

[task_list](#)

[task_schedule](#)

[task_self](#)

[task_status](#)

[task_suspend](#)

[task_yield](#)

8.13.1 General Notes

For a complete overview of the task system, please see the mini-guide [Multitasking in EUPHORIA](#).

8.13.2 Warning

The task system does not yet function in a shared library. Task routine calls that are compiled into a shared library are emitted as a NOP (no operation) and will therefore have no effect.

It is planned to allow the task system to function in shared libraries in future versions of OpenEUPHORIA.

8.13.3 Routines

8.13.3.1 task_delay

```
include std/task.e
public procedure task_delay(atom delaytime)
```

Suspends a task for a short period, allowing other tasks to run in the meantime.

8.13.3.1.1 Parameters:

1. `delaytime` : an atom, the duration of the delay in seconds.

8.13.3.1.2 Comments:

This procedure is similar to [sleep\(\)](#), but allows for other tasks to run by yielding on a regular basis. Like [sleep\(\)](#), its argument needs not being an integer.

8.13.3.1.3 See Also:

[sleep](#)

8.13.3.2 task_clock_start

```
<built-in> procedure task_clock_start()
```

Restart the clock used for scheduling real-time tasks.

8.13.3.2.1 Comments:

Call this routine, some time after calling [task_clock_stop\(\)](#), when you want scheduling of real-time tasks to continue.

[task_clock_stop\(\)](#) and [task_clock_start\(\)](#) can be used to freeze the scheduling of real-time tasks.

[task_clock_start\(\)](#) causes the scheduled times of all real-time tasks to be incremented by the amount of time since [task_clock_stop\(\)](#) was called. This allows a game, simulation, or other program to continue smoothly.

Time-shared tasks are not affected.

8.13.3.2.2 Example 1:

```
-- freeze the game while the player answers the phone
task_clock_stop()
while get_key() = -1 do
```

```
end while
task_clock_start()
```

8.13.3.2.3 See Also:

[task_clock_stop](#), [task_schedule](#), [task_yield](#), [task_suspend](#), [task_delay](#)

8.13.3.3 task_clock_stop

```
<built-in> procedure task_clock_stop()
```

Stop the scheduling of real-time tasks.

8.13.3.3.1 Comments:

Call `task_clock_stop()` when you want to take time out from scheduling real-time tasks. For instance, you want to temporarily suspend a game or simulation for a period of time.

Scheduling will resume when [task_clock_start\(\)](#) is called.

Time-shared tasks can continue. The current task can also continue, unless it's a real-time task and it yields.

The [time\(\)](#) function is not affected by this.

8.13.3.3.2 See Also:

[task_clock_start](#), [task_schedule](#), [task_yield](#), [task_suspend](#), [task_delay](#)

8.13.3.4 task_create

```
<built-in> function task_create(integer rid, sequence args)
```

Create a new task, given a home procedure and the arguments passed to it.

8.13.3.4.1 Parameters:

1. `rid`: an integer, the `routine_id` of a user-defined EUPHORIA procedure.
2. `args`: a sequence, the list of arguments that will be passed to this procedure when the task starts executing.

8.13.3.4.2 Returns:

An **atom**, a task identifier, created by the system. It can be used to identify this task to the other EUPHORIA multitasking routines.

8.13.3.4.3 Errors:

There must be at most 12 parameters in `args`.

8.13.3.4.4 Comments:

`task_create()` creates a new task, but does not start it executing. You must call `task_schedule()` for this purpose.

Each task has its own set of private variables and its own call stack. Global and local variables are shared between all tasks.

If a run-time error is detected, the traceback will include information on all tasks, with the offending task listed first.

Many tasks can be created that all run the same procedure, possibly with different parameters.

A task cannot be based on a function, since there would be no way of using the function result.

Each task id is unique. `task_create()` never returns the same task id as it did before. Task id's are integer-valued atoms and can be as large as the largest integer-valued atom (15 digits).

8.13.3.4.5 Example 1:

```
mytask = task_create(routine_id("myproc"), {5, 9, "ABC"})
```

8.13.3.4.6 See Also:

`task_schedule`, `task_yield`, `task_suspend`, `task_self`

8.13.3.5 task_list

```
<built-in> function task_list()
```

Get a sequence containing the task id's for all active or suspended tasks.

8.13.3.5.1 Returns:

A **sequence**, of atoms, the list of all task that are or may be scheduled.

8.13.3.5.2 Comments:

This function lets you find out which tasks currently exist. Tasks that have terminated are not included. You can pass a task id to [task_status\(\)](#) to find out more about a particular task.

8.13.3.5.3 Example 1:

```
sequence tasks

tasks = task_list()
for i = 1 to length(tasks) do
    if task_status(tasks[i]) > 0 then
        printf(1, "task %d is active\n", tasks[i])
    end if
end for
```

8.13.3.5.4 See Also:

[task_status](#), [task_create](#), [task_schedule](#), [task_yield](#), [task_suspend](#)

8.13.3.6 task_schedule

<built-in> `procedure task_schedule(atom task_id, object schedule)`

Schedule a task to run using a scheduling parameter.

8.13.3.6.1 Parameters:

1. `task_id`: an atom, the identifier of a task that did not terminate yet.
2. `schedule`: an object, describing when and how often to run the task.

8.13.3.6.2 Comments:

`task_id` must have been returned by [task_create\(\)](#).

The task scheduler, which is built-in to the EUPHORIA run-time system, will use `schedule` as a guide when scheduling this task. It may not always be possible to achieve the desired number of consecutive runs, or the desired time frame. For instance, a task might take so long before yielding control, that another task misses its desired time window.

`schedule` is being interpreted as follows:

`schedule` is an integer:

This defines `task_id` as time shared, and tells the task scheduler how many times it should the task in one burst before it considers running other tasks. `schedule` must be greater than zero then.

Increasing this count will increase the percentage of CPU time given to the selected task, while decreasing the percentage given to other time-shared tasks. Use trial and error to find the optimal trade off. It will also increase the efficiency of the program, since each actual task switch wastes a bit of time.

`schedule` is a sequence:

In this case, it must be a pair of positive atoms, the first one not being less than the second one. This defines `task_id` as a real time task. The pair states the minimum and maximum times, in seconds, to wait before running the task. The pair also sets the time interval for subsequent runs of the task, until the next call to `task_schedule()` or `task_suspend()`.

Real-time tasks have a higher priority. Time-shared tasks are run when no real-time task is ready to execute.

A task can switch back and forth between real-time and time-shared. It all depends on the last call to `task_schedule()` for that task. The scheduler never runs a real-time task before the start of its time frame (min value in the {min, max} pair), and it tries to avoid missing the task's deadline (max value).

For precise timing, you can specify the same value for min and max. However, by specifying a range of times, you give the scheduler some flexibility. This allows it to schedule tasks more efficiently, and avoid non-productive delays. When the scheduler must delay, it calls `sleep()`, unless the required delay is very short. `sleep()` lets the operating system run other programs.

The min and max values can be fractional. If the min value is smaller than the resolution of the scheduler's clock (currently 0.01 seconds on *Windows* or *Unix*) then accurate time scheduling cannot be performed, but the scheduler will try to run the task several times in a row to approximate what is desired.

For example, if you ask for a min time of 0.002 seconds, then the scheduler will try to run your task $.01/.002 = 5$ times in a row before waiting for the clock to "click" ahead by .01. During the next 0.01 seconds it will run your task (up to) another 5 times etc. provided your task can be completed 5 times in one clock period.

At program start-up there is a single task running. Its task id is 0, and initially it's a time-shared task allowed 1 run per `task_yield()`. No other task can run until task 0 executes a `task_yield()`.

If task 0 (top-level) runs off the end of the main file, the whole program terminates, regardless of what other tasks may still be active.

If the scheduler finds that no task is active, i.e. no task will ever run again (not even task 0), it terminates the program with a 0 exit code, similar to `abort(0)`.

8.13.3.6.3 Example 1:

```
-- Task t1 will be executed up to 10 times in a row before
-- other time-shared tasks are given control. If a real-time
-- task needs control, t1 will lose control to the real-time task.
task_schedule(t1, 10)

-- Task t2 will be scheduled to run some time between 4 and 5 seconds
-- from now. Barring any rescheduling of t2, it will continue to
-- execute every 4 to 5 seconds thereafter.
task_schedule(t2, {4, 5})
```

8.13.3.6.4 See Also:

[task_create](#), [task_yield](#), [task_suspend](#)

8.13.3.7 task_self

```
<built-in> function task_self()
```

Return the task id of the current task.

8.13.3.7.1 Comments:

This value may be needed, if a task wants to schedule or suspend itself.

8.13.3.7.2 Example 1:

```
-- schedule self
task_schedule(task_self(), {5.9, 6.0})
```

8.13.3.7.3 See Also:

[task_create](#), [task_schedule](#), [task_yield](#), [task_suspend](#)

8.13.3.8 task_status

```
<built-in> function task_status(atom task_id)
```

Return the status of a task.

8.13.3.8.1 Parameters:

1. `task_id`: an atom, the id of the task being queried.

8.13.3.8.2 Returns:

An **integer**,

- -1 -- task does not exist, or terminated
- 0 -- task is suspended
- 1 -- task is active

8.13.3.8.3 Comments:

A task might want to know the status of one or more other tasks when deciding whether to proceed with some processing.

8.13.3.8.4 Example 1:

```
integer s

s = task_status(tid)
if s = 1 then
    puts(1, "ACTIVE\n")
elseif s = 0 then
    puts(1, "SUSPENDED\n")
else
    puts(1, "DOESN'T EXIST\n")
end if
```

8.13.3.8.5 See Also:

[task_list](#), [task_create](#), [task_schedule](#), [task_suspend](#)

8.13.3.9 task_suspend

<built-in> `procedure task_suspend(atom task_id)`

Suspend execution of a task.

8.13.3.9.1 Parameters:

1. `task_id`: an atom, the id of the task to suspend.

8.13.3.9.2 Comments:

A suspended task will not be executed again unless there is a call to [task_schedule\(\)](#) for the task.

`task_id` is a task id returned from [task_create\(\)](#). - Any task can suspend any other task. If a task suspends itself, the suspension will start as soon as the task calls [task_yield\(\)](#).

Suspending a task and never scheduling it again is how to kill a task. There is no `task_kill()` primitives because undead tasks were creating too much trouble and confusion. As a general fact, nothing that impacts a running task can be effective as long as the task has not yielded.

8.13.3.9.3 Example 1:

```
-- suspend task 15
task_suspend(15)

-- suspend current task
task_suspend(task_self())
```

8.13.3.9.4 See Also:

[task_create](#), [task_schedule](#), [task_self](#), [task_yield](#)

8.13.3.10 task_yield

```
<built-in> procedure task_yield()
```

Yield control to the scheduler. The scheduler can then choose another task to run, or perhaps let the current task continue running.

8.13.3.10.1 Comments:

Tasks should call `task_yield()` periodically so other tasks will have a chance to run. Only when `task_yield()` is called, is there a way for the scheduler to take back control from a task. This is what's known as cooperative multitasking.

A task can have calls to `task_yield()` in many different places in its code, and at any depth of subroutine call.

The scheduler will use the current scheduling parameter (see [task_schedule](#)), in determining when to return to the current task.

When control returns, execution will continue with the statement that follows `task_yield()`. The call-stack and all private variables will remain as they were when `task_yield()` was called. Global and local variables may have changed, due to the execution of other tasks.

Tasks should try to call `task_yield()` often enough to avoid causing real-time tasks to miss their time window, and to avoid blocking time-shared tasks for an excessive period of time. On the other hand, there is a bit of overhead in calling `task_yield()`, and this overhead is slightly larger when an actual switch to a different task takes place. A `task_yield()` where the same task continues executing takes less time.

A task should avoid calling `task_yield()` when it is in the middle of a delicate operation that requires exclusive access to some data. Otherwise a race condition could occur, where one task might interfere with an operation being carried out by another task. In some cases a task might need to mark some data as "locked" or "unlocked" in order to prevent this possibility. With cooperative multitasking, these concurrency issues are much less of a problem than with the preemptive multitasking that other languages support.

8.13.3.10.2 Example 1:

```
-- From Language war game.
-- This small task deducts life support energy from either the
-- large EUPHORIA ship or the small shuttle.
-- It seems to run "forever" in an infinite loop,
-- but it's actually a real-time task that is called
-- every 1.7 to 1.8 seconds throughout the game.
-- It deducts either 3 units or 13 units of life support energy each time.

procedure task_life()
-- independent task: subtract life support energy
  while TRUE do
    if shuttle then
      p_energy(-3)
    else
      p_energy(-13)
    end if
    task_yield()
  end while
end procedure
```

8.13.3.10.3 See Also:

[task_create](#), [task_schedule](#), [task_suspend](#)

8.14 Types - Extended

[OBJ_UNASSIGNED](#)
[OBJ_INTEGER](#)
[OBJ_ATOM](#)
[OBJ_SEQUENCE](#)
[object](#)
[integer](#)
[atom](#)
[sequence](#)
[FALSE](#)
[TRUE](#)

CS_FIRST

Support Functions

- `char_test`
- `set_default_charsets`
- `get_charsets`
- `set_charsets`

Types

- `boolean`
- `t_boolean`
- `t_alnum`
- `t_identifier`
- `t_alpha`
- `t_ascii`
- `t_cntrl`
- `t_digit`
- `t_graph`
- `t_specword`
- `t_bytearray`
- `t_lower`
- `t_print`
- `t_display`
- `t_punct`
- `t_space`
- `t_upper`
- `t_xdigit`
- `t_vowel`
- `t_consonant`
- `integer_array`
- `t_text`
- `number_array`
- `sequence_array`
- `ascii_string`
- `string`

8.14.1 OBJ_UNASSIGNED

```
include std/types.e
public constant OBJ_UNASSIGNED
```

Object not assigned

8.14.1.1 OBJ_INTEGER

```
include std/types.e
public constant OBJ_INTEGER
```

Object is integer

8.14.1.2 OBJ_ATOM

```
include std/types.e
public constant OBJ_ATOM
```

Object is atom

8.14.1.3 OBJ_SEQUENCE

```
include std/types.e
public constant OBJ_SEQUENCE
```

Object is sequence

8.14.1.4 object

```
<built-in> function object(object x)
```

Returns information about the object type of the supplied argument x.

8.14.1.4.1 Returns:

1. An integer.
 - ◆ OBJ_UNASSIGNED if x has not been assigned anything yet.
 - ◆ OBJ_INTEGER if x holds an integer value.
 - ◆ OBJ_ATOM if x holds a number that is not an integer.
 - ◆ OBJ_SEQUENCE if x holds a sequence value.

8.14.1.4.2 Example 1:

```
? object(1) --> OBJ_INTEGER
? object(1.1) --> OBJ_ATOM
? object("1") --> OBJ_SEQUENCE
object x
? object(x) --> OBJ_UNASSIGNED
```

8.14.1.4.3 See Also:

[sequence\(\)](#), [integer\(\)](#), [atom\(\)](#)

8.14.1.5 integer

```
<built-in> function integer(object x)
```

Tests the supplied argument *x* to see if it is an integer or not.

8.14.1.5.1 Returns:

1. An integer.
 - ◆ 1 if *x* is an integer.
 - ◆ 0 if *x* is not an integer.

8.14.1.5.2 Example 1:

```
? integer(1) --> 1
? integer(1.1) --> 0
? integer("1") --> 0
```

8.14.1.5.3 See Also:

[sequence\(\)](#), [object\(\)](#), [atom\(\)](#)

8.14.1.6 atom

```
<built-in> function atom(object x)
```

Tests the supplied argument *x* to see if it is an atom or not.

8.14.1.6.1 Returns:

1. An integer.
 - ◆ 1 if *x* is an atom.
 - ◆ 0 if *x* is not an atom.

8.14.1.6.2 Example 1:

```
? atom(1) --> 1
? atom(1.1) --> 1
? atom("1") --> 0
```


8.14.1.6.3 See Also:

[sequence\(\)](#), [object\(\)](#), [integer\(\)](#)

8.14.1.7 sequence

```
<built-in> function sequence( object x)
```

Tests the supplied argument *x* to see if it is a sequence or not.

8.14.1.7.1 Returns:

1. An integer.
 - ◆ 1 if *x* is a sequence.
 - ◆ 0 if *x* is not an sequence.

8.14.1.7.2 Example 1:

```
? integer(1) --> 0
? integer(1.1) --> 0
? integer("1") --> 1
```

8.14.1.7.3 See Also:

[integer\(\)](#), [object\(\)](#), [atom\(\)](#)

8.14.1.8 FALSE

```
include std/types.e
public constant FALSE
```

Boolean FALSE value

8.14.1.9 TRUE

```
include std/types.e
public constant TRUE
```

Boolean TRUE value

8.14.1.10 CS_FIRST

```
include std/types.e
public enum CS_FIRST
```

8.14.1.10.1 Predefined character sets:

8.14.2 Support Functions

8.14.2.1 char_test

```
include std/types.e
public function char_test(object test_data, sequence char_set)
```

Determine whether one or more characters are in a given character set.

8.14.2.1.1 Parameters:

1. `test_data` : an object to test, either a character or a string
2. `char_set` : a sequence, either a list of allowable characters, or a list of pairs representing allowable ranges.

8.14.2.1.2 Returns:

An **integer**, 1 if all characters are allowed, else 0.

8.14.2.1.3 Comments:

`pCharset` is either a simple sequence of characters eg. "qwertyuiop[]" or a sequence of character pairs, which represent allowable ranges of characters. eg. Alphabetic is defined as `{{'a','z'}, {'A','Z'}}`

To add an isolated character to a character set which is defined using ranges, present it as a range of length 1, like in `{%, %}`.

8.14.2.1.4 Example 1:

```
char_test("ABCD", {{'A', 'D'}})
-- TRUE, every char is in the range 'A' to 'D'

char_test("ABCD", {{'A', 'C'}})
-- FALSE, not every char is in the range 'A' to 'C'
```

```
char_test("Harry", {{'a', 'z'}, {'D', 'J'}})
-- TRUE, every char is either in the range 'a' to 'z', or in the range 'D' to 'J'

char_test("Potter", "novel")
-- FALSE, not every character is in the set 'n', 'o', 'v', 'e', 'l'
```

8.14.2.2 set_default_charsets

```
include std/types.e
public procedure set_default_charsets()
```

Sets all the defined character sets to their default definitions.

8.14.2.2.1 Example 1:

```
set_default_charsets()
```

8.14.2.3 get_charsets

```
include std/types.e
public function get_charsets()
```

Gets the definition for each of the defined character sets.

8.14.2.3.1 Returns:

A **sequence**, of pairs. The first element of each pair is the character set id , eg. CS_Whitespace, and the second is the definition of that character set.

8.14.2.3.2 Comments:

This is the same format required for the [set_charsets\(\)](#) routine.

8.14.2.3.3 Example 1:

```
sequence sets
sets = get_charsets()
```

8.14.2.3.4 See Also:

[set_charsets](#), [set_default_charsets](#)

8.14.2.4 set_charsets

```
include std/types.e
public procedure set_charsets(sequence charset_list)
```

Sets the definition for one or more defined character sets.

8.14.2.4.1 Parameters:

1. `charset_list` : a sequence of zero or more character set definitions.

8.14.2.4.2 Comments:

`charset_list` must be a sequence of pairs. The first element of each pair is the character set id , eg. `CS_Whitespace`, and the second is the definition of that character set.

This is the same format returned by the [get_charsets\(\)](#) routine.

You cannot create new character sets using this routine.

8.14.2.4.3 Example 1:

```
set_charsets({{CS_Whitespace, " \t"}})
t_space('\n') --> FALSE

t_specword('$') --> FALSE
set_charsets({{CS_SpecWord, "_-#$$"}})
t_specword('$') --> TRUE
```

8.14.2.4.4 See Also:

[get_charsets](#)

8.14.3 Types

8.14.3.1 boolean

```
include std/types.e
public type boolean(object test_data)
```

Returns TRUE if argument is 1 or 0

Returns FALSE if the argument is anything else other than 1 or 0.

8.14.3.1.1 Example 1:

```
boolean(-1)           -- FALSE
boolean(0)            -- TRUE
boolean(1)            -- TRUE
boolean(1.234)        -- FALSE
boolean('A')         -- FALSE
boolean('9')         -- FALSE
boolean('?')         -- FALSE
boolean("abc")        -- FALSE
boolean("ab3")        -- FALSE
boolean({1, 2, "abc"}) -- FALSE
boolean({1, 2, 9.7})  -- FALSE
boolean({})           -- FALSE (empty sequence)
```

8.14.3.2 t_boolean

```
include std/types.e
public type t_boolean(object test_data)
```

Returns TRUE if argument is boolean (1 or 0) or if every element of the argument is boolean.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-boolean elements

8.14.3.2.1 Example 1:

```
t_boolean(-1)           -- FALSE
t_boolean(0)            -- TRUE
t_boolean(1)            -- TRUE
t_boolean({1, 1, 0})    -- TRUE
t_boolean({1, 1, 9.7})  -- FALSE
t_boolean({})           -- FALSE (empty sequence)
```

8.14.3.3 t_alnum

```
include std/types.e
public type t_alnum(object test_data)
```

Returns TRUE if argument is an alphanumeric character or if every element of the argument is an alphanumeric character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-alphanumeric elements

8.14.3.3.1 Example 1:

```
t_alnum(-1)           -- FALSE
t_alnum(0)            -- FALSE
t_alnum(1)            -- FALSE
t_alnum(1.234)        -- FALSE
t_alnum('A')          -- TRUE
t_alnum('9')          -- TRUE
t_alnum('?')          -- FALSE
t_alnum("abc")        -- TRUE (every element is alphabetic or a digit)
t_alnum("ab3")        -- TRUE
t_alnum({1, 2, "abc"}) -- FALSE (contains a sequence)
t_alnum({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_alnum({})           -- FALSE (empty sequence)
```

8.14.3.4 t_identifier

```
include std/types.e
public type t_identifier(object test_data)
```

Returns TRUE if argument is an alphanumeric character or if every element of the argument is an alphanumeric character and that the first character is not numeric and the whole group of characters are not all numeric.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-alphanumeric elements

8.14.3.4.1 Example 1:

```
t_identifier(-1)           -- FALSE
t_identifier(0)            -- FALSE
t_identifier(1)            -- FALSE
t_identifier(1.234)        -- FALSE
t_identifier('A')          -- TRUE
t_identifier('9')          -- FALSE
t_identifier('?')          -- FALSE
t_identifier("abc")        -- TRUE (every element is alphabetic or a digit)
t_identifier("ab3")        -- TRUE
t_identifier("ab_3")       -- TRUE (underscore is allowed)
t_identifier("1abc")       -- FALSE (identifier cannot start with a number)
t_identifier("102")        -- FALSE (identifier cannot be all numeric)
t_identifier({1, 2, "abc"}) -- FALSE (contains a sequence)
t_identifier({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_identifier({})           -- FALSE (empty sequence)
```

8.14.3.5 t_alpha

```
include std/types.e
public type t_alpha(object test_data)
```

Returns TRUE if argument is an alphabetic character or if every element of the argument is an alphabetic character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-alphabetic elements

8.14.3.5.1 Example 1:

```
t_alpha(-1)           -- FALSE
t_alpha(0)            -- FALSE
t_alpha(1)            -- FALSE
t_alpha(1.234)         -- FALSE
t_alpha('A')          -- TRUE
t_alpha('9')          -- FALSE
t_alpha('?')          -- FALSE
t_alpha("abc")         -- TRUE (every element is alphabetic)
t_alpha("ab3")         -- FALSE
t_alpha({1, 2, "abc"}) -- FALSE (contains a sequence)
t_alpha({1, 2, 9.7})   -- FALSE (contains a non-integer)
t_alpha({})           -- FALSE (empty sequence)
```

8.14.3.6 t_ascii

```
include std/types.e
public type t_ascii(object test_data)
```

Returns TRUE if argument is an ASCII character or if every element of the argument is an ASCII character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-ASCII elements

8.14.3.6.1 Example 1:

```
t_ascii(-1)           -- FALSE
t_ascii(0)            -- TRUE
t_ascii(1)            -- TRUE
t_ascii(1.234)         -- FALSE
t_ascii('A')          -- TRUE
t_ascii('9')          -- TRUE
t_ascii('?')          -- TRUE
t_ascii("abc")         -- TRUE (every element is ascii)
t_ascii("ab3")         -- TRUE
t_ascii({1, 2, "abc"}) -- FALSE (contains a sequence)
t_ascii({1, 2, 9.7})   -- FALSE (contains a non-integer)
t_ascii({})           -- FALSE (empty sequence)
```

8.14.3.7 t_cntrl

```
include std/types.e
public type t_cntrl(object test_data)
```

Returns TRUE if argument is an Control character or if every element of the argument is an Control character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-Control elements

8.14.3.7.1 Example 1:

```
t_cntrl(-1)           -- FALSE
t_cntrl(0)            -- TRUE
t_cntrl(1)            -- TRUE
t_cntrl(1.234)        -- FALSE
t_cntrl('A')          -- FALSE
t_cntrl('9')          -- FALSE
t_cntrl('?')          -- FALSE
t_cntrl("abc")        -- FALSE (every element is ascii)
t_cntrl("ab3")        -- FALSE
t_cntrl({1, 2, "abc"}) -- FALSE (contains a sequence)
t_cntrl({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_cntrl({1, 2, 'a'})  -- FALSE (contains a non-control)
t_cntrl({})           -- FALSE (empty sequence)
```

8.14.3.8 t_digit

```
include std/types.e
public type t_digit(object test_data)
```

Returns TRUE if argument is an digit character or if every element of the argument is an digit character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-digits

8.14.3.8.1 Example 1:

```
t_digit(-1)           -- FALSE
t_digit(0)            -- FALSE
t_digit(1)            -- FALSE
t_digit(1.234)        -- FALSE
t_digit('A')          -- FALSE
t_digit('9')          -- TRUE
t_digit('?')          -- FALSE
t_digit("abc")        -- FALSE
t_digit("ab3")        -- FALSE
t_digit("123")        -- TRUE
t_digit({1, 2, "abc"}) -- FALSE (contains a sequence)
t_digit({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_digit({1, 2, 'a'})  -- FALSE (contains a non-digit)
```



```
t_digit({})          -- FALSE (empty sequence)
```

8.14.3.9 t_graph

```
include std/types.e
public type t_graph(object test_data)
```

Returns TRUE if argument is a glyph character or if every element of the argument is a glyph character. (One that is visible when displayed)

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-glyph

8.14.3.9.1 Example 1:

```
t_graph(-1)          -- FALSE
t_graph(0)           -- FALSE
t_graph(1)           -- FALSE
t_graph(1.234)        -- FALSE
t_graph('A')         -- TRUE
t_graph('9')         -- TRUE
t_graph('?')         -- TRUE
t_graph(' ')         -- FALSE
t_graph("abc")       -- TRUE
t_graph("ab3")       -- TRUE
t_graph("123")       -- TRUE
t_graph({1, 2, "abc"}) -- FALSE (contains a sequence)
t_graph({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_graph({1, 2, 'a'})  -- FALSE (control chars (1,2) don't have glyphs)
t_graph({})          -- FALSE (empty sequence)
```

8.14.3.10 t_specword

```
include std/types.e
public type t_specword(object test_data)
```

Returns TRUE if argument is a special word character or if every element of the argument is a special word character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-special-word characters.

8.14.3.10.1 Comments:

A *special word character* is any character that is not normally part of a word but in certain cases may be considered. This is most commonly used when looking for words in programming source code which allows an underscore as a word character.

8.14.3.10.2 Example 1:

```
t_specword(-1)           -- FALSE
t_specword(0)            -- FALSE
t_specword(1)            -- FALSE
t_specword(1.234)        -- FALSE
t_specword('A')          -- FALSE
t_specword('9')          -- FALSE
t_specword('?')          -- FALSE
t_specword('_')          -- TRUE
t_specword("abc")        -- FALSE
t_specword("ab3")        -- FALSE
t_specword("123")        -- FALSE
t_specword({1, 2, "abc"}) -- FALSE (contains a sequence)
t_specword({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_specword({1, 2, 'a'})  -- FALSE (control chars (1,2) don't have glyphs)
t_specword({})           -- FALSE (empty sequence)
```

8.14.3.11 t_bytearray

```
include std/types.e
public type t_bytearray(object test_data)
```

Returns TRUE if argument is a byte or if every element of the argument is a byte. (Integers from 0 to 255)

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-byte

8.14.3.11.1 Example 1:

```
t_bytearray(-1)           -- FALSE (contains value less than zero)
t_bytearray(0)            -- TRUE
t_bytearray(1)            -- TRUE
t_bytearray(10)           -- TRUE
t_bytearray(100)          -- TRUE
t_bytearray(1000)         -- FALSE (greater than 255)
t_bytearray(1.234)        -- FALSE (contains a floating number)
t_bytearray('A')          -- TRUE
t_bytearray('9')          -- TRUE
t_bytearray('?')          -- TRUE
t_bytearray(' ')          -- TRUE
t_bytearray("abc")        -- TRUE
t_bytearray("ab3")        -- TRUE
t_bytearray("123")        -- TRUE
t_bytearray({1, 2, "abc"}) -- FALSE (contains a sequence)
t_bytearray({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_bytearray({1, 2, 'a'})  -- TRUE
t_bytearray({})           -- FALSE (empty sequence)
```

8.14.3.12 t_lower

```
include std/types.e
public type t_lower(object test_data)
```

Returns TRUE if argument is a lowercase character or if every element of the argument is an lowercase character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-lowercase

8.14.3.12.1 Example 1:

```
t_lower(-1)           -- FALSE
t_lower(0)            -- FALSE
t_lower(1)            -- FALSE
t_lower(1.234)         -- FALSE
t_lower('A')          -- FALSE
t_lower('9')          -- FALSE
t_lower('?')          -- FALSE
t_lower("abc")        -- TRUE
t_lower("ab3")        -- FALSE
t_lower("123")        -- TRUE
t_lower({1, 2, "abc"}) -- FALSE (contains a sequence)
t_lower({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_lower({1, 2, 'a'})  -- FALSE (contains a non-digit)
t_lower({})           -- FALSE (empty sequence)
```

8.14.3.13 t_print

```
include std/types.e
public type t_print(object test_data)
```

Returns TRUE if argument is a character that has an ASCII glyph or if every element of the argument is a character that has an ASCII glyph.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains characters that do not have an ASCII glyph.

8.14.3.13.1 Example 1:

```
t_print(-1)           -- FALSE
t_print(0)            -- FALSE
t_print(1)            -- FALSE
t_print(1.234)         -- FALSE
t_print('A')          -- TRUE
t_print('9')          -- TRUE
t_print('?')          -- TRUE
t_print("abc")        -- TRUE
t_print("ab3")        -- TRUE
t_print("123")        -- TRUE
```

```
t_print("123 ")      -- FALSE (contains a space)
t_print("123\n")     -- FALSE (contains a new-line)
t_print({1, 2, "abc"}) -- FALSE (contains a sequence)
t_print({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_print({1, 2, 'a'})  -- FALSE
t_print({})           -- FALSE (empty sequence)
```

8.14.3.14 t_display

```
include std/types.e
public type t_display(object test_data)
```

Returns TRUE if argument is a character that can be displayed or if every element of the argument is a character that can be displayed.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains characters that cannot be displayed.

8.14.3.14.1 Example 1:

```
t_display(-1)      -- FALSE
t_display(0)       -- FALSE
t_display(1)       -- FALSE
t_display(1.234)   -- FALSE
t_display('A')     -- TRUE
t_display('9')     -- TRUE
t_display('?')     -- TRUE
t_display("abc")   -- TRUE
t_display("ab3")   -- TRUE
t_display("123")   -- TRUE
t_display("123 ")  -- TRUE
t_display("123\n") -- TRUE
t_display({1, 2, "abc"}) -- FALSE (contains a sequence)
t_display({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_display({1, 2, 'a'})  -- FALSE
t_display({})        -- FALSE (empty sequence)
```

8.14.3.15 t_punct

```
include std/types.e
public type t_punct(object test_data)
```

Returns TRUE if argument is a punctuation character or if every element of the argument is a punctuation character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-punctuation symbols.

8.14.3.15.1 Example 1:

```
t_punct(-1)      -- FALSE
t_punct(0)       -- FALSE
t_punct(1)       -- FALSE
t_punct(1.234)   -- FALSE
t_punct('A')     -- FALSE
t_punct('9')     -- FALSE
t_punct('?')     -- TRUE
t_punct("abc")   -- FALSE
t_punct("(-)")   -- TRUE
t_punct("123")   -- TRUE
t_punct({1, 2, "abc"}) -- FALSE (contains a sequence)
t_punct({1, 2, 9.7}) -- FALSE (contains a non-integer)
t_punct({1, 2, 'a'}) -- FALSE (contains a non-digit)
t_punct({})      -- FALSE (empty sequence)
```

8.14.3.16 t_space

```
include std/types.e
public type t_space(object test_data)
```

Returns TRUE if argument is a whitespace character or if every element of the argument is an whitespace character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-whitespace character.

8.14.3.16.1 Example 1:

```
t_space(-1)      -- FALSE
t_space(0)       -- FALSE
t_space(1)       -- FALSE
t_space(1.234)   -- FALSE
t_space('A')     -- FALSE
t_space('9')     -- FALSE
t_space('\t')    -- TRUE
t_space("abc")   -- FALSE
t_space("123")   -- FALSE
t_space({1, 2, "abc"}) -- FALSE (contains a sequence)
t_space({1, 2, 9.7}) -- FALSE (contains a non-integer)
t_space({1, 2, 'a'}) -- FALSE (contains a non-digit)
t_space({})      -- FALSE (empty sequence)
```

8.14.3.17 t_upper

```
include std/types.e
public type t_upper(object test_data)
```

Returns TRUE if argument is an uppercase character or if every element of the argument is an uppercase character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-uppercase characters.

8.14.3.17.1 Example 1:

```
t_upper(-1)           -- FALSE
t_upper(0)            -- FALSE
t_upper(1)            -- FALSE
t_upper(1.234)         -- FALSE
t_upper('A')          -- TRUE
t_upper('9')          -- FALSE
t_upper('?')          -- FALSE
t_upper("abc")         -- FALSE
t_upper("ABC")         -- TRUE
t_upper("123")         -- FALSE
t_upper({1, 2, "abc"}) -- FALSE (contains a sequence)
t_upper({1, 2, 9.7})   -- FALSE (contains a non-integer)
t_upper({1, 2, 'a'})   -- FALSE (contains a non-digit)
t_upper({})           -- FALSE (empty sequence)
```

8.14.3.18 t_xdigit

```
include std/types.e
public type t_xdigit(object test_data)
```

Returns TRUE if argument is a hexadecimal digit character or if every element of the argument is a hexadecimal digit character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-hexadecimal character.

8.14.3.18.1 Example 1:

```
t_xdigit(-1)           -- FALSE
t_xdigit(0)            -- FALSE
t_xdigit(1)            -- FALSE
t_xdigit(1.234)         -- FALSE
t_xdigit('A')          -- TRUE
t_xdigit('9')          -- TRUE
t_xdigit('?')          -- FALSE
t_xdigit("abc")         -- TRUE
t_xdigit("fgh")         -- FALSE
t_xdigit("123")         -- TRUE
t_xdigit({1, 2, "abc"}) -- FALSE (contains a sequence)
t_xdigit({1, 2, 9.7})   -- FALSE (contains a non-integer)
t_xdigit({1, 2, 'a'})   -- FALSE (contains a non-digit)
t_xdigit({})           -- FALSE (empty sequence)
```

8.14.3.19 t_vowel

```
include std/types.e
public type t_vowel(object test_data)
```

Returns TRUE if argument is a vowel or if every element of the argument is a vowel character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-vowels

8.14.3.19.1 Example 1:

```
t_vowel(-1)           -- FALSE
t_vowel(0)            -- FALSE
t_vowel(1)            -- FALSE
t_vowel(1.234)        -- FALSE
t_vowel('A')          -- TRUE
t_vowel('9')          -- FALSE
t_vowel('?')          -- FALSE
t_vowel("abc")        -- FALSE
t_vowel("aiu")        -- TRUE
t_vowel("123")        -- FALSE
t_vowel({1, 2, "abc"}) -- FALSE (contains a sequence)
t_vowel({1, 2, 9.7})  -- FALSE (contains a non-integer)
t_vowel({1, 2, 'a'})  -- FALSE (contains a non-digit)
t_vowel({})           -- FALSE (empty sequence)
```

8.14.3.20 t_consonant

```
include std/types.e
public type t_consonant(object test_data)
```

Returns TRUE if argument is a consonant character or if every element of the argument is an consonant character.

Returns FALSE if the argument is an empty sequence, or contains sequences, or contains non-consonant character.

8.14.3.20.1 Example 1:

```
t_consonant(-1)       -- FALSE
t_consonant(0)        -- FALSE
t_consonant(1)        -- FALSE
t_consonant(1.234)    -- FALSE
t_consonant('A')      -- FALSE
t_consonant('9')      -- FALSE
t_consonant('?')      -- FALSE
t_consonant("abc")    -- FALSE
t_consonant("rTfM")   -- TRUE
t_consonant("123")    -- FALSE
t_consonant({1, 2, "abc"}) -- FALSE (contains a sequence)
```

```
t_consonant({1, 2, 9.7})    -- FALSE (contains a non-integer)
t_consonant({1, 2, 'a'})    -- FALSE (contains a non-digit)
t_consonant({})             -- FALSE (empty sequence)
```

8.14.3.21 integer_array

```
include std/types.e
public type integer_array(object x)
```

8.14.3.21.1 Returns:

TRUE if argument is a sequence that only contains zero or more integers.

8.14.3.21.2 Example 1:

```
integer_array(-1)           -- FALSE (not a sequence)
integer_array("abc")        -- TRUE (all single characters)
integer_array({1, 2, "abc"}) -- FALSE (contains a sequence)
integer_array({1, 2, 9.7})  -- FALSE (contains a non-integer)
integer_array({1, 2, 'a'})  -- TRUE
integer_array({})           -- TRUE
```

8.14.3.22 t_text

```
include std/types.e
public type t_text(object x)
```

8.14.3.22.1 Returns:

TRUE if argument is a sequence that only contains zero or more characters.

8.14.3.22.2 Comment:

A 'character' is defined as a positive integer or zero. This is a broad definition that may be refined once proper UNICODE support is implemented.

8.14.3.22.3 Example 1:

```
t_text(-1)                  -- FALSE (not a sequence)
t_text("abc")               -- TRUE (all single characters)
t_text({1, 2, "abc"})       -- FALSE (contains a sequence)
t_text({1, 2, 9.7})         -- FALSE (contains a non-integer)
t_text({1, 2, 'a'})         -- TRUE
t_text({1, -2, 'a'})        -- FALSE (contains a negative integer)
```



```
t_text({})          -- TRUE
```

8.14.3.23 number_array

```
include std/types.e
public type number_array(object x)
```

8.14.3.23.1 Returns:

TRUE if argument is a sequence that only contains zero or more numbers.

8.14.3.23.2 Example 1:

```
number_array(-1)          -- FALSE (not a sequence)
number_array("abc")       -- TRUE  (all single characters)
number_array({1, 2, "abc"}) -- FALSE (contains a sequence)
number_array(1, 2, 9.7)   -- TRUE
number_array(1, 2, 'a')   -- TRUE
number_array({})          -- TRUE
```

8.14.3.24 sequence_array

```
include std/types.e
public type sequence_array(object x)
```

8.14.3.24.1 Returns:

TRUE if argument is a sequence that only contains zero or more sequences.

8.14.3.24.2 Example 1:

```
sequence_array(-1)          -- FALSE (not a sequence)
sequence_array("abc")       -- FALSE (all single characters)
sequence_array({1, 2, "abc"}) -- FALSE (contains some atoms)
sequence_array({1, 2, 9.7})  -- FALSE
sequence_array({1, 2, 'a'})  -- FALSE
sequence_array({"abc", {3.4, 99182.78737}}) -- TRUE
sequence_array({})          -- TRUE
```

8.14.3.25 ascii_string

```
include std/types.e
public type ascii_string(object x)
```

8.14.3.25.1 Returns:

TRUE if argument is a sequence that only contains zero or more ASCII characters.

8.14.3.25.2 Comment:

An ASCII 'character' is defined as a integer in the range [0 to 127].

8.14.3.25.3 Example 1:

```
ascii_string(-1)           -- FALSE (not a sequence)
ascii_string("abc")        -- TRUE  (all single ASCII characters)
ascii_string({1, 2, "abc"}) -- FALSE (contains a sequence)
ascii_string({1, 2, 9.7})   -- FALSE (contains a non-integer)
ascii_string({1, 2, 'a'})   -- TRUE
ascii_string({1, -2, 'a'})  -- FALSE (contains a negative integer)
ascii_string({})           -- TRUE
```

8.14.3.26 string

```
include std/types.e
public type string(object x)
```

8.14.3.26.1 Returns:

TRUE if argument is a sequence that only contains zero or more byte characters.

8.14.3.26.2 Comment:

A byte 'character' is defined as a integer in the range [0 to 255].

8.14.3.26.3 Example 1:

```
string(-1)           -- FALSE (not a sequence)
string("abc'6")       -- TRUE  (all single byte characters)
string({1, 2, "abc'6"}) -- FALSE (contains a sequence)
string({1, 2, 9.7})    -- FALSE (contains a non-integer)
string({1, 2, 'a'})    -- TRUE
string({1, -2, 'a'})   -- FALSE (contains a negative integer)
string({})           -- TRUE
```

9 Sequence Centric Routines

Data type conversion

Routines

Input Routines

Error Status Constants

Answer Types

Routines

Searching

Equality

Finding

Matching

Sequence Manipulation

Constants

Basic routines

Building sequences

Adding to sequences

Extracting, removing, replacing from/into a sequence

Changing the shape of a sequence

Serialization of EUPHORIA Objects

Routines

Sorting

Constants

Routines

9.1 Data type conversion

Routines

int_to_bytes

bytes_to_int

int_to_bits

bits_to_int

atom_to_float64

atom_to_float32

float64_to_atom

float32_to_atom

hex_text

set_decimal_mark

to_number

to_integer

9.1.1 Routines

9.1.1.1 int_to_bytes

```
include std/convert.e
public function int_to_bytes(atom x)
```

Converts an atom that represents an integer to a sequence of 4 bytes.

9.1.1.1.1 Parameters:

1. `x` : an atom, the value to convert.

9.1.1.1.2 Returns:

A **sequence**, of 4 bytes, lowest significant byte first.

9.1.1.1.3 Comments:

If the atom does not fit into a 32-bit integer, things may still work right:

- If there is a fractional part, the first element in the returned value will carry it. If you poke the sequence to RAM, that fraction will be discarded anyway.
- If `x` is simply too big, the first three bytes will still be correct, and the 4th element will be `floor(x/power(2, 24))`. If this is not a byte sized integer, some truncation may occur, but usually no error.

The integer can be negative. Negative byte-values will be returned, but after poking them into memory you will have the correct (two's complement) representation for the 386+.

9.1.1.1.4 Example 1:

```
s = int_to_bytes(999)
-- s is {231, 3, 0, 0}
```

9.1.1.1.5 Example 2:

```
s = int_to_bytes(-999)
-- s is {-231, -4, -1, -1}
```

9.1.1.1.6 See Also:

[bytes_to_int](#), [int_to_bits](#), [atom_to_float64](#), [poke4](#)

9.1.1.2 bytes_to_int

```
include std/convert.e
public function bytes_to_int(sequence s)
```

Converts a sequence of at most 4 bytes into an atom.

9.1.1.2.1 Parameters:

1. *s* : the sequence to convert

9.1.1.2.2 Returns:

An **atom**, the value of the concatenated bytes of *s*.

9.1.1.2.3 Comments:

This performs the reverse operation from [int_to_bytes](#)

An atom is being returned, because the converted value may be bigger than what can fit in an EUPHORIA integer.

9.1.1.2.4 Example 1:

```
atom int32

int32 = bytes_to_int({37,1,0,0})
-- int32 is 37 + 256*1 = 293
```

9.1.1.2.5 See Also:

[bits_to_int](#), [float64_to_atom](#), [int_to_bytes](#), [peek](#), [peek4s](#), [peek4u](#), [poke4](#)

9.1.1.3 int_to_bits

```
include std/convert.e
public function int_to_bits(atom x, integer nbits = 32)
```

Extracts the lower bits from an integer.

9.1.1.3.1 Parameters:

1. `x` : the atom to convert
2. `nbits` : the number of bits requested. The default is 32.

9.1.1.3.2 Returns:

A **sequence**, of length `nbits`, made of 1's and 0's.

9.1.1.3.3 Comments:

`x` should have no fractional part. If it does, then the first "bit" will be an atom between 0 and 2.

The bits are returned lowest first.

For negative numbers the two's complement bit pattern is returned.

You can use subscripting, slicing, and/or/xor/not of entire sequences etc. to manipulate sequences of bits. Shifting of bits and rotating of bits are easy to perform.

9.1.1.3.4 Example 1:

```
s = int_to_bits(177, 8)
-- s is {1,0,0,0,1,1,0,1} -- "reverse" order
```

9.1.1.3.5 See Also:

[bits_to_int](#), [int_to_bytes](#), [Relational operators](#), [operations on sequences](#)

9.1.1.4 bits_to_int

```
include std/convert.e
public function bits_to_int(sequence bits)
```

Converts a sequence of bits to an atom that has no fractional part.

9.1.1.4.1 Parameters:

1. `bits` : the sequence to convert.

9.1.1.4.2 Returns:

A positive **atom**, whose machine representation was given by `bits`.

9.1.1.4.3 Comments:

An element in `bits` can be any atom. If nonzero, it counts for 1, else for 0.

The first elements in `bits` represent the bits with the least weight in the returned value. Only the 52 last bits will matter, as the PC hardware cannot hold an integer with more digits than this.

If you print `s` the bits will appear in "reverse" order, but it is convenient to have increasing subscripts access bits of increasing significance.

9.1.1.4.4 Example 1:

```
a = bits_to_int({1,1,1,0,1})  
-- a is 23 (binary 10111)
```

9.1.1.4.5 See Also:

[bytes_to_int](#), [int_to_bits](#), [operations on sequences](#)

9.1.1.5 atom_to_float64

```
include std/convert.e  
public function atom_to_float64(atom a)
```

Convert an atom to a sequence of 8 bytes in IEEE 64-bit format

9.1.1.5.1 Parameters:

1. `a` : the atom to convert:

9.1.1.5.2 Returns:

A **sequence**, of 8 bytes, which can be poked in memory to represent `a`.

9.1.1.5.3 Comments:

All EUPHORIA atoms have values which can be represented as 64-bit IEEE floating-point numbers, so you can convert any atom to 64-bit format without losing any precision.

Integer values will also be converted to 64-bit floating-point format.

9.1.1.5.4 Example:

```
fn = open("numbers.dat", "wb")
puts(fn, atom_to_float64(157.82)) -- write 8 bytes to a file
```

9.1.1.5.5 See Also:

[float64_to_atom](#), [int_to_bytes](#), [atom_to_float32](#)

9.1.1.6 atom_to_float32

```
include std/convert.e
public function atom_to_float32(atom a)
```

Convert an atom to a sequence of 4 bytes in IEEE 32-bit format

9.1.1.6.1 Parameters:

1. *a* : the atom to convert:

9.1.1.6.2 Returns:

A **sequence**, of 4 bytes, which can be poked in memory to represent *a*.

9.1.1.6.3 Comments:

EUPHORIA atoms can have values which are 64-bit IEEE floating-point numbers, so you may lose precision when you convert to 32-bits (16 significant digits versus 7). The range of exponents is much larger in 64-bit format (10 to the 308, versus 10 to the 38), so some atoms may be too large or too small to represent in 32-bit format. In this case you will get one of the special 32-bit values: *inf* or *-inf* (infinity or -infinity). To avoid this, you can use [atom_to_float64\(\)](#).

Integer values will also be converted to 32-bit floating-point format.

On modern computers, computations on 64 bit floats are no slower than on 32 bit floats. Internally, the PC stores them in 80 bit registers anyway. EUPHORIA does not support these so called long doubles. Not all C compilers do.

9.1.1.6.4 Example 1:

```
fn = open("numbers.dat", "wb")
puts(fn, atom_to_float32(157.82)) -- write 4 bytes to a file
```

9.1.1.6.5 See Also:

[float32_to_atom](#), [int_to_bytes](#), [atom_to_float64](#)

9.1.1.7 float64_to_atom

```
include std/convert.e
public function float64_to_atom(sequence_8 ieee64)
```

Convert a sequence of 8 bytes in IEEE 64-bit format to an atom

9.1.1.7.1 Parameters:

1. `ieee64` : the sequence to convert:

9.1.1.7.2 Returns:

An **atom**, the same value as the FPU would see by peeking `ieee64` from RAM.

9.1.1.7.3 Comments:

Any 64-bit IEEE floating-point number can be converted to an atom.

9.1.1.7.4 Example 1:

```
f = repeat(0, 8)
fn = open("numbers.dat", "rb") -- read binary
for i = 1 to 8 do
    f[i] = getc(fn)
end for
a = float64_to_atom(f)
```

9.1.1.7.5 See Also:

[float32_to_atom](#), [bytes_to_int](#), [atom_to_float64](#)

9.1.1.8 float32_to_atom

```
include std/convert.e
public function float32_to_atom(sequence_4 ieee32)
```

Convert a sequence of 4 bytes in IEEE 32-bit format to an atom

9.1.1.8.1 Parameters:

1. `ieee32` : the sequence to convert:

9.1.1.8.2 Returns:

An **atom**, the same value as the FPU would see by peeking `ieee64` from RAM.

9.1.1.8.3 Comments:

Any 32-bit IEEE floating-point number can be converted to an atom.

9.1.1.8.4 Example 1:

```
f = repeat(0, 4)
fn = open("numbers.dat", "rb") -- read binary
f[1] = getc(fn)
f[2] = getc(fn)
f[3] = getc(fn)
f[4] = getc(fn)
a = float32_to_atom(f)
```

9.1.1.8.5 See Also:

[float64_to_atom](#), [bytes_to_int](#), [atom_to_float32](#)

9.1.1.9 hex_text

```
include std/convert.e
public function hex_text(sequence text)
```

Convert a text representation of a hexadecimal number to an atom

9.1.1.9.1 Parameters:

1. `text` : the text to convert.

9.1.1.9.2 Returns:

An **atom**, the numeric equivalent to `text`

9.1.1.9.3 Comments:

- The text can optionally begin with '#' which is ignored.
- The text can have any number of underscores, all of which are ignored.
- The text can have one leading '-', indicating a negative number.
- The text can have any number of underscores, all of which are ignored.

9.1.1.9.4 Example 1:

```
atom h = hex_text("-#3_4FA.00E_1BD")
-- h is now -13562.003444492816925
atom h = hex_text("DEADBEEF")
-- h is now 3735928559
```

9.1.1.9.5 See Also:

[value](#)

9.1.1.10 set_decimal_mark

```
include std/convert.e
public function set_decimal_mark(integer new_mark)
```

Gets, and possibly sets, the decimal mark that [to_number\(\)](#) uses.

9.1.1.10.1 Parameters:

1. `new_mark` : An integer: Either a comma (,), a period (.) or any other integer.

9.1.1.10.2 Returns:

An **integer**, The current value, before `new_mark` changes it.

9.1.1.10.3 Comments:

- When `new_mark` is a *period* it will cause `to_number()` to interpret a dot (.) as the decimal point symbol. The pre-changed value is returned.
- When `new_mark` is a *comma* it will cause `to_number()` to interpret a comma (,) as the decimal point symbol. The pre-changed value is returned.
- Any other value does not change the current setting. Instead it just returns the current value.
- The initial value of the decimal marker is a period.

9.1.1.11 to_number

```
include std/convert.e
public function to_number(sequence text_in, integer return_bad_pos = 0)
```

Converts the text into a number.

9.1.1.11.1 Parameters:

1. `text_in`: A string containing the text representation of a number.
2. `return_bad_pos`: An integer.
 - ◆ If 0 (the default) then this will return a number based on the supplied text and it will **not** return any position in `text_in` that caused an incomplete conversion.
 - ◆ If `return_bad_pos` is -1 then if the conversion of `text_in` was complete the resulting number is returned otherwise a single-element sequence containing the position within `text_in` where the conversion stopped.
 - ◆ If not 0 then this returns both the converted value up to the point of failure (if any) and the position in `text_in` that caused the failure. If that position is 0 then there was no failure.

9.1.1.11.2 Returns:

- an **atom**, If `return_bad_pos` is zero, the number represented by `text_in`. If `text_in` contains invalid characters, zero is returned.
- a **sequence**, If `return_bad_pos` is non-zero. If `return_bad_pos` is -1 it returns a 1-element sequence containing the spot inside `text_in` where conversion stopped. Otherwise it returns a 2-element sequence containing the number represented by `text_in` and either 0 or the position in `text_in` where conversion stopped.

9.1.1.11.3 Comments:

1. You can supply **Hexadecimal** values if the value is preceded by a '#' character, **Octal** values if the value is preceded by a '@' character, and **Binary** values if the value is preceded by a '!' character. With hexadecimal values, the case of the digits 'A' - 'F' is not important. Also, any decimal marker embedded in the number is used with the correct base.
2. Any underscore characters or thousands separators, that are embedded in the text number are ignored. These can be used to help visual clarity for long numbers. The thousands separator is a ',' when the

decimal mark is '.' (the default), or ',' if the decimal mark is ','. You inspect and set it using `set_decimal_mark()`.

3. You can supply a single leading or trailing sign. Either a minus (-) or plus (+).
4. You can supply one or more trailing adjacent percentage signs. The first one causes the resulting value to be divided by 100, and each subsequent one divides the result by a further 10. Thus 3845% gives a value of (3845 / 100) ==> 38.45, and 3845%% gives a value of (3845 / 1000) ==> 3.845.
5. You can have single currency symbol before the first digit or after the last digit. A currency symbol is any character of the string: "\$£¥".
6. You can have any number of whitespace characters before the first digit and after the last digit.
7. The currency, sign and base symbols can appear in any order. Thus "\$ -21.10" is the same as "-\$21.10", which is also the same as "21.10\$", etc.
8. This function can optionally return information about invalid numbers. If `return_bad_pos` is not zero, a two-element sequence is returned. The first element is the converted number value, and the second is the position in the text where conversion stopped. If no errors were found then the second element is zero.
9. When converting floating point text numbers to atoms, you need to be aware that many numbers cannot be accurately converted to the exact value expected due to the limitations of the 64-bit IEEE Floating point format.

9.1.1.11.4 Examples:

```
object val
val = to_number("12.34", 1) ---> {12.34, 0} -- No errors.
val = to_number("12.34", -1) ---> 12.34 -- No errors.
val = to_number("12.34a", 1) ---> {12.34, 6} -- Error at position 6
val = to_number("12.34a", -1) ---> {6} -- Error at position 6
val = to_number("12.34a") ---> 0 because its not a valid number
val = to_number("#f80c") --> 63500
val = to_number("#f80c.7aa") --> 63500.47900390625
val = to_number("@1703") --> 963
val = to_number("!101101") --> 45
val = to_number("12_583_891") --> 12583891
val = to_number("12_583_891%") --> 125838.91
val = to_number("12,583,891%%") --> 12583.891
```

9.1.1.12 to_integer

```
include std/convert.e
public function to_integer(object data_in, integer def_value = 0)
```

Converts an object into a integer.

9.1.1.12.1 Parameters:

1. `data_in`: Any EUPHORIA object.
2. `def_value`: An integer. This is returned if `data_in` cannot be converted into an integer. If omitted, zero is returned.

9.1.1.12.2 Returns:

An **integer**, either the integer rendition of `data_in` or `def_value` if it has no integer value.

9.1.1.12.3 Comments:

The returned value is guaranteed to be a valid EUPHORIA integer.

9.1.1.12.4 Examples:

```
? to_integer(12)           --> 12
? to_integer(12.4)        --> 12
? to_integer("12")        --> 12
? to_integer("12.9")      --> 12
? to_integer("a12")       --> 0 (not a valid number)
? to_integer("a12",-1)    --> -1 (not a valid number)
? to_integer({"12"})      --> 0 (sub-sequence found)
? to_integer(#3FFFFFFF)   --> 1073741823
? to_integer(#3FFFFFFF + 1) --> 0 (too big for a EUPHORIA integer)
```

9.2 Input Routines

Error Status Constants

GET_SUCCESS

GET_EOF

GET_FAIL

GET_NOTHING

Answer Types

GET_SHORT_ANSWER

GET_LONG_ANSWER

Routines

get

value

defaulted_value

9.2.1 Error Status Constants

These are returned from `get` and `value`.

9.2.1.1 GET_SUCCESS

```
include std/get.e
public constant GET_SUCCESS
```

9.2.1.2 GET_EOF

```
include std/get.e
public constant GET_EOF
```

9.2.1.3 GET_FAIL

```
include std/get.e
public constant GET_FAIL
```

9.2.1.4 GET_NOTHING

```
include std/get.e
public constant GET_NOTHING
```

9.2.2 Answer Types

9.2.2.1 GET_SHORT_ANSWER

```
include std/get.e
public constant GET_SHORT_ANSWER
```

9.2.2.2 GET_LONG_ANSWER

```
include std/get.e
public constant GET_LONG_ANSWER
```

9.2.3 Routines

9.2.3.1 get

```
include std/get.e
public function get(integer file, integer offset = 0, integer answer = GET_SHORT_ANSWER)
```

Input, from an open file, a human-readable string of characters representing a EUPHORIA object. Convert the string into the numeric value of that object.

9.2.3.1.1 Parameters:

1. `file` : an integer, the handle to an open file from which to read
2. `offset` : an integer, an offset to apply to file position before reading. Defaults to 0.
3. `answer` : an integer, either `GET_SHORT_ANSWER` (the default) or `GET_LONG_ANSWER`.

9.2.3.1.2 Returns:

A **sequence**, of length 2 (`GET_SHORT_ANSWER`) or 4 (`GET_LONG_ANSWER`), made of

- an integer, the return status. This is any of:
 - ◆ `GET_SUCCESS` -- object was read successfully
 - ◆ `GET_EOF` -- end of file before object was read completely
 - ◆ `GET_FAIL` -- object is not syntactically correct
 - ◆ `GET_NOTHING` -- nothing was read, even a partial object string, before end of input
- an object, the value that was read. This is valid only if return status is `GET_SUCCESS`.
- an integer, the number of characters read. On an error, this is the point at which the error was detected.
- an integer, the amount of initial whitespace read before the first active character was found

9.2.3.1.3 Comments:

When `answer` is not specified, or explicitly `GET_SHORT_ANSWER`, only the first two elements in the returned sequence are actually returned.

The `GET_NOTHING` return status will not be returned if `answer` is `GET_SHORT_ANSWER`.

`get()` can read arbitrarily complicated EUPHORIA objects. You could have a long sequence of values in braces and separated by commas and comments, e.g. `{23, {49, 57}, 0.5, -1, 99, 'A', "john"}`. A single call to `get()` will read in this entire sequence and return its value as a result, as well as complementary information.

If a nonzero offset is supplied, it is interpreted as an offset to the current file position, and the file will be `seek()`ed there first.

`get()` returns a 2 or 4 element sequence, like `value()` does:

- a status code (success/error/end of file/no value at all)
- the value just read (meaningful only when the status code is `GET_SUCCESS`) (optionally)
- the total number of characters read
- the amount of initial whitespace read.

Using the default value for `answer`, or setting it to `GET_SHORT_ANSWER`, returns 2 elements. Setting it to `GET_LONG_ANSWER` causes 4 elements to be returned.

Each call to `get()` picks up where the previous call left off. For instance, a series of 5 calls to `get()` would be needed to read in


```
"99 5.2 {1, 2, 3} "Hello" -1"
```

On the sixth and any subsequent call to `get ()` you would see a `GET_EOF` status. If you had something like

```
{1, 2, xxx}
```

in the input stream you would see a `GET_FAIL` error status because `xxx` is not a EUPHORIA object. And seeing

```
-- something\nBut no value
```

and the input stream stops right there, you'll receive a status code of `GET_NOTHING`, because nothing but whitespace or comments was read. If you had opted for a short answer, you'd get `GET_EOF` instead.

Multiple "top-level" objects in the input stream must be separated from each other with one or more "whitespace" characters (blank, tab, `\r` or `\n`). At the very least, a top level number must be followed by a white space from the following object. Whitespace is not necessary *within* a top-level object. Comments, terminated by either `\n` or `\r`, are allowed anywhere inside sequences, and ignored if at the top level. A call to `get ()` will read one entire top-level object, plus possibly one additional (whitespace) character, after a top level number, even though the next object may have an identifiable starting point.

The combination of `print ()` and `get ()` can be used to save a EUPHORIA object to disk and later read it back. This technique could be used to implement a database as one or more large EUPHORIA sequences stored in disk files. The sequences could be read into memory, updated and then written back to disk after each series of transactions is complete. Remember to write out a whitespace character (using `puts ()`) after each call to `print ()`, at least when a top level number was just printed.

The value returned is not meaningful unless you have a `GET_SUCCESS` status.

9.2.3.1.4 Example 1:

```
-- If he types 77.5, get(0) would return:  
{GET_SUCCESS, 77.5}
```

```
-- whereas gets(0) would return:  
"77.5\n"
```

9.2.3.1.5 Example 2:

See `bin\mydata.ex`

9.2.3.1.6 See Also:

[value](#)

9.2.3.2 value

```
include std/get.e
public function value(sequence st, integer start_point = 1, integer answer = GET_SHORT_ANSWER)
```

Read, from a string, a human-readable string of characters representing a EUPHORIA object. Convert the string into the numeric value of that object.

9.2.3.2.1 Parameters:

1. `st` : a sequence, from which to read text
2. `offset` : an integer, the position at which to start reading. Defaults to 1.
3. `answer` : an integer, either `GET_SHORT_ANSWER` (the default) or `GET_LONG_ANSWER`.

9.2.3.2.2 Returns:

A **sequence**, of length 2 (`GET_SHORT_ANSWER`) or 4 (`GET_LONG_ANSWER`), made of

- an integer, the return status. This is any of
 - ◆ `GET_SUCCESS` -- object was read successfully
 - ◆ `GET_EOF` -- end of file before object was read completely
 - ◆ `GET_FAIL` -- object is not syntactically correct
 - ◆ `GET_NOTHING` -- nothing was read, even a partial object string, before end of input
- an object, the value that was read. This is valid only if return status is `GET_SUCCESS`.
- an integer, the number of characters read. On an error, this is the point at which the error was detected.
- an integer, the amount of initial whitespace read before the first active character was found

9.2.3.2.3 Comments:

When `answer` is not specified, or explicitly `GET_SHORT_ANSWER`, only the first two elements in the returned sequence are actually returned.

This works the same as `get()`, but it reads from a string that you supply, rather than from a file or device.

After reading one valid representation of a EUPHORIA object, `value()` will stop reading and ignore any additional characters in the string. For example, "36" and "36P" will both give you `{GET_SUCCESS, 36}`.

The function returns `{return_status, value}` if the answer type is not passed or set to `GET_SHORT_ANSWER`. If set to `GET_LONG_ANSWER`, the number of characters read and the amount of leading whitespace are returned in 3rd and 4th position. The `GET_NOTHING` return status can occur only on a long answer.

9.2.3.2.4 Example 1:

```
s = value("12345")
s is {GET_SUCCESS, 12345}
```

9.2.3.2.5 Example 2:

```
s = value("{0, 1, -99.9}")
-- s is {GET_SUCCESS, {0, 1, -99.9}}
```

9.2.3.2.6 Example 3:

```
s = value("+++")
-- s is {GET_FAIL, 0}
```

9.2.3.2.7 See Also:

[get](#)

9.2.3.3 defaulted_value

```
include std/get.e
public function defaulted_value(object st, object def, integer start_point = 1)
```

Perform a value() operation on a sequence, returning the value on success or the default on failure.

9.2.3.3.1 Parameters:

1. st : object to retrieve value from.
2. def : the value returned if st is an atom or value(st) fails.
3. start_point : an integer, the position in st at which to start getting the value from. Defaults to 1

9.2.3.3.2 Returns:

- If st, is an atom then def is returned.
- If value(st), call is a success, then value() [2], otherwise it will return the parameter #def#.

9.2.3.3.3 Examples:

```
object i = defaulted_value("10", 0)
-- i is 10
```

```
i = defaulted_value("abc", 39)
-- i is 39
```

```
i = defaulted_value(12, 42)
-- i is 42

i = defaulted_value("{1,2}", 42)
-- i is {1,2}
```

9.2.3.3.4 See Also:

[value](#)

9.3 Searching

Page Contents

- Equality
 - [compare](#)
 - [equal](#)
- Finding
 - [find](#)
 - [find_from](#)
 - [find_any](#)
 - [find_all](#)
 - [NESTED_ANY](#)
 - [NESTED_ALL](#)
 - [NESTED_INDEX](#)
 - [NESTED_BACKWARD](#)
 - [find_nested](#)
 - [rfind](#)
 - [find_replace](#)
 - [match_replace](#)
 - [binary_search](#)
- Matching
 - [match](#)
 - [match_from](#)
 - [match_all](#)
 - [rmatch](#)
 - [begins](#)
 - [ends](#)
 - [is_in_range](#)
 - [is_in_list](#)
 - [lookup](#)
 - [vlookup](#)

9.3.1 Equality

9.3.1.1 compare

```
<built-in> function compare(object compared, object reference)
```

Compare two items returning less than, equal or greater than.

9.3.1.1.1 Parameters:

1. `compared` : the compared object
2. `reference` : the reference object

9.3.1.1.2 Returns:

An integer,

- 0 -- if objects are identical
- 1 -- if `compared` is greater than `reference`
- -1 -- if `compared` is less than `reference`

9.3.1.1.3 Comments:

Atoms are considered to be less than sequences. Sequences are compared alphabetically starting with the first element until a difference is found or one of the sequences is exhausted. Atoms are compared as ordinary reals.

9.3.1.1.4 Example 1:

```
x = compare({1,2,{3,{4}}},5), {2-1,1+1,{3,{4}}},6-1))
-- identical, x is 0
```

9.3.1.1.5 Example 2:

```
if compare("ABC", "ABCD") < 0 then -- -1
    -- will be true: ABC is "less" because it is shorter
end if
```

9.3.1.1.6 Example 3:

```
x = compare('a', "a")
-- x will be -1 because 'a' is an atom
-- while "a" is a sequence
```

9.3.1.1.7 See Also:

[equal](#), [relational operators](#), [operations on sequences](#), [sort](#)

9.3.1.2 equal

<built-in> `function equal(object left, object right)`

Compare two EUPHORIA objects to see if they are the same.

9.3.1.2.1 Parameters:

1. `left` : one of the objects to test
2. `right` : the other object

9.3.1.2.2 Returns:

An **integer**, 1 if the two objects are identical, else 0.

9.3.1.2.3 Comments:

This is equivalent to the expression: `compare(left, right) = 0`.

This routine, like most other built-in routines, is very fast. It does not have any subroutine call overhead.

9.3.1.2.4 Example 1:

```
if equal(PI, 3.14) then
    puts(1, "give me a better value for PI!\n")
end if
```

9.3.1.2.5 Example 2:

```
if equal(name, "George") or equal(name, "GEORGE") then
    puts(1, "name is George\n")
end if
```

9.3.1.2.6 See Also:

[compare](#)

9.3.2 Finding

9.3.2.1 find

<built-in> `function find(object needle, sequence haystack, integer start)`

Find the first occurrence of a "needle" as an element of a "haystack", starting from position "start"..

9.3.2.1.1 Parameters:

1. `needle` : an object whose presence is being queried
2. `haystack` : a sequence, which is being looked up for `needle`
3. `start` : an integer, the position at which to start searching. Defaults to 1.

9.3.2.1.2 Returns:

An **integer**, 0 if `needle` is not on `haystack`, else the smallest index of an element of `haystack` that equals `needle`.

9.3.2.1.3 Comments:

`find()` and `find_from()` are identical, but you can omit giving `find()` a starting point.

9.3.2.1.4 Example 1:

```
location = find(11, {5, 8, 11, 2, 3})
-- location is set to 3
```

9.3.2.1.5 Example 2:

```
names = {"fred", "rob", "george", "mary", ""}
location = find("mary", names)
-- location is set to 4
```

9.3.2.1.6 See Also:

[find_from](#), [match](#), [match_from](#), [compare](#)

9.3.2.2 find_from

```
<built-in> function find_from(object needle, object haystack, integer start)
```

Find the first occurrence of a "needle" as an element of a "haystack". Search starts at a specified index.

9.3.2.2.1 Parameters:

1. `needle` : an object whose presence is being queried
2. `haystack` : a sequence, which is being looked up for `needle`
3. `start` : an integer, the index in `haystack` at which to start searching.

9.3.2.2.2 Returns:

An **integer**, 0 if `needle` is not on `haystack` past position `start`, else the smallest index, not less than `start`, of an element of `haystack` that equals `needle`.

9.3.2.2.3 Comments:

`start` may have any value from 1 to the length of `haystack` plus 1. (Analogous to the first index of a slice of `haystack`).

`find()` and `find_from()` are identical, but you can omit giving `find()` a starting point.

9.3.2.2.4 Example 1:

```
location = find_from(11, {11, 8, 11, 2, 3}, 2)
-- location is set to 3
```

9.3.2.2.5 Example 2:

```
names = {"mary", "rob", "george", "mary", ""}
location = find_from("mary", names, 3)
-- location is set to 4
```

9.3.2.2.6 See Also:

[find](#), [match](#), [match_from](#), [compare](#)

9.3.2.3 find_any

```
include std/search.e
public function find_any(sequence needles, sequence haystack, integer start = 1)
```


Find any element from a list inside a sequence. Returns the location of the first hit.

9.3.2.3.1 Parameters:

1. `needles` : a sequence, the list of items to look for
2. `haystack` : a sequence, in which "needles" are looked for
3. `start` : an integer, the starting point of the search. Defaults to 1.

9.3.2.3.2 Returns:

An **integer**, the smallest index in `haystack` of an element of `needles`, or 0 if no needle is found.

9.3.2.3.3 Comments:

This function may be applied to a string sequence or a complex sequence.

9.3.2.3.4 Example 1:

```
location = find_any("aeiou", "John Smith", 3)
-- location is 8
```

9.3.2.3.5 Example 2:

```
location = find_any("aeiou", "John Doe")
-- location is 2
```

9.3.2.3.6 See Also:

[find](#), [find_from](#)

9.3.2.4 find_all

```
include std/search.e
public function find_all(object needle, sequence haystack, integer start = 1)
```

Find all occurrences of an object inside a sequence, starting at some specified point.

9.3.2.4.1 Parameters:

1. `needle` : an object, what to look for
2. `haystack` : a sequence to search in
3. `start` : an integer, the starting index position (defaults to 1)

9.3.2.4.2 Returns:

A **sequence**, the list of all indexes no less than `start` of elements of `haystack` that equal `needle`. This sequence is empty if no match found.

9.3.2.4.3 Example 1:

```
s = find_all('A', "ABCABAB")
-- s is {1,4,6}
```

9.3.2.4.4 See Also:

[find](#), [match](#), [match_all](#)

9.3.2.5 NESTED_ANY

```
include std/search.e
public constant NESTED_ANY
```

9.3.2.6 NESTED_ALL

```
include std/search.e
public constant NESTED_ALL
```

9.3.2.7 NESTED_INDEX

```
include std/search.e
public constant NESTED_INDEX
```

9.3.2.8 NESTED_BACKWARD

```
include std/search.e
public constant NESTED_BACKWARD
```

9.3.2.9 find_nested

```
include std/search.e
public function find_nested(object needle, sequence haystack, integer flags = 0, integer rtn_idx)
```

Find any object (among a list) in a sequence of arbitrary shape at arbitrary nesting.

9.3.2.9.1 Parameters:

1. `needle` : an object, either what to look up, or a list of items to look up
2. `haystack` : a sequence, where to look up
3. `flags` : options to the function, see Comments section. Defaults to 0.
4. `routine` : an integer, the `routine_id` of an user supplied equal function. Defaults to -1.

9.3.2.9.2 Returns:

A possibly empty **sequence**, of results, one for each hit.

9.3.2.9.3 Comments:

Each item in the returned sequence is either a sequence of indexes, or a pair {sequence of indexes, index in `needle`}.

9.3.2.9.4 The following flags are available to fine tune the search:

- `NESTED_BACKWARD` -- if on `flags`, search is performed backward. Default is forward.
- `NESTED_ALL` -- if on `flags`, all occurrences are looked for. Default is one hit only.
- `NESTED_ANY` -- if present on `flags`, `needle` is a list of items to look for. Not the default.
- `NESTED_INDEXES` -- if present on `flags`, an individual result is a pair {position, index in `needle`}. Default is just return the position.

If `s` is a single index list, or position, from the returned sequence, then `fetch(haystack, s) = needle`.

If a routine id is supplied, the routine must behave like `equal()` if the `NESTED_ANY` flag is not supplied, and like `find()` if it is. The routine is being passed the current `haystack` item and `needle`. The returned integer is interpreted as if returned by `equal()` or `find()`.

If the `NESTED_ANY` flag is specified, and `needle` is an atom, then the flag is removed.

9.3.2.9.5 Example 1:

```
sequence s = find_nested(3, {5, {4, {3, {2}}}})
-- s is {2, 2, 1}
```

9.3.2.9.6 Example 2:

```
sequence s = find_nested({3, 2}, {1, 3, {2, 3}}, NESTED_ANY + NESTED_BACKWARD + NESTED_ALL)
-- s is {{3, 2}, {3, 1}, {2}}
```

9.3.2.9.7 Example 3:

```
sequence s = find_nested({3, 2}, {1, 3, {2,3}}, NESTED_ANY + NESTED_INDEXES + NESTED_ALL)
-- s is {{{2}, 1}, {{3, 1}, 2}, {{3, 2}, 1}}
```

9.3.2.9.8 See Also:

[find](#), [rfind](#), [find_any](#), [fetch](#)

9.3.2.10 rfind

```
include std/search.e
public function rfind(object needle, sequence haystack, integer start = length(haystack))
```

Find a needle in a haystack in reverse order.

9.3.2.10.1 Parameters:

1. `needle` : an object to search for
2. `haystack` : a sequence to search in
3. `start` : an integer, the starting index position (defaults to `length(haystack)`)

9.3.2.10.2 Returns:

An **integer**, 0 if no instance of `needle` can be found on `haystack` before index `start`, or the highest such index otherwise.

9.3.2.10.3 Comments:

If `start` is less than 1, it will be added once to `length(haystack)` to designate a position counted backwards. Thus, if `start` is -1, the first element to be queried in `haystack` will be `haystack[$-1]`, then `haystack[$-2]` and so on.

9.3.2.10.4 Example 1:

```
location = rfind(11, {5, 8, 11, 2, 11, 3})
-- location is set to 5
```

9.3.2.10.5 Example 2:

```
names = {"fred", "rob", "rob", "george", "mary"}
location = rfind("rob", names)
-- location is set to 3
location = rfind("rob", names, -4)
```

```
-- location is set to 2
```

9.3.2.10.6 See Also:

[find](#), [rmatch](#)

9.3.2.11 find_replace

```
include std/search.e
public function find_replace(object needle, sequence haystack, object replacement, integer max
```

Finds a "needle" in a "haystack", and replace any, or only the first few, occurrences with a replacement.

9.3.2.11.1 Parameters:

1. `needle` : an object to search and perhaps replace
2. `haystack` : a sequence to be inspected
3. `replacement` : an object to substitute for any (first) instance of `needle`
4. `max` : an integer, 0 to replace all occurrences

9.3.2.11.2 Returns:

A `sequence`, the modified `haystack`.

9.3.2.11.3 Comments:

Replacements will not be made recursively on the part of `haystack` that was already changed.

If `max` is 0 or less, any occurrence of `needle` in `haystack` will be replaced by `replacement`. Otherwise, only the first `max` occurrences are.

9.3.2.11.4 Example 1:

```
s = find_replace('b', "The batty book was all but in Canada.", 'c', 0)
-- s is "The catty cook was all cut in Canada."
```

9.3.2.11.5 Example 2:

```
s = find_replace('/', "/euphoria/demo/unix", '\\', 2)
-- s is "\\euphoria\\demo/unix"
```

9.3.2.11.6 Example 3:

```
s = find_replace("theater", { "the", "theater", "theif" }, "theatre")
-- s is { "the", "theatre", "theif" }
```

9.3.2.11.7 See Also:

[find](#), [replace](#), [match_replace](#)

9.3.2.12 match_replace

```
include std/search.e
public function match_replace(object needle, sequence haystack, object replacement, integer max)
```

Finds a "needle" in a "haystack", and replace any, or only the first few, occurrences with a replacement.

9.3.2.12.1 Parameters:

1. `needle` : an object to search and perhaps replace
2. `haystack` : a sequence to be inspected
3. `replacement` : an object to substitute for any (first) instance of `needle`
4. `max` : an integer, 0 to replace all occurrences

9.3.2.12.2 Returns:

A **sequence**, the modified `haystack`.

9.3.2.12.3 Comments:

Replacements will not be made recursively on the part of `haystack` that was already changed.

If `max` is 0 or less, any occurrence of `needle` in `haystack` will be replaced by `replacement`. Otherwise, only the first `max` occurrences are.

If either `needle` or `replacement` are atoms they will be treated as if you had passed in a length-1 sequence containing the said atom.

9.3.2.12.4 Example 1:

```
s = match_replace("the", "the cat ate the food under the table", "THE", 0)
-- s is "THE cat ate THE food under THE table"
```

9.3.2.12.5 Example 2:

```
s = match_replace("the", "the cat ate the food under the table", "THE", 2)
-- s is "THE cat ate THE food under the table"
```

9.3.2.12.6 Example 3:

```
s = match_replace('/', "/euphoria/demo/unix", '\\', 2)
-- s is "\\euphoria\\demo/unix"
```

9.3.2.12.7 See Also:

[find](#), [replace](#), [find_replace](#)

9.3.2.13 binary_search

```
include std/search.e
public function binary_search(object needle, sequence haystack, integer start_point = 1, integer end_point = length(haystack))
```

Finds a "needle" in an ordered "haystack". Start and end point can be given for the search.

9.3.2.13.1 Parameters:

1. `needle` : an object to look for
2. `haystack` : a sequence to search in
3. `start_point` : an integer, the index at which to start searching. Defaults to 1.
4. `end_point` : an integer, the end point of the search. Defaults to 0, ie search to end.

9.3.2.13.2 Returns:

An **integer**, either:

1. a positive integer `i`, which means `haystack[i]` equals `needle`.
2. a negative integer, `-i`, with `i` between adjusted start and end points. This means that `needle` is not in the searched slice of `haystack`, but would be at index `i` if it were there.
3. a negative integer `-i` with `i` out of the searched range. This means that `needle` might be either below the start point if `i` is below the start point, or above the end point if `i` is.

9.3.2.13.3 Comments:

- If `end_point` is not greater than zero, it is added to `length(haystack)` once only. Then, the end point of the search is adjusted to `length(haystack)` if out of bounds.
- The start point is adjusted to 1 if below 1.

- The way this function returns is very similar to what [db_find_key](#) does. They use variants of the same algorithm. The latter is all the more efficient as `haystack` is long.
- `haystack` is assumed to be in ascending order. Results are undefined if it is not.
- If duplicate copies of `needle` exist in the range searched on `haystack`, any of the possible contiguous indexes may be returned.

9.3.2.13.4 See Also:

[find](#), [db_find_key](#)

9.3.3 Matching

9.3.3.1 match

```
<built-in> function match(sequence needle, sequence haystack, integer start)
```

Try to match a "needle" against some slice of a "haystack", starting at position "start".

9.3.3.1.1 Parameters:

1. `needle` : a sequence whose presence as a "substring" is being queried
2. `haystack` : a sequence, which is being looked up for `needle` as a sub-sequence
3. `start` : an integer, the point from which matching is attempted. Defaults to 1.

9.3.3.1.2 Returns:

An **integer**, 0 if no slice of `haystack` is `needle`, else the smallest index at which such a slice starts.

9.3.3.1.3 Comments:

`match()` and [match_from\(\)](#) are identical, but you can omit giving `match()` a starting point.

9.3.3.1.4 Example 1:

```
location = match("pho", "EUPHORIA")  
-- location is set to 3
```

9.3.3.1.5 See Also:

[find](#), [find_from](#), [compare](#), [match_from](#), [wildcard_match](#)

9.3.3.2 match_from

```
<built-in> function match_from(sequence needle, sequence haystack, integer start)
```

Try to match a "needle" against some slice of a "haystack", starting from some index.

9.3.3.2.1 Parameters:

1. `needle` : an sequence whose presence as a sub-sequence is being queried
2. `haystack` : a sequence, which is being looked up for `needle` as a sub-sequence
3. `start` : an integer, the index in `haystack` at which to start searching.

9.3.3.2.2 Returns:

An **integer**, 0 if no slice of `haystack` with lower index at least `start` is `needle`, else the smallest such index.

9.3.3.2.3 Comments:

`start` may have any value from 1 to the length of `haystack` plus 1. (Just like the first index of a slice of `haystack`.)

`match()` and `match_from()` are identical, but you can omit giving `match()` a starting point.

9.3.3.2.4 Example 1:

```
location = match_from("pho", "phoEUPHORIA", 4)
-- location is set to 6
```

9.3.3.2.5 See Also:

[find](#), [find_from](#), [match](#), [compare](#), [wildcard_match](#), [regex:find](#)

9.3.3.3 match_all

```
include std/search.e
public function match_all(sequence needle, sequence haystack, integer start = 1)
```

Match all items of `haystack` in `needle`.

9.3.3.3.1 Parameters:

1. `needle` : a sequence, what to look for
2. `haystack` : a sequence to search in
3. `start` : an integer, the starting index position (defaults to 1)

9.3.3.3.2 Returns:

A **sequence**, of integers, the list of all lower indexes, not less than `start`, of all slices in `haystack` that equal `needle`. The list may be empty.

9.3.3.3.3 Example 1:

```
s = match_all("the", "the dog chased the cat under the table.")
-- s is {1,16,30}
```

9.3.3.3.4 See Also:

[match](#), [find](#), [find_all](#)

9.3.3.4 rmatch

```
include std/search.e
public function rmatch(sequence needle, sequence haystack, integer start = length(haystack))
```

Try to match a needle against some slice of a haystack in reverse order.

9.3.3.4.1 Parameters:

1. `needle` : a sequence to search for
2. `haystack` : a sequence to search in
3. `start` : an integer, the starting index position (defaults to `length(haystack)`)

9.3.3.4.2 Returns:

An **integer**, either 0 if no slice of `haystack` starting before `start` equals `needle`, else the highest lower index of such a slice.

9.3.3.4.3 Comments:

If `start` is less than 1, it will be added once to `length(haystack)` to designate a position counted backwards. Thus, if `start` is -1, the first element to be queried in `haystack` will be `haystack[$-1]`, then `haystack[$-2]` and so on.

9.3.3.4.4 Example 1:

```
location = rmatch("the", "the dog ate the steak from the table.")
-- location is set to 28 (3rd 'the')
location = rmatch("the", "the dog ate the steak from the table.", -11)
-- location is set to 13 (2nd 'the')
```

9.3.3.4.5 See Also:

[rfind](#), [match](#)

9.3.3.5 begins

```
include std/search.e
public function begins(object sub_text, sequence full_text)
```

Test whether a sequence is the head of another one.

9.3.3.5.1 Parameters:

1. `sub_text` : an object to be looked for
2. `full_text` : a sequence, the head of which is being inspected.

9.3.3.5.2 Returns:

An **integer**, 1 if `sub_text` begins `full_text`, else 0.

9.3.3.5.3 Example 1:

```
s = begins("abc", "abcdef")
-- s is 1
s = begins("bcd", "abcdef")
-- s is 0
```

9.3.3.5.4 See Also:

[ends](#), [head](#)

9.3.3.6 ends

```
include std/search.e
public function ends(object sub_text, sequence full_text)
```

Test whether a sequence ends another one.

9.3.3.6.1 Parameters:

1. `sub_text` : an object to be looked for
2. `full_text` : a sequence, the tail of which is being inspected.

9.3.3.6.2 Returns:

An **integer**, 1 if `sub_text` ends `full_text`, else 0.

9.3.3.6.3 Example 1:

```
s = ends("def", "abcdef")
-- s is 1
s = begins("bcd", "abcdef")
-- s is 0
```

9.3.3.6.4 See Also:

[begins](#), [tail](#)

9.3.3.7 is_in_range

```
include std/search.e
public function is_in_range(object item, sequence range_limits, sequence boundries = "[")
```

Tests to see if the `item` is in a range of values supplied by `range_limits`

9.3.3.7.1 Parameters:

1. `item` : The object to test for.
2. `range_limits` : A sequence of two or more elements. The first is assumed to be the smallest value and the last is assumed to be the highest value.
3. `boundries` : a sequence. This determines if the range limits are inclusive or not. Must be one of "[", "]", "(", or ")".

9.3.3.7.2 Returns:

An **integer**, 0 if `item` is not in the `range_limits` otherwise it returns 1.

9.3.3.7.3 Comments:

- In `boundries#`, square brackets mean *inclusive* and round brackets mean *exclusive*. Thus `[]` includes both limits in the range, while `()` excludes both limits. And, `[)` includes the lower limit and excludes the upper limits while `])` does the reverse.

9.3.3.7.4 Example 1:

```
if is_in_range(2, {2, 75}) then
    procA(user_data) -- Gets run (both limits included)
end if
if is_in_range(2, {2, 75}, "()") then
    procA(user_data) -- Does not get run
end if
```

9.3.3.8 is_in_list

```
include std/search.e
public function is_in_list(object item, sequence list)
```

Tests to see if the `item` is in a list of values supplied by `list`

9.3.3.8.1 Parameters:

1. `item`: The object to test for.
2. `list`: A sequence of elements that `item` could be a member of.

9.3.3.8.2 Returns:

An **integer**, 0 if `item` is not in the `list`, otherwise it returns 1.

9.3.3.8.3 Example 1:

```
if is_in_list(user_data, {100, 45, 2, 75, 121}) then
    procA(user_data)
end if
```

9.3.3.9 lookup

```
include std/search.e
public function lookup(object find_item, sequence source_list, sequence target_list, object def
```

If the supplied item is in the source list, this returns the corresponding element from the target list.

9.3.3.9.1 Parameters:

1. `find_item`: an object that might exist in `source_list`.
2. `source_list`: a sequence that might contain `pItem`.
3. `target_list`: a sequence from which the corresponding item will be returned.
4. `def_value`: an object (defaults to zero). This is returned when `find_item` is not in `source_list` **and** `target_list` is not longer than `source_list`.

9.3.3.9.2 Returns:

an object

- If `find_item` is found in `source_list` then this is the corresponding element from `target_list`
- If `find_item` is not in `source_list` then if `target_list` is longer than `source_list` then the last item in `target_list` is returned otherwise `def_value` is returned.

9.3.3.9.3 Examples:

```
lookup('a', "cat", "dog") --> 'o'
lookup('d', "cat", "dogx") --> 'x'
lookup('d', "cat", "dog") --> 0
lookup('d', "cat", "dog", -1) --> -1
lookup("ant", {"ant", "bear", "cat"}, {"spider", "seal", "dog", "unknown"}) --> "spider"
lookup("dog", {"ant", "bear", "cat"}, {"spider", "seal", "dog", "unknown"}) --> "unknown"
```

9.3.3.10 vlookup

```
include std/search.e
public function vlookup(object find_item, sequence grid_data, integer source_col, integer target_col)
```

If the supplied item is in a source grid column, this returns the corresponding element from the target column.

9.3.3.10.1 Parameters:

1. `find_item`: an object that might exist in `source_col`.
2. `grid_data`: a 2D grid sequence that might contain `pItem`.
3. `source_col`: an integer. The column number to look for `find_item`.
4. `target_col`: an integer. The column number from which the corresponding item will be returned.
5. `def_value`: an object (defaults to zero). This is returned when `find_item` is not found in the `source_col` column, or if found but the target column does not exist.

9.3.3.10.2 Comments:

- If a row in the grid is actually a single atom, the row is ignored.
- If a row's length is less than the `source_col`, the row is ignored.

9.3.3.10.3 Returns:

an object

- If `find_item` is found in the `source_col` column then this is the corresponding element from the `target_col` column.

9.3.3.10.4 Examples:

```
sequence grid
grid = {
    {"ant", "spider", "mortein"},
    {"bear", "seal", "gun"},
    {"cat", "dog", "ranger"},
    $
}
vlookup("ant", grid, 1, 2, "?") --> "spider"
vlookup("ant", grid, 1, 3, "?") --> "mortein"
vlookup("seal", grid, 2, 3, "?") --> "gun"
vlookup("seal", grid, 2, 1, "?") --> "bear"
vlookup("mouse", grid, 2, 3, "?") --> "?"
```

9.4 Sequence Manipulation

Constants

ADD_PREPEND
ADD_APPEND
ADD_SORT_UP
ADD_SORT_DOWN
ROTATE_LEFT
ROTATE_RIGHT

Basic routines

can_add
fetch
store
valid_index
rotate
columnize
apply
mapping
length
reverse
shuffle

Building sequences

- linear
- repeat_pattern
- repeat

Adding to sequences

- append
- prepend
- insert
- splice
- pad_head
- pad_tail
- add_item
- remove_item

Extracting, removing, replacing from/into a sequence

- head
- tail
- mid
- slice
- vslice
- remove
- patch
- remove_all
- retain_all
- filter
- STDFLTR_ALPHA
- replace
- replace_all
- extract
- project

Changing the shape of a sequence

- split
- split_any
- join
- BK_LEN
- BK_PIECES
- breakup
- flatten
- pivot
- build_list
- transform
- sim_index
- SEQ_NOALT
- remove_subseq
- RD_INPLACE
- remove_dups
- COMBINE_UNSORTED
- COMBINE_SORTED
- combine
- minsize

9.4.1 Constants

9.4.1.1 ADD_PREPEND

```
include std/sequence.e
public enum ADD_PREPEND
```

9.4.1.2 ADD_APPEND

```
include std/sequence.e
public enum ADD_APPEND
```

9.4.1.3 ADD_SORT_UP

```
include std/sequence.e
public enum ADD_SORT_UP
```

9.4.1.4 ADD_SORT_DOWN

```
include std/sequence.e
public enum ADD_SORT_DOWN
```

9.4.1.5 ROTATE_LEFT

```
include std/sequence.e
public constant ROTATE_LEFT
```

9.4.1.6 ROTATE_RIGHT

```
include std/sequence.e
public constant ROTATE_RIGHT
```

9.4.2 Basic routines

9.4.2.1 can_add

```
include std/sequence.e
public function can_add(object a, object b)
```

Checks whether two objects can be legally added together.

9.4.2.1.1 Parameters:

1. a : one of the objects to test for compatible shape
2. b : the other object

9.4.2.1.2 Returns:

An **integer**, 1 if an addition (or any of the [Relational operators](#)) are possible between a and b, else 0.

9.4.2.1.3 Example 1:

```
i = can_add({1,2,3},{4,5})
-- i is 0

i = can_add({1,2,3},4)
-- i is 1

i = can_add({1,2,3},{4,{5,6},7})
-- i is 1
```

9.4.2.1.4 See Also:

[linear](#)

9.4.2.2 fetch

```
include std/sequence.e
public function fetch(sequence source, sequence indexes)
```

Retrieves an element nested arbitrarily deep into a sequence.

9.4.2.2.1 Parameters:

1. source : the sequence from which to fetch
2. indexes : a sequence of integers, the path to follow to reach the element to return.

9.4.2.2.2 Returns:

An **object**, which is `source[indexes[1]][indexes[2]]...[indexes[$]]`

9.4.2.2.3 Errors:

If the path cannot be followed to its end, an error about reading a nonexistent element, or subscripting an atom, will occur.

9.4.2.2.4 Comments:

The last element of `indexes` may be a pair `{lower,upper}`, in which case a slice of the innermost referenced sequence is returned.

9.4.2.2.5 Example 1:

```
x = fetch({0,1,2,3, {"abc", "def", "ghi"}}, 6), {5,2,3})
-- x is 'f', or 102.
```

9.4.2.2.6 See Also:

[store](#), [Subscripting of Sequences](#)

9.4.2.3 store

```
include std/sequence.e
public function store(sequence target, sequence indexes, object x)
```

Stores something at a location nested arbitrarily deep into a sequence.

9.4.2.3.1 Parameters:

1. `target` : the sequence in which to store something
2. `indexes` : a sequence of integers, the path to follow to reach the place where to store
3. `x` : the object to store.

9.4.2.3.2 Returns:

A **sequence**, a **copy** of `target` with the specified place `indexes` modified by storing `x` into it.

9.4.2.3.3 Errors:

If the path to storage location cannot be followed to its end, or an index is not what one would expect or is not valid, an error about illegal sequence operations will occur.

9.4.2.3.4 Comments:

If the last element of `indexes` is a pair of integers, `x` will be stored as a slice three, the bounding indexes being given in the pair as `{lower,upper}`..

In EUPHORIA, you can never modify an object by passing it to a routine. You have to get a modified copy and then assign it back to the original.

9.4.2.3.5 Example 1:

```
s = store({0,1,2,3,{ "abc", "def", "ghi"}, 6},{5,2,3},108)
-- s is {0,1,2,3,{ "abc", "del", "ghi"}, 6}
```

9.4.2.3.6 See Also:

[fetch](#), [Subscripting of Sequences](#)

9.4.2.4 `valid_index`

```
include std/sequence.e
public function valid_index(sequence st, object x)
```

Checks whether an index exists on a sequence.

9.4.2.4.1 Parameters:

1. `s` : the sequence for which to check
2. `x` : an object, the index to check.

9.4.2.4.2 Returns:

An **integer**, 1 if `s[x]` makes sense, else 0.

9.4.2.4.3 Example 1:

```
i = valid_index({51,27,33,14},2)
-- i is 1
```

9.4.2.4.4 See Also:

[Subscripting of Sequences](#)

9.4.2.5 rotate

```
include std/sequence.e
public function rotate(sequence source, integer shift, integer start = 1, integer stop = length
```

Rotates a slice of a sequence.

9.4.2.5.1 Parameters:

1. `source` : sequence to be rotated
2. `shift` : direction and count to be shifted (ROTATE_LEFT or ROTATE_RIGHT)
3. `start` : starting position for shift, defaults to 1
4. `stop` : stopping position for shift, defaults to length (source)

9.4.2.5.2 Comments:

Use `amount * direction` to specify the shift. `direction` is either ROTATE_LEFT or ROTATE_RIGHT. This enables to shift multiple places in a single call. For instance, use ROTATE_LEFT * 5 to rotate left, 5 positions.

A null shift does nothing and returns source unchanged.

9.4.2.5.3 Example 1:

```
s = rotate({1, 2, 3, 4, 5}, ROTATE_LEFT)
-- s is {2, 3, 4, 5, 1}
```

9.4.2.5.4 Example 2:

```
s = rotate({1, 2, 3, 4, 5}, ROTATE_RIGHT * 2)
-- s is {4, 5, 1, 2, 3}
```

9.4.2.5.5 Example 3:

```
s = rotate({11,13,15,17,19,23}, ROTATE_LEFT, 2, 5)
-- s is {11,15,17,19,13,23}
```

9.4.2.5.6 Example 4:

```
s = rotate({11,13,15,17,19,23}, ROTATE_RIGHT, 2, 5)
-- s is {11,19,13,15,17,23}
```

9.4.2.5.7 See Also:

[slice](#), [head](#), [tail](#)

9.4.2.6 columnize

```
include std/sequence.e
public function columnize(sequence source, object cols = {}, object defval = 0)
```

Converts a set of sub sequences into a set of 'columns'.

9.4.2.6.1 Parameters:

1. `source` : sequence containing the sub-sequences
2. `cols` : either a specific column number or a set of column numbers. Default is 0, which returns the maximum number of columns.
3. `defval` : an object. Used when a column value is not available. Default is 0

9.4.2.6.2 Comments:

Any atoms found in `source` are treated as if they are a 1-element sequence.

9.4.2.6.3 Example 1:

```
s = columnize({{1, 2}, {3, 4}, {5, 6}})
-- s is { {1,3,5}, {2,4,6} }
```

9.4.2.6.4 Example 2:

```
s = columnize({{1, 2}, {3, 4}, {5, 6, 7}})
-- s is { {1,3,5}, {2,4,6}, {0,0,7} }
s = columnize({{1, 2}, {3, 4}, {5, 6, 7}}, -999) -- Change the not-available value.
-- s is { {1,3,5}, {2,4,6}, {-999,-999,7} }
```

9.4.2.6.5 Example 3:

```
s = columnize({{1, 2}, {3, 4}, {5, 6, 7}}, 2)
-- s is { {2,4,6} } -- Column 2 only
```

9.4.2.6.6 Example 4:

```
s = columnize({{1, 2}, {3, 4}, {5, 6, 7}}, {2,1})
-- s is { {2,4,6}, {1,3,5} } -- Column 2 then column 1
```

9.4.2.6.7 Example 5:

```
s = columnize({"abc", "def", "ghi"})
-- s is {"adg", "beh", "cfi" }
```

9.4.2.7 apply

```
include std/sequence.e
public function apply(sequence source, integer rid, object userdata = {})
```

Apply a function to every element of a sequence returning a new sequence of the same size.

9.4.2.7.1 Parameters:

- `source` : the sequence to map
- `rid` : the [routine_id](#) of function to use as converter
- `userdata` : an object passed to each invocation of `rid`. If omitted, `{}` is used.

9.4.2.7.2 Returns:

A **sequence**, the length of `source`. Each element there is the corresponding element in `source` mapped using the routine referred to by `rid`.

9.4.2.7.3 Comments:

The supplied routine must take two parameters. The type of the first parameter must be compatible with all the elements in `source`. The second parameter is an object containing `userdata`.

9.4.2.7.4 Example 1:

```
function greeter(object o, object d)
    return o[1] & ", " & o[2] & d
end function

s = apply({"Hello", "John"}, {"Goodbye", "John"}, routine_id("greeter"), "!")
-- s is {"Hello, John!", "Goodbye, John!"}
```

9.4.2.7.5 See Also:

[filter](#)

9.4.2.8 mapping

```
include std/sequence.e
public function mapping(object source_arg, sequence from_set, sequence to_set, integer one_level)
```

Each item from `source_arg` found in `from_set` is changed into the corresponding item in `to_set`

9.4.2.8.1 Parameters:

1. `source_arg` : Any EUPHORIA object to be transformed.
2. `from_set` : A sequence of objects representing the only items from `source_arg` that are actually transformed.
3. `to_set` : A sequence of objects representing the transformed equivalents of those found in `from_set`.
4. `one_level` : An integer. 0 (the default) means that mapping applies to every atom in every level of sub-sequences. 1 means that mapping only applies to the items at the first level in `source_arg`.

9.4.2.8.2 Returns:

An **object**, The transformed version of `source_arg`.

9.4.2.8.3 Comments:

- When `one_level` is zero or omitted, for each item in `source_arg`,
 - ◆ if it is an atom then it may be transformed
 - ◆ if it is a sequence, then the mapping is performed recursively on the sequence.
 - ◆ This option required `from_set` to only contain atoms and contain no sub-sequences.
- When `one_level` is not zero, for each item in `source_arg`,
 - ◆ regardless of whether it is an atom or sequence, if it is found in `from_set` then it is mapped to the corresponding object in `to_set`..
- Mapping occurs when an item in `source_arg` is found in `from_set`, then it is replaced by the corresponding object in `to_set`.

9.4.2.8.4 Example 1:

```
res = mapping("The Cat in the Hat", "aeiou", "AEIOU")
-- res is now "ThE CAT In thE HAt"
```


9.4.2.9 length

<built-in> `function length(sequence target)`

Return the length of a sequence.

9.4.2.9.1 Parameters:

1. `target` : the sequence being queried

9.4.2.9.2 Returns:

An **integer**, the number of elements `target` has.

9.4.2.9.3 Comments:

The length of each sequence is stored internally by the interpreter for fast access. In some other languages this operation requires a search through memory for an end marker.

9.4.2.9.4 Example 1:

```
length({{1,2}, {3,4}, {5,6}})  -- 3
length("") -- 0
length({}) -- 0
```

9.4.2.9.5 See Also:

[append](#), [prepend](#), &

9.4.2.10 reverse

```
include std/sequence.e
public function reverse(object target, integer pFrom = 1, integer pTo = 0)
```

Reverse the order of elements in a sequence.

9.4.2.10.1 Parameters:

1. `target` : the sequence to reverse.
2. `pFrom` : an integer, the starting point. Defaults to 1.
3. `pTo` : an integer, the end point. Defaults to 0.

9.4.2.10.2 Returns:

A **sequence**, if `target` is a sequence, the same length as `target` and the same elements, but those with index between `pFrom` and `pTo` appear in reverse order.

9.4.2.10.3 Comments:

In the result sequence, some or all top-level elements appear in reverse order compared to the original sequence. This does not reverse any sub-sequences found in the original sequence.

The `pTo` parameter can be negative, which indicates an offset from the last element. Thus -1 means the second-last element and 0 means the last element.

9.4.2.10.4 Example 1:

```
reverse({1,3,5,7})           -- {7,5,3,1}
reverse({1,3,5,7,9}, 2, -1) -- {1,7,5,3,9}
reverse({1,3,5,7,9}, 2)     -- {1,9,7,5,3}
reverse({{1,2,3}, {4,5,6}}) -- {{4,5,6}, {1,2,3}}
reverse({99})               -- {99}
reverse({})                 -- {}
reverse(42)                 -- 42
```

9.4.2.11 shuffle

```
include std/sequence.e
public function shuffle(sequence seq)
```

Shuffle the elements of a sequence.

9.4.2.11.1 Parameters:

1. `seq`: the sequence to shuffle.

9.4.2.11.2 Returns:

A **sequence**

9.4.2.11.3 Comments:

The input sequence does not have to be in any specific order and can contain duplicates. The output will be in an unpredictable order, which might even be the same as the input order.

9.4.2.11.4 Example 1:

```
shuffle({1,2,3,3}) -- {3,1,3,2}
shuffle({1,2,3,3}) -- {2,3,1,3}
shuffle({1,2,3,3}) -- {1,2,3,3}
```

9.4.3 Building sequences

9.4.3.1 linear

```
include std/sequence.e
public function linear(object start, object increment, integer count)
```

Returns a sequence in arithmetic progression.

9.4.3.1.1 Parameters:

1. `start` : the initial value from which to start
2. `increment` : the value to recursively add to `start` to get new elements
3. `count` : an integer, the number of additions to perform.

9.4.3.1.2 Returns:

An **object**, either 0 on failure or `{start, start+increment, ..., start+count*increment}`

9.4.3.1.3 Comments:

If `count` is negative, or if adding `start` to `increment` would prove to be impossible, then 0 is returned. Otherwise, a sequence, of length `count+1`, starting with `start` and whose adjacent elements differ exactly by `increment`, is returned.

9.4.3.1.4 Example 1:

```
s = linear({1,2,3},4,3)
-- s is {{1,2,3},{5,6,7},{9,10,11}}
```

9.4.3.1.5 See Also:

[repeat_pattern](#)

9.4.3.2 repeat_pattern

```
include std/sequence.e
public function repeat_pattern(sequence pattern, integer count)
```

Returns a periodic sequence, given a pattern and a count.

9.4.3.2.1 Parameters:

1. `pattern` : the sequence whose elements are to be repeated
2. `count` : an integer, the number of times the pattern is to be repeated.

9.4.3.2.2 Returns:

A **sequence**, empty on failure, and of length `count*length(pattern)` otherwise. The first elements of the returned sequence are those of `pattern`. So are those that follow, on to the end.

9.4.3.2.3 Example 1:

```
s = repeat_pattern({1,2,5},3)
-- s is {1,2,5,1,2,5,1,2,5}
```

9.4.3.2.4 See Also:

[repeat](#), [linear](#)

9.4.3.3 repeat

```
<built-in> function repeat(object item, atom count)
```

Create a sequence whose all elements are identical, with given length.

9.4.3.3.1 Parameters:

1. `item` : an object, to which all elements of the result will be equal
2. `count` : an atom, the requested length of the result sequence. This must be a value from zero to 0x3FFFFFFF. Any floating point values are first floored.

9.4.3.3.2 Returns:

A **sequence**, of length `count` each element of which is `item`.

9.4.3.3.3 Errors:

`count` cannot be less than zero and cannot be greater than 1,073,741,823.

9.4.3.3.4 Comments:

When you `repeat()` a sequence or a floating-point number the interpreter does not actually make multiple copies in memory. Rather, a single copy is "pointed to" a number of times.

9.4.3.3.5 Example 1:

```
repeat(0, 10) -- {0,0,0,0,0,0,0,0,0,0}

repeat("JOHN", 4) -- {"JOHN", "JOHN", "JOHN", "JOHN"}
-- The interpreter will create only one copy of "JOHN"
-- in memory and create a sequence containing four references to it.
```

9.4.3.3.6 See Also:

[repeat_pattern](#), [linear](#)

9.4.4 Adding to sequences

9.4.4.1 append

```
<built-in> function append(sequence target, object x)
```

Adds an object as the last element of a sequence.

9.4.4.1.1 Parameters:

1. `source` : the sequence to add to
2. `x` : the object to add

9.4.4.1.2 Returns:

A **sequence**, whose first elements are those of `target` and whose last element is `x`.

9.4.4.1.3 Comments:

The length of the resulting sequence will be `length(target) + 1`, no matter what `x` is.

If `x` is an atom this is equivalent to `result = target & x`. If `x` is a sequence it is not equivalent.

The extra storage is allocated automatically and very efficiently with EUPHORIA's dynamic storage allocation. The case where `target` itself is `append()`ed to (as in Example 1 below) is highly optimized.

9.4.4.1.4 Example 1:

```
sequence x

x = {}
for i = 1 to 10 do
    x = append(x, i)
end for
-- x is now {1,2,3,4,5,6,7,8,9,10}
```

9.4.4.1.5 Example 2:

```
sequence x, y, z

x = {"fred", "barney"}
y = append(x, "wilma")
-- y is now {"fred", "barney", "wilma"}

z = append(append(y, "betty"), {"bam", "bam"})
-- z is now {"fred", "barney", "wilma", "betty", {"bam", "bam"}}
```

9.4.4.1.6 See Also:

[prepend](#), [&](#)

9.4.4.2 prepend

```
<built-in> function prepend(sequence target, object x)
```

Adds an object as the first element of a sequence.

9.4.4.2.1 Parameters:

1. `source` : the sequence to add to
2. `x` : the object to add

9.4.4.2.2 Returns:

A **sequence**, whose last elements are those of `target` and whose first element is `x`.

9.4.4.2.3 Comments:

The length of the returned sequence will be `length(target) + 1` always.

If `x` is an atom this is the same as `result = x & target`. If `x` is a sequence it is not the same.

The case where `target` itself is `prepend()`ed to is handled very efficiently.

9.4.4.2.4 Example 1:

```
prepend({1,2,3}, {0,0})-- {{0,0}, 1, 2, 3}
-- Compare with concatenation:
{0,0} & {1,2,3}-- {0, 0, 1, 2, 3}
```

9.4.4.2.5 Example 2:

```
s = {}
for i = 1 to 10 do
    s = prepend(s, i)
end for
-- s is {10,9,8,7,6,5,4,3,2,1}
```

9.4.4.2.6 See Also:

[append](#), [&](#)

9.4.4.3 insert

```
<built-in> function insert(sequence target, object what, integer index)
```

Insert an object into a sequence as a new element at a given location.

9.4.4.3.1 Parameters:

1. `target` : the sequence to insert into
2. `what` : the object to insert
3. `index` : an integer, the position in `target` where `what` should appear

9.4.4.3.2 Returns:

A **sequence**, which is `target` with one more element at `index`, which is `what`.

9.4.4.3.3 Comments:

`target` can be a sequence of any shape, and `what` any kind of object.

The length of the returned sequence is `length(target)+1` always.

`insert()`ing a sequence into a string returns a sequence which is no longer a string.

9.4.4.3.4 Example 1:

```
s = insert("John Doe", " Middle", 5)
-- s is {'J','o','h','n'," Middle ","D','o','e'}
```

9.4.4.3.5 Example 2:

```
s = insert({10,30,40}, 20, 2)
-- s is {10,20,30,40}
```

9.4.4.3.6 See Also:

[remove](#), [splice](#), [append](#), [prepend](#)

9.4.4.4 splice

<built-in> `function splice(sequence target, object what, integer index)`

Inserts an object as a new slice in a sequence at a given position.

9.4.4.4.1 Parameters:

1. `target` : the sequence to insert into
2. `what` : the object to insert
3. `index` : an integer, the position in `target` where `what` should appear

9.4.4.4.2 Returns:

A **sequence**, which is `target` with one or more elements, those of `what`, inserted at locations starting at `index`.

9.4.4.4.3 Comments:

`target` can be a sequence of any shape, and `what` any kind of object.

The length of this new sequence is the sum of the lengths of `target` and `what` (atoms are of length 1 for this purpose). `splice()` is equivalent to `insert()` when `what` is an atom, but not when it is a sequence.

Splicing a string into a string results into a new string.

9.4.4.4.4 Example 1:

```
s = splice("John Doe", " Middle", 5)
-- s is "John Middle Doe"
```

9.4.4.4.5 Example 2:

```
s = splice({10,30,40}, 20, 2)
-- s is {10,20,30,40}
```

9.4.4.4.6 See Also:

[insert](#), [remove](#), [replace](#), &

9.4.4.5 pad_head

```
include std/sequence.e
public function pad_head(sequence target, integer size, object ch = ' ')
```

Pad the beginning of a sequence with an object so as to meet a minimum length condition.

9.4.4.5.1 Parameters:

1. `target` : the sequence to pad.
2. `size` : an integer, the target minimum size for `target`
3. `padding` : an object, usually the character to pad to (defaults to ' ').

9.4.4.5.2 Returns:

A **sequence**, either `target` if it was long enough, or a sequence of length `size` whose last elements are those of `target` and whose first few head elements all equal padding.

9.4.4.5.3 Comments:

`pad_head()` will not remove characters. If `length(target)` is greater than `size`, this function simply returns `target`. See [head\(\)](#) if you wish to truncate long sequences.

9.4.4.5.4 Example 1:

```
s = pad_head("ABC", 6)
-- s is "    ABC"

s = pad_head("ABC", 6, '-')
-- s is "---ABC"
```

9.4.4.5.5 See Also:

[trim_head](#), [pad_tail](#), [head](#)

9.4.4.6 pad_tail

```
include std/sequence.e
public function pad_tail(sequence target, integer size, object ch = ' ')
```

Pad the end of a sequence with an object so as to meet a minimum length condition.

9.4.4.6.1 Parameters:

1. `target` : the sequence to pad.
2. `size` : an integer, the target minimum size for `target`
3. `padding` : an object, usually the character to pad to (defaults to `' '`).

9.4.4.6.2 Returns:

A **sequence**, either `target` if it was long enough, or a sequence of length `size` whose first elements are those of `target` and whose last few head elements all equal padding.

9.4.4.6.3 Comments:

`pad_tail()` will not remove characters. If `length(target)` is greater than `size`, this function simply returns `target`. See [tail\(\)](#) if you wish to truncate long sequences.

9.4.4.6.4 Comments:

`pad_tail()` will not remove characters. If `length(str)` is greater than `params`, this function simply returns `str`. see `tail()` if you wish to truncate long sequences.

9.4.4.6.5 Example 1:

```
s = pad_tail("ABC", 6)
-- s is "ABC   "

s = pad_tail("ABC", 6, '-')
-- s is "ABC---"
```

9.4.4.6.6 See Also:

[trim_tail](#), [pad_head](#), [tail](#)

9.4.4.7 add_item

```
include std/sequence.e
public function add_item(object needle, sequence haystack, integer pOrder = 1)
```

Adds an item to the sequence if its not already there. If it already exists in the list, the list is returned unchanged.

9.4.4.7.1 Parameters:

1. `needle` : object to add.
2. `haystack` : sequence to add it to.
3. `order` : an integer; determines how the `needle` affects the `haystack`. It can be added to the front (prepended), to the back (appended), or sorted after adding. The default is to prepend it.

9.4.4.7.2 Returns:

A **sequence**, which is `haystack` with `needle` added to it.

9.4.4.7.3 Comments:

An error occurs if an invalid `order` argument is supplied.

The following enum is provided for specifying `order`:

- `ADD_PREPEND` -- prepend `needle` to `haystack`. This is the default option.
- `ADD_APPEND` -- append `needle` to `haystack`.

- ADD_SORT_UP -- sort haystack in ascending order after inserting needle
- ADD_SORT_DOWN -- sort haystack in descending order after inserting needle

9.4.4.7.4 Example 1:

```
s = add_item( 1, {3,4,2}, ADD_PREPEND ) -- prepend
-- s is {1,3,4,2}
```

9.4.4.7.5 Example 2:

```
s = add_item( 1, {3,4,2}, ADD_APPEND ) -- append
-- s is {3,4,2,1}
```

9.4.4.7.6 Example 3:

```
s = add_item( 1, {3,4,2}, ADD_SORT_UP ) -- ascending
-- s is {1,2,3,4}
```

9.4.4.7.7 Example 4:

```
s = add_item( 1, {3,4,2}, ADD_SORT_DOWN ) -- descending
-- s is {4,3,2,1}
```

9.4.4.7.8 Example 5:

```
s = add_item( 1, {3,1,4,2} )
-- s is {3,1,4,2} -- Item was already in list so no change.
```

9.4.4.8 remove_item

```
include std/sequence.e
public function remove_item(object needle, sequence haystack)
```

Removes an item from the sequence.

9.4.4.8.1 Parameters:

1. needle : object to remove.
2. haystack : sequence to remove it from.

9.4.4.8.2 Returns:

A **sequence**, which is `haystack` with `needle` removed from it.

9.4.4.8.3 Comments:

If `needle` is not in `haystack` then `haystack` is returned unchanged.

9.4.4.8.4 Example 1:

```
s = remove_item( 1, {3,4,2,1} ) --> {3,4,2}
s = remove_item( 5, {3,4,2,1} ) --> {3,4,2,1}
```

9.4.5 Extracting, removing, replacing from/into a sequence

9.4.5.1 head

```
<built-in> function head(sequence source, atom size=1)
```

Return the first item(s) of a sequence.

9.4.5.1.1 Parameters:

1. `source` : the sequence from which elements will be returned
2. `size` : an integer, how many head elements at most will be returned. Defaults to 1.

9.4.5.1.2 Returns:

A **sequence**, `source` if its length is not greater than `size`, or the `size` first elements of `source` otherwise.

9.4.5.1.3 Example 1:

```
s2 = head("John Doe", 4)
-- s2 is John
```

9.4.5.1.4 Example 2:

```
s2 = head("John Doe", 50)
-- s2 is John Doe
```

9.4.5.1.5 Example 3:

```
s2 = head({1, 5.4, "John", 30}, 3)
-- s2 is {1, 5.4, "John"}
```

9.4.5.1.6 See Also:

[tail](#), [mid](#), [slice](#)

9.4.5.2 tail

```
<built-in> function tail(sequence source, atom n=length(source) - 1)
```

Return the last items of a sequence.

9.4.5.2.1 Parameters:

1. `source` : the sequence to get the tail of.
2. `size` : an integer, the number of items to return. (defaults to `length(source) - 1`)

9.4.5.2.2 Returns:

A **sequence**, of length at most `size`. If the length is less than `size`, then `source` was returned. Otherwise, the `size` last elements of `source` were returned.

9.4.5.2.3 Comments:

`source` can be any type of sequence, including nested sequences.

9.4.5.2.4 Example 1:

```
s2 = tail("John Doe", 3)
-- s2 is "Doe"
```

9.4.5.2.5 Example 2:

```
s2 = tail("John Doe", 50)
-- s2 is "John Doe"
```

9.4.5.2.6 Example 3:

```
s2 = tail({1, 5.4, "John", 30}, 3)
-- s2 is {5.4, "John", 30}
```

9.4.5.2.7 See Also:

[head](#), [mid](#), [slice](#)

9.4.5.3 mid

```
include std/sequence.e
public function mid(sequence source, atom start, atom len)
```

Returns a slice of a sequence, given by a starting point and a length.

9.4.5.3.1 Parameters:

1. `source` : the sequence some elements of which will be returned
2. `start` : an integer, the lower index of the slice to return
3. `len` : an integer, the length of the slice to return

9.4.5.3.2 Returns:

A **sequence**, made of at most `len` elements of `source`. These elements are at contiguous positions in `source` starting at `start`.

9.4.5.3.3 Errors:

If `len` is less than `-length(source)`, an error occurs.

9.4.5.3.4 Comments:

`len` may be negative, in which case it is added `length(source)` once.

9.4.5.3.5 Example 1:

```
s2 = mid("John Middle Doe", 6, 6)
-- s2 is Middle
```

9.4.5.3.6 Example 2:

```
s2 = mid("John Middle Doe", 6, 50)
-- s2 is Middle Doe
```

9.4.5.3.7 Example 3:

```
s2 = mid({1, 5.4, "John", 30}, 2, 2)
-- s2 is {5.4, "John"}
```

9.4.5.3.8 See Also:

[head](#), [tail](#), [slice](#)

9.4.5.4 slice

```
include std/sequence.e
public function slice(sequence source, atom start = 1, atom stop = 0)
```

Return a portion of the supplied sequence.

9.4.5.4.1 Parameters:

1. `source` : the sequence from which to get a portion
2. `start` : an integer, the starting point of the portion. Default is 1.
3. `stop` : an integer, the ending point of the portion. Default is `length(source)`.

9.4.5.4.2 Returns:

A [sequence](#).

9.4.5.4.3 Comments:

- If the supplied `start` is less than 1 then it set to 1.
- If the supplied `stop` is less than 1 then `length(source)` is added to it. In this way, 0 represents the end of `source`, -1 represents one element in from the end of `source` and so on.
- If the supplied `stop` is greater than `length(source)` then it is set to the end.
- After these adjustments, and if `source[start..stop]` makes sense, it is returned, otherwise, `{ }` is returned.

9.4.5.4.4 Examples:

```
s2 = slice("John Doe", 6, 8) --> "Doe"
s2 = slice("John Doe", 6, 50) --> "Doe"
s2 = slice({1, 5.4, "John", 30}, 2, 3) --> {5.4, "John"}
s2 = slice({1,2,3,4,5}, 2, -1) --> {2,3,4}
s2 = slice({1,2,3,4,5}, 2) --> {2,3,4,5}
s2 = slice({1,2,3,4,5}, , 4) --> {1,2,3,4}
```

9.4.5.4.5 See Also:

[head](#), [mid](#), [tail](#)

9.4.5.5 vslice

```
include std/sequence.e
public function vslice(sequence source, atom colno, object error_control = 0)
```

Perform a vertical slice on a nested sequence

9.4.5.5.1 Parameters:

1. `source` : the sequence to take a vertical slice from
2. `colno` : an atom, the column number to extract (rounded down)
3. `error_control` : an object which says what to do if some element does not exist. Defaults to 0 (crash in such a circumstance).

9.4.5.5.2 Returns:

A **sequence**, usually of the same length as `source`, made of all the `source[x][colno]`.

9.4.5.5.3 Errors:

If an element is not defined and `error_control` is 0, an error occurs. If `colno` is less than 1, it cannot be any valid column, and an error occurs.

9.4.5.5.4 Comments:

If it is not possible to return the sequence of all `source[x][colno]` for all available `x`, the outcome is decided by `error_control`:

- If 0 (the default), program is aborted.
- If a nonzero atom, the short vertical slice is returned.

- Otherwise, elements of `error_control` will be taken to make for any missing element. A short vertical slice is returned if `error_control` is exhausted.

9.4.5.5.5 Example 1:

```
s = vslice({{5,1}, {5,2}, {5,3}}, 2)
-- s is {1,2,3}

s = vslice({{5,1}, {5,2}, {5,3}}, 1)
-- s is {5,5,5}
```

9.4.5.5.6 See Also:

[slice](#), [project](#)

9.4.5.6 remove

```
<built-in> function remove(sequence target, atom start, atom stop=start)
```

Remove an item, or a range of items from a sequence.

9.4.5.6.1 Parameters:

1. `target` : the sequence to remove from.
2. `start` : an atom, the (starting) index at which to remove
3. `stop` : an atom, the index at which to stop removing (defaults to `start`)

9.4.5.6.2 Returns:

A **sequence**, obtained from `target` by carving the `start .. stop` slice out of it.

9.4.5.6.3 Comments:

A new sequence is created. `target` can be a string or complex sequence.

9.4.5.6.4 Example 1:

```
s = remove("Johnn Doe", 4)
-- s is "John Doe"
```

9.4.5.6.5 Example 2:

```
s = remove({1,2,3,3,4}, 4)
-- s is {1,2,3,4}
```

9.4.5.6.6 Example 3:

```
s = remove("John Middle Doe", 6, 12)
-- s is "John Doe"
```

9.4.5.6.7 Example 4:

```
s = remove({1,2,3,3,4,4}, 4, 5)
-- s is {1,2,3,4}
```

9.4.5.6.8 See Also:

[replace](#), [insert](#), [splice](#), [remove_all](#)

9.4.5.7 patch

```
include std/sequence.e
public function patch(sequence target, sequence source, integer start, object filler = '')
```

Changes a sequence slice, possibly with padding

9.4.5.7.1 Parameters:

1. `target` : a sequence, a modified copy of which will be returned
2. `source` : a sequence, to be patched inside or outside `target`
3. `start` : an integer, the position at which to patch
4. `filler` : an object, used for filling gaps. Defaults to ''

9.4.5.7.2 Returns:

A **sequence**, which looks like `target`, but a slice starting at `start` equals `source`.

9.4.5.7.3 Comments:

In some cases, this call will result in the same result as [replace\(\)](#).

If `source` doesn't fit into `target` because of the lengths and the supplied `start` value, gaps will be created, and `filler` is used to fill them in.

Notionally, `target` has an infinite amount of `filler` on both sides, and `start` counts position relative to where `target` actually starts. Then, notionally, a `replace()` operation is performed.

9.4.5.7.4 Example 1:

```
sequence source = "abc", target = "John Doe"
sequence s = patch(target, source, 11, '0')
-- s is now "John Doe00abc"
```

9.4.5.7.5 Example 2:

```
sequence source = "abc", target = "John Doe"
sequence s = patch(target, source, -1)
-- s is now "abcohn Doe"
```

Note that there was no gap to fill.

Since $-1 = 1 - 2$, the patching started 2 positions before the initial 'J'.

9.4.5.7.6 Example 3:

```
sequence source = "abc", target = "John Doe"
sequence s = patch(target, source, 6)
-- s is now "John Dabc"
```

9.4.5.7.7 See Also:

[mid](#), [replace](#)

9.4.5.8 `remove_all`

```
include std/sequence.e
public function remove_all(object needle, sequence haystack)
```

Removes all occurrences of some object from a sequence.

9.4.5.8.1 Parameters:

1. `needle` : the object to remove.
2. `haystack` : the sequence to remove from.

9.4.5.8.2 Returns:

A **sequence**, of length at most `length(haystack)`, and which has the same elements, without any copy of `needle` left

9.4.5.8.3 Comments:

This function weeds elements out, not sub-sequences.

9.4.5.8.4 Example 1:

```
s = remove_all( 1, {1,2,4,1,3,2,4,1,2,3} )
-- s is {2,4,3,2,4,2,3}
```

9.4.5.8.5 Example 2:

```
s = remove_all("x", "I'm toox secxksy for my shixrt.")
-- s is "I'm too secksy for my shirt."
```

9.4.5.8.6 See Also:

[remove](#), [replace](#)

9.4.5.9 retain_all

```
include std/sequence.e
public function retain_all(object needles, sequence haystack)
```

Keeps all occurrences of a set of objects from a sequence and removes all others.

9.4.5.9.1 Parameters:

1. `needles` : the set of objects to retain.
2. `haystack` : the sequence to remove items not in `needles`.

9.4.5.9.2 Returns:

A **sequence** containing only those objects from `haystack` that are also in `needles`.

9.4.5.9.3 Example:

```
s = retain_all( {1,3,5}, {1,2,4,1,3,2,4,1,2,3} ) --> {1,1,3,1,3}
s = retain_all("0123456789", "+34 (04) 555-44392") -> "340455544392"
```

9.4.5.9.4 See Also:

[remove](#), [replace](#), [remove_all](#)

9.4.5.10 filter

```
include std/sequence.e
public function filter(sequence source, object rid, object userdata = {}, object rangetype = "")
```

Filter a sequence based on a user supplied comparator function.

9.4.5.10.1 Parameters:

- `source` : sequence to filter
- `rid` : Either a [routine_id](#) of function to use as comparator or one of the predefined comparitors.
- `userdata` : an object passed to each invocation of `rid`. If omitted, `{}` is used.
- `rangetype`: A sequence. Only used when `rid` is "in" or "out". This is used to let the function know how to interpret `userdata`. When `rangetype` is an empty string (which is the default), then `userdata` is treated as a set of zero or more discrete items such that "in" will only return items from `source` that are in the set of item in `userdata` and "out" returns those not in `userdata`. The other values for `rangetype` mean that `userdata` must be a set of exactly two items, that represent the lower and upper limits of a range of values.

9.4.5.10.2 Returns:

A **sequence**, made of the elements in `source` which passed the comparator test.

9.4.5.10.3 Comments:

- The only items from `source` that are returned are those that pass the test.
- When `rid` is a `routine_id`, that user defined routine must be a function. Each item in `source`, along with the `userdata` is passed to the function. The function must return a non-zero atom if the item is to be included in the result sequence, otherwise it should return zero to exclude it from the result.
- The predefined comparitors are...

"<" or "lt"	return items in <code>source</code> that are less than <code>userdata</code>
"<=" or "le"	return items in <code>source</code> that are less than or equal to <code>userdata</code>
"=" or "==" or "eq"	return items in <code>source</code> that are equal to <code>userdata</code>
"!=" or "ne"	return items in <code>source</code> that are not equal to <code>userdata</code>
">" or "gt"	return items in <code>source</code> that are greater than <code>userdata</code>
">=" or "ge"	return items in <code>source</code> that are greater than or equal to <code>userdata</code>
"in"	return items in <code>source</code> that are in <code>userdata</code>
"out"	return items in <code>source</code> that are not in <code>userdata</code>

- Range Type Usage

Range Type	Meaning
"["	Inclusive range. Lower and upper are in the range.
"["	Low Inclusive range. Lower is in the range but upper is not.
"]"	High Inclusive range. Lower is not in the range but upper is.
"("	Exclusive range. Lower and upper are not in the range.

9.4.5.10.4 Example 1:

```
function mask_nums(atom a, object t)
    if sequence(t) then
        return 0
    end if
    return and_bits(a, t) != 0
end function

function even_nums(atom a, atom t)
    return and_bits(a,1) = 0
end function

constant data = {5,8,20,19,3,2,10}
filter(data, routine_id("mask_nums"), 1) --> {5,19,3}
filter(data, routine_id("mask_nums"), 2) --> {19, 3, 2, 10}
filter(data, routine_id("even_nums")) --> {8, 20, 2, 10}

-- Using 'in' and 'out' with sets.
filter(data, "in", {3,4,5,6,7,8}) --> {5,8,3}
filter(data, "out", {3,4,5,6,7,8}) --> {20,19,2,10}

-- Using 'in' and 'out' with ranges.
filter(data, "in", {3,8}, "[") --> {5,8,3}
filter(data, "in", {3,8}, "[") --> {5,3}
filter(data, "in", {3,8}, "[") --> {5,8}
filter(data, "in", {3,8}, "(") --> {5}
filter(data, "out", {3,8}, "[") --> {20,19,2,10}
filter(data, "out", {3,8}, "[") --> {8,20,19,2,10}
filter(data, "out", {3,8}, "[") --> {20,19,3,2,10}
filter(data, "out", {3,8}, "(") --> {8,20,19,3,2,10}
```

9.4.5.10.5 Example 3:

```
function quicksort(sequence s)
    length(s) < 2 then
        return s
    end if
    quicksort( filter(s[2..$], "<=", s[1]) ) & s[1] & quicksort(filter(s[2..$], ">", s[1]))
end function
? quicksort( {5,4,7,2,4,9,1,0,4,32,7,54,2,5,8,445,67} )
--> {0,1,2,2,4,4,4,5,5,7,7,8,9,32,54,67,445}
```

9.4.5.10.6 See Also:

[apply](#)

9.4.5.11 STDFLTR_ALPHA

```
public constant STDFLTR_ALPHA
```

Predefined routine_id for use with [filter\(\)](#).

9.4.5.11.1 Comments:

Used to filter out non-alphabetic characters from a string.

9.4.5.11.2 Example:

```
-- Collect only the alphabetic characters from 'text'
result = filter(text, STDFLTR_ALPHA)
```

9.4.5.12 replace

```
<built-in> function replace(sequence target, object replacement, integer start, integer stop=st
```

Replace a slice in a sequence by an object.

9.4.5.12.1 Parameters:

1. `target` : the sequence in which replacement will be done.
2. `replacement` : an object, the item to replace with.
3. `start` : an integer, the starting index of the slice to replace.
4. `stop` : an integer, the stopping index of the slice to replace.

9.4.5.12.2 Returns:

A **sequence**, which is made of `target` with the `start`..`stop` slice removed and replaced by `replacement`, which is [splice\(\)](#)d in.

9.4.5.12.3 Comments:

- A new sequence is created. `target` can be a string or complex sequence of any shape.

- To replace by just one element, enclose replacement in curly braces, which will be removed at replace time.

9.4.5.12.4 Example 1:

```
s = replace("John Middle Doe", "Smith", 6, 11)
-- s is "John Smith Doe"

s = replace({45.3, "John", 5, {10, 20}}, 25, 2, 3)
-- s is {45.3, 25, {10, 20}}
```

9.4.5.12.5 See Also:

[splice](#), [remove](#), [remove_all](#)

9.4.5.13 replace_all

```
include std/sequence.e
public function replace_all(sequence source, object olddata, object newdata)
```

Replaces all occurrences of olddata with newdata

9.4.5.13.1 Parameters:

1. source : the sequence in which replacements will be done.
2. olddata : the sequence/item which is going to be replaced. If this is an empty sequence, the source is returned as is.
3. newdata : the sequence/item which will be the replacement.

9.4.5.13.2 Returns:

A **sequence**, which is made of source with all olddata occurrences replaced by newdata.

9.4.5.13.3 Comments:

This also removes all olddata occurrences when newdata is "".

9.4.5.13.4 Example:

```
s = replace_all("abracadabra", 'a', 'X')
-- s is now "XbrXcXdXbrX"
s = replace_all("abracadabra", "ra", 'X')
-- s is now "abXcadabX"
s = replace_all("abracadabra", "a", "aa")
```

```
-- s is now "aabracadaabraa"
s = replace_all("abracadabra", "a", "")
-- s is now "brcdbr"
```

9.4.5.13.5 See Also:

[replace](#), [remove_all](#)

9.4.5.14 extract

```
include std/sequence.e
public function extract(sequence source, sequence indexes)
```

Turns a sequences of indexes into the sequence of elements in a source that have such indexes.

9.4.5.14.1 Parameters:

1. `source` : the sequence from which to extract elements
2. `indexes` : a sequence of atoms, the indexes of the elements to be fetched in `source`.

9.4.5.14.2 Returns:

A **sequence**, of length at most `length(indexes)`. If `p` is the `r`-th element of `indexes` which is valid on `source`, then `result[r]` is `source[p]`.

9.4.5.14.3 Example 1:

```
s = extract({11,13,15,17},{3,1,2,1,4})
-- s is {15,11,13,11,17}
```

9.4.5.14.4 See Also:

[slice](#)

9.4.5.15 project

```
include std/sequence.e
public function project(sequence source, sequence coords)
```

Creates a list of sequences based on selected elements from sequences in the source.

9.4.5.15.1 Parameters:

1. `source` : a list of sequences.
2. `coords` : a list of index lists.

9.4.5.15.2 Returns:

A **sequence**, with the same length as `source`. Each of its elements is a sequence, the length of `coords`. Each innermost sequence is made of the elements from the corresponding source sub-sequence.

9.4.5.15.3 Comments:

For each sequence in `source`, a set of sub-sequences is created; one for each index list in `coords`. An index list is just a sequence containing indexes for items in a sequence.

9.4.5.15.4 Example 1:

```
s = project({ "ABCD", "789"}, {{1,2}, {3,1}, {2}})
-- s is {"AB", "CA", "B"}, {"78", "97", "8"}
```

9.4.5.15.5 See Also:

[vslice](#), [extract](#)

9.4.6 Changing the shape of a sequence

9.4.6.1 split

```
include std/sequence.e
public function split(sequence st, object delim = ' ', integer limit = 0, integer no_empty = 0)
```

Split a sequence on separator delimiters into a number of sub-sequences.

9.4.6.1.1 Parameters:

1. `source` : the sequence to split.
2. `delim` : an object (default is ' '). The delimiter that separates items in `source`.
3. `limit` : an integer (default is 0). The maximum number of sub-sequences to create. If zero, there is no limit.
4. `no_empty` : an integer (default is 0). If not zero then all zero-length sub-sequences are removed from the returned sequence. Use this when leading, trailing and duplicated delimiters are not significant.

9.4.6.1.2 Returns:

A **sequence**, of sub-sequences of `source`. Delimiters are removed.

9.4.6.1.3 Comments:

This function may be applied to a string sequence or a complex sequence.

If `limit` is > 0 , this is the maximum number of sub-sequences that will be created, otherwise there is no limit.

9.4.6.1.4 Example 1:

```
result = split("John Middle Doe")
-- result is {"John", "Middle", "Doe"}
```

9.4.6.1.5 Example 2:

```
result = split("John,Middle,Doe", ",", 2)
-- result is {"John", "Middle,Doe"}
```

9.4.6.1.6 See Also:

[split_any](#), [breakup](#), [join](#)

9.4.6.2 split_any

```
include std/sequence.e
public function split_any(sequence source, object delim, integer limit = 0, integer no_empty =
```

Split a sequence by any of the separators in the list of delimiters.

If `limit` is > 0 then limit the number of tokens that will be split to `limit`.

9.4.6.2.1 Parameters:

1. `source` : the sequence to split.
2. `delim` : a list of delimiters to split by.
3. `limit` : an integer (default is 0). The maximum number of sub-sequences to create. If zero, there is no limit.
4. `no_empty` : an integer (default is 0). If not zero then all zero-length sub-sequences removed from the returned sequence. Use this when leading, trailing and duplicated delimiters are not significant.

9.4.6.2.2 Comments:

This function may be applied to a string sequence or a complex sequence.

It works like `split()`, but in this case `delim` is a set of potential delimiters rather than a single delimiter.

9.4.6.2.3 Example 1:

```
result = split_any("One,Two|Three.Four", ".,|")
-- result is {"One", "Two", "Three", "Four"}
result = split_any(",One,,Two|.Three||.Four,", ".,|",,1) -- No Empty option
-- result is {"One", "Two", "Three", "Four"}
```

9.4.6.2.4 See Also:

[split](#), [breakup](#), [join](#)

9.4.6.3 join

```
include std/sequence.e
public function join(sequence items, object delim = " ")
```

Join sequences together using a delimiter.

9.4.6.3.1 Parameters:

1. `items`: the sequence of items to join.
2. `delim`: an object, the delimiter to join by. Defaults to " ".

9.4.6.3.2 Comments:

This function may be applied to a string sequence or a complex sequence

9.4.6.3.3 Example 1:

```
result = join({"John", "Middle", "Doe"})
-- result is "John Middle Doe"
```

9.4.6.3.4 Example 2:

```
result = join({"John", "Middle", "Doe"}, ",")
-- result is "John,Middle,Doe"
```

9.4.6.3.5 See Also:

[split](#), [split_any](#), [breakup](#)

9.4.6.4 BK_LEN

```
include std/sequence.e
public enum BK_LEN
```

Indicates that `size` parameter is maximum length of sub-sequence. See [breakup](#)

9.4.6.5 BK_PIECES

```
include std/sequence.e
public enum BK_PIECES
```

Indicates that `size` parameter is maximum number of sub-sequence. See [breakup](#)

9.4.6.6 breakup

```
include std/sequence.e
public function breakup(sequence source, object size, integer style = BK_LEN)
```

Breaks up a sequence into multiple sequences of a given length.

9.4.6.6.1 Parameters:

1. `source` : the sequence to be broken up into sub-sequences.
2. `size` : an object, if an integer it is either the maximum length of each resulting sub-sequence or the maximum number of sub-sequences to break `source` into.
If `size` is a sequence, it is a list of element counts for the sub-sequences it creates.
3. `style` : an integer, Either `BK_LEN` if `size` integer represents the sub-sequences' maximum length, or `BK_PIECES` if the `size` integer represents the maximum number of sub-sequences (pieces) to break `source` into.

9.4.6.6.2 Returns:

A **sequence**, of sequences.

9.4.6.6.3 Comments:

When **size** is an integer and **style** is **BK_LEN**...

The sub-sequences have length `size`, except possibly the last one, which may be shorter. For example if `source` has 11 items and `size` is 3, then the first three sub-sequences will get 3 items each and the remaining 2 items will go into the last sub-sequence. If `size` is less than 1 or greater than the length of the `source`, the `source` is returned as the only sub-sequence.

When **size** is an integer and **style** is **BK_PIECES**...

There is exactly `size` sub-sequences created. If the `source` is not evenly divisible into that many pieces, then the lefthand sub-sequences will contain one more element than the right-hand sub-sequences. For example, if `source` contains 10 items and we break it into 3 pieces, piece #1 gets 4 elements, piece #2 gets 3 items and piece #3 gets 3 items - a total of 10. If `source` had 11 elements then the pieces will have 4, 4, and 3 respectively.

When **size** is a sequence...

The `style` parameter is ignored in this case. The `source` will be broken up according to the counts contained in the `size` parameter. For example, if `size` was `{3,4,0,1}` then piece #1 gets 3 items, #2 gets 4 items, #3 gets 0 items, and #4 gets 1 item. Note that if not all items from `source` are placed into the sub-sequences defined by `size`, and *extra* sub-sequence is appended that contains the remaining items from `source`.

In all cases, when concatenated these sub-sequences will be identical to the original `source`.

9.4.6.6.4 Example 1:

```
s = breakup("5545112133234454", 4)
-- s is {"5545", "1121", "3323", "4454"}
```

9.4.6.6.5 Example 2:

```
s = breakup("12345", 2)
-- s is {"12", "34", "5"}
```

9.4.6.6.6 Example 3:

```
s = breakup({1,2,3,4,5,6}, 3)
-- s is {{1,2,3}, {4,5,6}}
```

9.4.6.6.7 Example 4:

```
s = breakup("ABCDEF", 0)
-- s is {"ABCDEF"}
```

9.4.6.6.8 See Also:

[split flatten](#)

9.4.6.7 flatten

```
include std/sequence.e
public function flatten(sequence s, object delim = "")
```

Remove all nesting from a sequence.

9.4.6.7.1 Parameters:

1. `s` : the sequence to flatten out.
2. `delim` : An optional delimiter to place after each flattened sub-sequence (except the last one).

9.4.6.7.2 Returns:

A **sequence**, of atoms, all the atoms in `s` enumerated.

9.4.6.7.3 Comments:

- If you consider a sequence as a tree, then the enumeration is performed by left-right reading of the tree. The elements are simply read left to right, without any care for braces.
- Empty sub-sequences are stripped out entirely.

9.4.6.7.4 Example 1:

```
s = flatten({{18, 19}, 45, {18.4, 29.3}})
-- s is {18, 19, 45, 18.4, 29.3}
```

9.4.6.7.5 Example 2:

```
s = flatten({18, { 19, {45}}, {18.4, {}, 29.3}})
-- s is {18, 19, 45, 18.4, 29.3}
```

9.4.6.7.6 Example 3:

```
Using the delimiter argument.
s = flatten({"abc", "def", "ghi"}, ", ")
-- s is "abc, def, ghi"
```


9.4.6.8 pivot

```
include std/sequence.e
public function pivot(object data_p, object pivot_p = 0)
```

Returns a sequence of three sub-sequences. The sub-sequences contain all the elements less than the supplied pivot value, equal to the pivot, and greater than the pivot.

9.4.6.8.1 Parameters:

1. data_p : Either an atom or a list. An atom is treated as if it is one-element sequence.
2. pivot_p : An object. Default is zero.

9.4.6.8.2 Returns:

A **sequence**, { {less than pivot}, {equal to pivot}, {greater than pivot} }

9.4.6.8.3 Comments:

pivot() is used as a split up a sequence relative to a specific value.

9.4.6.8.4 Example 1:

```
? pivot( {7, 2, 8.5, 6, 6, -4.8, 6, 6, 3.341, -8, "text"}, 6 )
-- Ans: {{2, -4.8, 3.341, -8}, {6, 6, 6, 6}, {7, 8.5, "text"}}
? pivot( {4, 1, -4, 6, -1, -7, 9, 10} )
-- Ans: {{-4, -1, -7}, {}, {4, 1, 6, 9, 10}}
? pivot( 5 )
-- Ans: {}, {}, {5}
```

9.4.6.8.5 Example 2:

```
function quicksort(sequence s)
    length(s) < 2 then
        s
    return
    end if
    sequence
        k(s, pivot and (length(s)))

    quicksort(k[1]) & k[2] & quicksort(k[3])
end function
sequence t2 = {5, 4, 7, 2, 4, 9, 1, 0, 4, 32, 7, 54, 2, 5, 8, 445, 67}
? quicksort(t2) --> {0, 1, 2, 2, 4, 4, 4, 5, 5, 7, 7, 8, 9, 32, 54, 67, 445}
```

9.4.6.9 build_list

```
include std/sequence.e
public function build_list(sequence source, object transformer, integer singleton = 1, object u
```

Implements "List Comprehension" or building a list based on the contents of another list.

9.4.6.9.1 Parameters:

1. `source` : A sequence. The list of items to base the new list upon.
2. `transformer` : One or more routine_ids. These are [routine ids](#) of functions that must receive three parameters (object x, sequence i, object u) where 'x' is an item in the `source` list, 'i' contains the position that 'x' is found in the `source` list and the length of `source`, and 'u' is the `user_data` value. Each transformer must return a two-element sequence. If the first element is zero, then `build_list()` continues on with the next transformer function for the same 'x'. If the first element is not zero, the second element is added to the new list being built (other elements are ignored) and `build_list` skips the rest of the transformers and processes the next element in `source`.
3. `singleton` : An integer. If zero then the transformer functions return multiple list elements. If not zero then the transformer functions return a single item (which might be a sequence).
4. `user_data` : Any object. This is passed unchanged to each transformer function.

9.4.6.9.2 Returns:

A **sequence**, The new list of items.

9.4.6.9.3 Comments:

- If the transformer is -1, then the source item is just copied.

9.4.6.9.4 Example 1:

```
function remitem(object x, sequence i, object q)
    (x < 0) if then
        {0} -- no return
    else
        {1,x} -- return x
    end if
end function

sequence s
-- Remove negative elements (x < 0)
s = build_list({-3, 0, 1.1, -2, 2, 3, -1.5}, routine_id("remitem"), , 0)
-- s is {0, 1.1, 2, 3}
```

9.4.6.10 transform

```
include std/sequence.e
public function transform(sequence source_data, object transformer_rids)
```

Transforms the input sequence by using one or more user-supplied transformers.

9.4.6.10.1 Parameters:

1. `source_data` : A sequence to be transformed.
2. `transformer_rids` : An object. One or more routine_ids used to transform the input.

9.4.6.10.2 Returns:

The source **sequence**, that has been transformed.

9.4.6.10.3 Comments:

- This works by calling each transformer in order, passing to it the result of the previous transformation. Of course, the first transformer gets the original sequence as passed to this routine.
- Each transformer routine takes one or more parameters. The first is a source sequence to be transformed and others are any user data that may have been supplied to the `transform` routine.
- Each transformer routine returns a transformed sequence.
- The `transformer_rids` parameters is either a single routine_id or a sequence of routine_ids. In this second case, the routine_id may actually be a multi-element sequence containing the real routine_id and some user data to pass to the transformer routine. If there is no user data then the transformer is called with only one parameter.

9.4.6.10.4 Examples:

```
res = transform(" hello    ", {
    {routine_id("trim"), " ", 0},
    routine_id("upper"),
    {routine_id("replace_all"), "O", "A"}
})
--> "HELLA"
```

9.4.6.11 sim_index

```
include std/sequence.e
public function sim_index(sequence A, sequence B)
```

Calculates the similarity between two sequences.

9.4.6.11.1 Parameters:

1. A : A sequence.
2. B : A sequence.

9.4.6.11.2 Returns:

An **atom**, the closer to zero, the more the two sequences are alike.

9.4.6.11.3 Comments:

The calculation is weighted to give mismatched elements towards the front of the sequences larger scores. This means that sequences that differ near the beginning are considered more un-alike than mismatches towards the end of the sequences. Also, unmatched elements from the first sequence are weighted more than unmatched elements from the second sequence.

Two identical sequences return zero. A non-zero means that they are not the same and larger values indicate a larger differences.

9.4.6.11.4 Example 1:

```
? sim_index("sit", "sin")      --> 0.08784
? sim_index("sit", "sat")      --> 0.32394
? sim_index("sit", "skit")     --> 0.34324
? sim_index("sit", "its")      --> 0.68293
? sim_index("sit", "kit")      --> 0.86603

? sim_index("knitting", "knitting") --> 0.00000
? sim_index("kitting", "kitten")   --> 0.09068
? sim_index("knitting", "knotting") --> 0.27717
? sim_index("knitting", "kitten")   --> 0.35332
? sim_index("abacus", "zoological") --> 0.76304
```

9.4.6.12 SEQ_NOALT

```
include std/sequence.e
public constant SEQ_NOALT
```

Indicates that `remove_subseq()` must not replace removed sub-sequences with an alternative value.

9.4.6.13 remove_subseq

```
include std/sequence.e
public function remove_subseq(sequence source_list, object alt_value = SEQ_NOALT)
```

Removes all sub-sequences from the supplied sequence, optionally replacing them with a supplied alternative value. One common use is to remove all strings from a mixed set of numbers and strings.

9.4.6.13.1 Parameters:

1. `source_list` : A sequence from which sub-sequences are removed.
2. `alt_value` : An object. The default is `SEQ_NOALT`, which causes sub-sequences to be physically removed, otherwise any other value will be used to replace the sub-sequence.

9.4.6.13.2 Returns:

A **sequence**, which contains only the atoms from `source_list` and optionally the `alt_value` where sub-sequences used to be.

9.4.6.13.3 Example:

```
sequence s = remove_subseq({4,6,"Apple",0.1,{1,2,3},4})
-- 's' is now {4, 6, 0.1, 4} -- length now 4
s = remove_subseq({4,6,"Apple",0.1,{1,2,3},4}, -1)
-- 's' is now {4, 6, -1, 0.1, -1, 4} -- length unchanged.
```

9.4.6.14 RD_INPLACE

```
include std/sequence.e
public enum RD_INPLACE
```

These are used with the [remove_dups\(\)](#) function.

- ◆ `RD_INPLACE` removes items while preserving the original order of the unique items.
- ◆ `RD_PRESORTED` assumes that the elements in `source_data` are already sorted. If they are not already sorted, this option merely removed adjacent duplicate elements.
- ◆ `RD_SORT` will return the unique elements in ascending sorted order.

9.4.6.15 remove_dups

```
include std/sequence.e
public function remove_dups(sequence source_data, integer proc_option = RD_PRESORTED)
```

Removes duplicate elements

9.4.6.15.1 Parameters:

1. `source_data` : A sequence that may contain duplicated elements
2. `proc_option` : One of `RD_INPLACE`, `RD_PRESORTED`, or `RD_SORT`.
 - ◆ `RD_INPLACE` removes items while preserving the original order of the unique items.
 - ◆ `RD_PRESORTED` assumes that the elements in `source_data` are already sorted. If they are not already sorted, this option merely removed adjacent duplicate elements.
 - ◆ `RD_SORT` will return the unique elements in ascending sorted order.

9.4.6.15.2 Returns:

A **sequence**, that contains only the unique elements from `source_data`.

9.4.6.15.3 Example 1:

```
sequence s = { 4,7,9,7,2,5,5,9,0,4,4,5,6,5}
? remove_dups(s, RD_INPLACE) --> {4,7,9,2,5,0,6}
? remove_dups(s, RD_SORT) --> {0,2,4,5,6,7,9}
? remove_dups(s, RD_PRESORTED) --> {4,7,9,7,2,5,9,0,4,5,6,5}
? remove_dups(sort(s), RD_PRESORTED) --> {0,2,4,5,6,7,9}
```

9.4.6.16 COMBINE_UNSORTED

```
include std/sequence.e
public enum COMBINE_UNSORTED
```

9.4.6.17 COMBINE_SORTED

```
include std/sequence.e
public enum COMBINE_SORTED
```

9.4.6.18 combine

```
include std/sequence.e
public function combine(sequence source_data, integer proc_option = COMBINE_SORTED)
```

Combines all the sub-sequences into a single, optionally sorted, list

9.4.6.18.1 Parameters:

1. `source_data` : A sequence that contains sub-sequences to be combined.

2. `proc_option` : An integer; `COMBINE_UNSORTED` to return a non-sorted list and `COMBINE_SORTED` (the default) to return a sorted list.

9.4.6.18.2 Returns:

A **sequence**, that contains all the elements from all the first-level of sub-sequences from `source_data`.

9.4.6.18.3 Comments:

The elements in the sub-sequences do not have to be pre-sorted.

Only one level of sub-sequence is combined.

9.4.6.18.4 Example 1:

```
sequence s = { {4,7,9}, {7,2,5,9}, {0,4}, {5}, {6,5} }
? combine(s, COMBINE_SORTED)    --> {0,2,4,4,5,5,5,6,7,7,9,9}
? combine(s, COMBINE_UNSORTED) --> {4,7,9,7,2,5,9,0,4,5,6,5}
```

9.4.6.18.5 Example 2:

```
sequence s = { {"cat", "dog"}, {"fish", "whale"}, {"wolf"}, {"snail", "worm"} }
? combine(s)                --> {"cat", "dog", "fish", "snail", "whale", "wolf", "worm"}
? combine(s, COMBINE_UNSORTED) --> {"cat", "dog", "fish", "whale", "wolf", "snail", "worm"}
```

9.4.6.18.6 Example 3:

```
sequence s = { "cat", "dog", "fish", "whale", "wolf", "snail", "worm" }
? combine(s)                --> "aaacdeffghhiilllmnoorsstwww"
? combine(s, COMBINE_UNSORTED) --> "catdogfishwhalewolf snailworm"
```

9.4.6.19 minsize

```
include std/sequence.e
public function minsize(sequence source_data, integer min_size, object new_data)
```

Ensures that the supplied sequence is at least the supplied minimum length.

9.4.6.19.1 Parameters:

1. `source_data` : A sequence that might need extending.
2. `min_size`: An integer. The minimum length that `source_data` must be.
3. `new_data`: An object. This used to when `source_data` needs to be extended, in which case it is appended as many times as required to make the length equal to `min_size`.

9.4.6.19.2 Returns:

A **sequence**.

9.4.6.19.3 Example:

```
sequence s
s = minsize({4,3,6,2,7,1,2}, 10, -1) --> {4,3,6,2,7,1,2,-1,-1,-1}
s = minsize({4,3,6,2,7,1,2}, 5, -1) --> {4,3,6,2,7,1,2}
```

9.5 Serialization of EUPHORIA Objects

Routines

deserialize

serialize

dump

load

9.5.1 Routines

9.5.1.1 deserialize

```
include std/serialize.e
public function deserialize(object sdata, integer pos = 1)
```

Convert a serialized object in to a standard EUPHORIA object.

9.5.1.1.1 Parameters:

1. `sdata` : either a sequence containing one or more concatenated serialized objects or an open file handle. If this is a file handle, the current position in the file is assumed to be at a serialized object in the file.
2. `pos` : optional index into `sdata`. If omitted 1 is assumed. The index must point to the start of a serialized object.

9.5.1.1.2 Returns:

The return **value**, depends on the input type.

- If `sdata` is a file handle then this function returns a EUPHORIA object that had been stored in the file, and moves the current file to the first byte after the stored object.
- If `sdata` is a sequence then this returns a two-element sequence. The *first* element is the EUPHORIA object that corresponds to the serialized object that begins at index `pos`, and the *second*

element is the index position in the input parameter just after the serialized object.

9.5.1.1.3 Comments:

A serialized object is one that has been returned from the [serialize](#) function.

9.5.1.1.4 Example 1:

```
sequence objcache
objcache = serialize(FirstName) &
           serialize(LastName) &
           serialize(PhoneNumber) &
           serialize(Address)

sequence res
integer pos = 1
res = deserialize( objcache , pos)
FirstName = res[1] pos = res[2]
res = deserialize( objcache , pos)
LastName = res[1] pos = res[2]
res = deserialize( objcache , pos)
PhoneNumber = res[1] pos = res[2]
res = deserialize( objcache , pos)
Address = res[1] pos = res[2]
```

9.5.1.1.5 Example 2:

```
sequence objcache
objcache = serialize({FirstName,
                     LastName,
                     PhoneNumber,
                     Address})

sequence res
res = deserialize( objcache )
FirstName = res[1][1]
LastName = res[1][2]
PhoneNumber = res[1][3]
Address = res[1][4]
```

9.5.1.1.6 Example 3:

```
integer fh
fh = open("cust.dat", "wb")
puts(fh, serialize(FirstName))
puts(fh, serialize(LastName))
puts(fh, serialize(PhoneNumber))
puts(fh, serialize(Address))
close(fh)

fh = open("cust.dat", "rb")
FirstName = deserialize(fh)
```



```
LastName = deserialize(fh)
PhoneNumber = deserialize(fh)
Address = deserialize(fh)
close(fh)
```

9.5.1.1.7 Example 4:

```
integer fh
fh = open("cust.dat", "wb")
puts(fh, serialize({FirstName,
                    LastName,
                    PhoneNumber,
                    Address}))
close(fh)

sequence res
fh = open("cust.dat", "rb")
res = deserialize(fh)
close(fh)
FirstName = res[1]
LastName = res[2]
PhoneNumber = res[3]
Address = res[4]
```

9.5.1.2 serialize

```
include std/serialize.e
public function serialize(object x)
```

Convert a standard EUPHORIA object in to a serialized version of it.

9.5.1.2.1 Parameters:

1. `euobj` : any EUPHORIA object.

9.5.1.2.2 Returns:

A **sequence**, this is the serialized version of the input object.

9.5.1.2.3 Comments:

A serialized object is one that has been converted to a set of byte values. This can then be written directly out to a file for storage.

You can use the [deserialize](#) function to convert it back into a standard EUPHORIA object.

9.5.1.2.4 Example 1:

```
integer fh
fh = open("cust.dat", "wb")
puts(fh, serialize(FirstName))
puts(fh, serialize(LastName))
puts(fh, serialize(PhoneNumber))
puts(fh, serialize(Address))
close(fh)

fh = open("cust.dat", "rb")
FirstName = deserialize(fh)
LastName = deserialize(fh)
PhoneNumber = deserialize(fh)
Address = deserialize(fh)
close(fh)
```

9.5.1.2.5 Example 2:

```
integer fh
fh = open("cust.dat", "wb")
puts(fh, serialize({FirstName,
                    LastName,
                    PhoneNumber,
                    Address}))
close(fh)

sequence res
fh = open("cust.dat", "rb")
res = deserialize(fh)
close(fh)
FirstName = res[1]
LastName = res[2]
PhoneNumber = res[3]
Address = res[4]
```

9.5.1.3 dump

```
include std/serialize.e
public function dump(sequence data, sequence filename)
```

Saves a EUPHORIA object to disk in a binary format.

9.5.1.3.1 Parameters:

1. data : any EUPHORIA object.
2. filename : the name of the file to save it to.

9.5.1.3.2 Returns:

An **integer**, 0 if the function fails, otherwise the number of bytes in the created file.

9.5.1.3.3 Comments:

If the named file doesn't exist it is created, otherwise it is overwritten.

You can use the [load](#) function to recover the data from the file.

9.5.1.3.4 Example :

```
include std/serialize.e
integer size = dump(myData, theFileName)
if size = 0 then
    puts(1, "Failed to save data to file\n")
else
    printf(1, "Saved file is %d bytes long\n", size)
end if
```

9.5.1.4 load

```
include std/serialize.e
public function load(sequence filename)
```

Restores a EUPHORIA object that has been saved to disk by [dump](#).

9.5.1.4.1 Parameters:

1. `filename` : the name of the file to restore it from.

9.5.1.4.2 Returns:

A **sequence**, the first element is the result code. If the result code is 0 then it means that the function failed, otherwise the restored data is in the second element.

9.5.1.4.3 Comments:

This is used to load back data from a file created by the [dump](#) function.

9.5.1.4.4 Example :

```
include std/serialize.e
sequence mydata = load(theFileName)
if mydata[1] = 0 then
    puts(1, "Failed to load data from file\n")
else
    mydata = mydata[2] -- Restored data is in second element.
end if
```

9.6 Sorting

Constants

- ASCENDING
- NORMAL_ORDER
- DESCENDING
- REVERSE_ORDER

Routines

- sort
- custom_sort
- sort_columns
- merge
- insertion_sort

9.6.1 Constants

9.6.1.1 ASCENDING

```
include std/sort.e
public constant ASCENDING
```

ascending sort order, always the default.

When a sequence is sorted in `ASCENDING` order, its first element is the smallest as per the sort order and its last element is the largest

9.6.1.2 NORMAL_ORDER

```
include std/sort.e
public constant NORMAL_ORDER
```

The normal sort order used by the custom comparison routine.

9.6.1.3 DESCENDING

```
include std/sort.e
public constant DESCENDING
```

descending sort order, which is the reverse of `ASCENDING`.

9.6.1.4 REVERSE_ORDER

```
include std/sort.e
public constant REVERSE_ORDER
```

Reverses the sense of the order returned by a custom comparison routine.

9.6.2 Routines

9.6.2.1 sort

```
include std/sort.e
public function sort(sequence x, integer order = ASCENDING)
```

Sort the elements of a sequence into ascending order.

9.6.2.1.1 Parameters:

1. `x` : The sequence to be sorted.
2. `order` : the sort order. Default is `ASCENDING`.

9.6.2.1.2 Returns:

A **sequence**, a copy of the original sequence in ascending order

9.6.2.1.3 Comments:

The elements can be atoms or sequences.

The standard `compare()` routine is used to compare elements. This means that "`y` is greater than `x`" is defined by `compare(y, x)=1`.

This function uses the "Shell" sort algorithm. This sort is not "stable", i.e. elements that are considered equal might change position relative to each other.

9.6.2.1.4 Example 1:

```
constant student_ages = {18,21,16,23,17,16,20,20,19}
sequence sorted_ages
sorted_ages = sort( student_ages )
-- result is {16,16,17,18,19,20,20,21,23}
```

9.6.2.1.5 See Also:

[compare](#), [custom_sort](#)

9.6.2.2 custom_sort

```
include std/sort.e
public function custom_sort(integer custom_compare, sequence x, object data = {}, integer order)
```

Sort the elements of a sequence according to a user-defined order.

9.6.2.2.1 Parameters:

1. `custom_compare` : an integer, the routine-id of the user defined routine that compares two items which appear in the sequence to sort.
2. `x` : the sequence of items to be sorted.
3. `data` : an object, either `{}` (no custom data, the default), an atom or a non-empty sequence.
4. `order` : an integer, either `NORMAL_ORDER` (the default) or `REVERSE_ORDER`.

9.6.2.2.2 Returns:

A **sequence**, a copy of the original sequence in sorted order

9.6.2.2.3 Errors:

If the user defined routine does not return according to the specifications in the *Comments* section below, an error will occur.

9.6.2.2.4 Comments:

- If some user data is being provided, that data must be either an atom or a sequence with at least one element. **NOTE** only the first element is passed to the user defined comparison routine, any other elements are just ignored. The user data is not used or inspected in any way other than passing it to the user defined routine.
- The user defined routine must return an integer *comparison result*
 - ◆ a **negative** value if object A must appear before object B

- ◆ a **positive** value if object B must appear before object A
- ◆ 0 if the order does not matter

NOTE: The meaning of the value returned by the user-defined routine is reversed when `order = REVERSE_ORDER`. The default is `order = NORMAL_ORDER`, which sorts in order returned by the custom comparison routine.

- When no user data is provided, the user defined routine must accept two objects (A, B) and return just the *comparison result*.
- When some user data is provided, the user defined routine must take three objects (A, B , data). It must return either...
 - ◆ an integer, which is a *comparison result*
 - ◆ a two-element sequence, in which the first element is a *comparison result* and the second element is the updated user data that is to be used for the next call to the user defined routine.
- The elements of `x` can be atoms or sequences. Each time that the sort needs to compare two items in the sequence, it calls the user-defined function to determine the order.
- This function uses the "Shell" sort algorithm. This sort is not "stable", i.e. elements that are considered equal might change position relative to each other.

9.6.2.2.5 Example 1:

```
constant students = {{"Anne",18}, {"Bob",21},
                    {"Chris",16}, {"Diane",23},
                    {"Eddy",17}, {"Freya",16},
                    {"George",20}, {"Heidi",20},
                    {"Ian",19}}

sequence sorted_byage
function byage(object a, object b)
    ----- If the ages are the same, compare the names otherwise just compare ages.
    if equal(a[2], b[2]) then
        return compare(upper(a[1]), upper(b[1]))
    end if
    return compare(a[2], b[2])
end function

sorted_byage = custom_sort( routine_id("byage"), students )
-- result is {"Chris",16}, {"Freya",16},
--           {"Eddy",17}, {"Anne",18},
--           {"Ian",19}, {"George",20},
--           {"Heidi",20}, {"Bob",21},
--           {"Diane",23}}

sorted_byage = custom_sort( routine_id("byage"), students,, REVERSE_ORDER )
-- result is {"Diane",23}, {"Bob",21},
--           {"Heidi",20}, {"George",20},
--           {"Ian",19}, {"Anne",18},
--           {"Eddy",17}, {"Freya",16},
--           {"Chris",16}}
--
```


9.6.2.2.6 Example 2:

```

constant students = {{"Anne", "Baxter", 18}, {"Bob", "Palmer", 21},
                    {"Chris", "du Pont", 16}, {"Diane", "Fry", 23},
                    {"Eddy", "Ammon", 17}, {"Freya", "Brash", 16},
                    {"George", "Gungle", 20}, {"Heidi", "Smith", 20},
                    {"Ian", "Sidebottom", 19}}

sequence sorted
function colsort(object a, object b, sequence cols)
    integer sign
    for i = 1 to length(cols) do
        if cols[i] < 0 then
            sign = -1
            cols[i] = -cols[i]
        else
            sign = 1
        end if
        if not equal(a[cols[i]], b[cols[i]]) then
            return sign * compare(upper(a[cols[i]]), upper(b[cols[i]]))
        end if
    end for

    return 0
end function

-- Order is age:descending, Surname, Given Name
sequence column_order = {-3, 2, 1}
sorted = custom_sort( routine_id("colsort"), students, {column_order} )
-- result is
{
    {"Diane", "Fry", 23},
    {"Bob", "Palmer", 21},
    {"George", "Gungle", 20},
    {"Heidi", "Smith", 20},
    {"Ian", "Sidebottom", 19},
    {"Anne", "Baxter", 18 },
    {"Eddy", "Ammon", 17},
    {"Freya", "Brash", 16},
    {"Chris", "du Pont", 16}
}

sorted = custom_sort( routine_id("colsort"), students, {column_order}, REVERSE_ORDER )
-- result is
{
    {"Chris", "du Pont", 16},
    {"Freya", "Brash", 16},
    {"Eddy", "Ammon", 17},
    {"Anne", "Baxter", 18 },
    {"Ian", "Sidebottom", 19},
    {"Heidi", "Smith", 20},
    {"George", "Gungle", 20},
    {"Bob", "Palmer", 21},
    {"Diane", "Fry", 23}
}

```

9.6.2.2.7 See Also:

[compare](#), [sort](#)

9.6.2.3 sort_columns

```
include std/sort.e
public function sort_columns(sequence x, sequence column_list)
```

Sort the rows in a sequence according to a user-defined column order.

9.6.2.3.1 Parameters:

1. `x` : a sequence, holding the sequences to be sorted.
2. `column_list` : a list of columns indexes `x` is to be sorted by.

9.6.2.3.2 Returns:

A **sequence**, a copy of the original sequence in sorted order.

9.6.2.3.3 Comments:

`x` must be a sequence of sequences.

A non-existent column is treated as coming before an existing column. This allows sorting of records that are shorter than the columns in the column list.

By default, columns are sorted in ascending order. To sort in descending order, make the column number negative.

This function uses the "Shell" sort algorithm. This sort is not "stable", i.e. elements that are considered equal might change position relative to each other.

9.6.2.3.4 Example 1:

```
sequence dirlist
dirlist = dir("c:\\temp")
sequence sorted
-- Order is Size:descending, Name
sorted = sort_columns( dirlist, {-D_SIZE, D_NAME} )
```

9.6.2.3.5 See Also:

[compare](#), [sort](#)

9.6.2.4 merge

```
include std/sort.e
public function merge(sequence a, sequence b, integer compfunc = - 1, object userdata = "")
```

Merge two pre-sorted sequences into a single sequence.

9.6.2.4.1 Parameters:

1. *a* : a sequence, holding pre-sorted data.
2. *b* : a sequence, holding pre-sorted data.
3. *compfunc* : an integer, either -1 or the routine id of a user-defined comparison function.

9.6.2.4.2 Returns:

A **sequence**, consisting of *a* and *b* merged together.

9.6.2.4.3 Comments:

- If *a* or *b* is not already sorted, the resulting sequence might not be sorted either.
- The input sequences do not have to be the same size.
- The user-defined comparison function must accept two objects and return an integer. It returns -1 if the first object must appear before the second one, and 1 if the first object must after before the second one, and 0 if the order doesn't matter.

9.6.2.4.4 Example 1:

```
sequence X, Y
X = sort( {5,3,7,1,9,0} ) --> {0,1,3,5,7,9}
Y = sort( {6,8,10,2} ) --> {2,6,8,10}
? merge(X, Y) --> {0,1,2,3,5,6,7,8,9,10}
```

9.6.2.4.5 See Also:

[compare](#), [sort](#)

9.6.2.5 insertion_sort

```
include std/sort.e
public function insertion_sort(sequence s, object e = "", integer compfunc = - 1, object userda
```

Sort a sequence, and optionally another object together.

9.6.2.5.1 Parameters:

1. *s* : a sequence, holding data to be sorted.
2. *e* : an object. If this is an atom, it is sorted in with *s*. If this is a non-empty sequence then *s* and *e* are both sorted independantly using this `insertion_sort` function and then the results are merged and returned.
3. *compfunc* : an integer, either -1 or the routine id of a user-defined comparision function.

9.6.2.5.2 Returns:

A **sequence**, consisting of *s* and *e* sorted together.

9.6.2.5.3 Comments:

- This routine is usually a lot faster than the standard sort when *s* and *e* are (mostly) sorted before calling the function. For example, you can use this routine to quickly add to a sorted list.
- The input sequences do not have to be the same size.
- The user-defined comparision function must accept two objects and return an integer. It returns -1 if the first object must appear before the second one, and 1 if the first object must after before the second one, and 0 if the order doesn't matter.

9.6.2.5.4 Example 1:

```
sequence X = {}
while true do
    newdata = get_data()
    if compare(-1, newdata) then
        exit
    end if
    X = insertion_sort(X, newdata)
    process(new_data)
end while
```

9.6.2.5.5 See Also:

[compare](#), [sort](#), [merge](#)

10 String Centric Routines

Locale Routines

[Message translation functions](#)

[Time/Number Translation](#)

[Constants](#)

[Locale Name Translation](#)

Regular Expressions

[Introduction](#)

[General Use](#)

[Option Constants](#)

[Error Constants](#)

[Create/Destroy](#)

[Utility Routines](#)

[Find/Match](#)

[Splitting](#)

[Replacement](#)

Text Manipulation

[Routines](#)

Unicode

Wildcard Matching

[Routines](#)

10.1 Locale Routines

Page Contents

[Message translation functions](#)

[set_lang_path](#)

[get_lang_path](#)

[lang_load](#)

[set_def_lang](#)

[get_def_lang](#)

[translate](#)

[trsprintf](#)

[Time/Number Translation](#)

[set](#)

[get](#)

[money](#)

[number](#)

[datetime](#)

[Constants](#)

[w32_names](#)

[w32_name_canonical](#)

[posix_names](#)

locale_canonical
platform_locale
Locale Name Translation
canonical
decanonical
canon2win

10.1.1 Message translation functions

10.1.1.1 set_lang_path

```
include std/locale.e  
public procedure set_lang_path(object pp)
```

Set the language path.

10.1.1.1.1 Parameters:

1. pp : an object, either an actual path or an atom.

10.1.1.1.2 Comments:

When the language path is not set, and it is unset by default, [set\(\)](#) does not load any language file.

10.1.1.1.3 See Also:

[set](#)

10.1.1.2 get_lang_path

```
include std/locale.e  
public function get_lang_path()
```

Get the language path.

10.1.1.2.1 Returns:

An **object**, the current language path.

10.1.1.2.2 See Also:

[get_lang_path](#)

10.1.1.3 lang_load

```
include std/locale.e
public function lang_load(sequence filename)
```

Load a language file.

10.1.1.3.1 Parameters:

1. `filename` : a sequence, the name of the file to load. If no file extension is supplied, then ".lng" is used.

10.1.1.3.2 Returns:

A language **map**, if successful. This is to be used when calling [translate\(\)](#).

If the load fails it returns a zero.

10.1.1.3.3 Comments:

The language file must be made of lines which are either comments, empty lines or translations. Note that leading whitespace is ignored on all lines except continuation lines.

- **Comments** are lines that begin with a # character and extend to the end of the line.
- **Empty Lines** are ignored.
- **Translations** have two forms ...

```
keyword translation_text
```

In which the 'keyword' is a word that must not have any spaces in it.

```
keyphrase = translation_text
```

In which the 'keyphrase' is anything up to the first '=' symbol.

It is possible to have the translation text span multiple lines. You do this by having '&' as the last character of the line. These are placed by newline characters when loading.

10.1.1.3.4 Example:

```
# Example translation file
#

hello Hola
world Mundo
greeting %s, %s!

help text = &
This is an example of some &
translation text that spans &
multiple lines.

# End of example PO #2
```

10.1.1.3.5 See Also:

[translate](#)

10.1.1.4 set_def_lang

```
include std/locale.e
public procedure set_def_lang(object langmap)
```

Sets the default language (translation) map

10.1.1.4.1 Parameters:

1. langmap : A value returned by [lang_load\(\)](#), or zero to remove any default map.

10.1.1.4.2 Example:

```
set_def_lang( lang_load("appmsgs") )
```

10.1.1.5 get_def_lang

```
include std/locale.e
public function get_def_lang()
```

Gets the default language (translation) map

10.1.1.5.1 Parameters:

none.

10.1.1.5.2 Returns:

An **object**, a language map, or zero if there is no default language map yet.

10.1.1.5.3 Example:

```
object langmap = get_def_lang()
```

10.1.1.6 translate

```
include std/locale.e
public function translate(sequence word, object langmap = 0, object defval = "", integer mode =
```

Translates a word, using the current language file.

10.1.1.6.1 Parameters:

1. **word** : a sequence, the word to translate.
2. **langmap** : Either a value returned by [lang_load\(\)](#) or zero to use the default language map
3. **defval** : a object. The value to return if the word cannot be translated. Default is "". If **defval** is **PINF** then the **word** is returned if it can't be translated.
4. **mode** : an integer. If zero (the default) it uses **word** as the keyword and returns the translation text. If not zero it uses **word** as the translation and returns the keyword.

10.1.1.6.2 Returns:

A **sequence**, the value associated with **word**, or **defval** if there is no association.

10.1.1.6.3 Example 1:

```
sequence newword
newword = translate(msgtext)
if length(msgtext) = 0 then
    error_message(msgtext)
else
    error_message(newword)
end if
```

10.1.1.6.4 Example 2:

```
error_message(translate(msgtext, , PINF))
```

10.1.1.6.5 See Also:

[set](#), [lang_load](#)

10.1.1.7 trsprintf

```
include std/locale.e
public function trsprintf(sequence fmt, sequence data, object langmap = 0)
```

Returns a formatted string with automatic translation performed on the parameters.

10.1.1.7.1 Parameters:

1. `fmt` : A sequence. Contains the formatting string. see [printf\(\)](#) for details.
2. `data` : A sequence. Contains the data that goes into the formatted result. see [printf](#) for details.
3. `langmap` : An object. Either 0 (the default) to use the default language maps, or the result returned from [lang_load\(\)](#) to specify a particular language map.

10.1.1.7.2 Returns:

A **sequence**, the formatted result.

10.1.1.7.3 Comments:

This works very much like the [sprintf\(\)](#) function. The difference is that the `fmt` sequence and sequences contained in the `data` parameter are **translated** before passing them to [sprintf](#). If an item has no translation, it remains unchanged.

Further more, after the translation pass, if the result text begins with "`__`", the "`__`" is removed. This method can be used when you do not want an item to be translated.

10.1.1.7.4 Examples:

```
-- Assuming a language has been loaded and
-- "greeting" translates as '%s %s, %s'
-- "hello" translates as "G'day"
-- "how are you today" translates as "How's the family?"
sequence UserName = "Bob"
sequence result = trsprintf( "greeting", {"hello", "__" & UserName, "how are you today"})
--> "G'day Bob, How's the family?"
```

10.1.2 Time/Number Translation

10.1.2.1 set

```
include std/locale.e
public function set(sequence new_locale)
```

Set the computer locale, and possibly load appropriate translation file.

10.1.2.1.1 Parameters:

1. `new_locale` : a sequence representing a new locale.

10.1.2.1.2 Returns:

An **integer**, either 0 on failure or 1 on success.

10.1.2.1.3 Comments:

Locale strings have the following format: `xx_YY` or `xx_YY.xyz` . The `xx` part refers to a culture, or main language/script. For instance, "en" refers to English, "de" refers to German, and so on. For some language, a script may be specified, like in "mn_Cyrl_MN" (mongolian in cyrillic transcription).

The `YY` part refers to a subculture, or variant, of the main language. For instance, "fr_FR" refers to metropolitan France, while "fr_BE" refers to the variant spoken in Wallonie, the French speaking region of Belgium.

The optional `.xyz` part specifies an encoding, like `.utf8` or `.1252` . This is required in some cases.

10.1.2.2 get

```
include std/locale.e
public function get()
```

Get current locale string

10.1.2.2.1 Returns:

A **sequence**, a locale string.

10.1.2.2.2 See Also:

[set](#)

10.1.2.3 money

```
include std/locale.e
public function money(object amount)
```

Converts an amount of currency into a string representing that amount.

10.1.2.3.1 Parameters:

1. `amount` : an atom, the value to write out.

10.1.2.3.2 Returns:

A **sequence**, a string that writes out `amount` of current currency.

10.1.2.3.3 Example 1:

```
-- Assuming an en_US locale
? money(1020.5) -- returns"$1,020.50"
```

10.1.2.3.4 See Also:

[set](#), [number](#)

10.1.2.4 number

```
include std/locale.e
public function number(object num)
```

Converts a number into a string representing that number.

10.1.2.4.1 Parameters:

1. `num` : an atom, the value to write out.

10.1.2.4.2 Returns:

A **sequence**, a string that writes out num.

10.1.2.4.3 Example 1:

```
-- Assuming an en_US locale
? number(1020.5) -- returns "1,020.50"
```

10.1.2.4.4 See Also:

[set](#), [money](#)

10.1.2.5 datetime

```
include std/locale.e
public function datetime(sequence fmt, dt :datetime dtm)
```

Formats a date according to current locale.

10.1.2.5.1 Parameters:

1. `fmt` : A format string, as described in [datetime:format](#)
2. `dtm` : the datetime to write out.

10.1.2.5.2 Returns:

A **sequence**, representing the formatted date.

10.1.2.5.3 Example 1:

```
? datetime("Today is a %A", dt:now())
```

10.1.2.5.4 See Also:

[datetime:format](#)

10.1.3 Constants

Win32 locale names:

af-ZA	sq-AL	gsw-FR	am-ET	ar-DZ	ar-BH	ar-EG	ar-IQ
ar-JO	ar-KW	ar-LB	ar-LY	ar-MA	ar-OM	ar-QA	ar-SA
ar-SY	ar-TN	ar-AE	ar-YE	hy-AM	as-IN	az-Cyrl-AZ	az-Latn-AZ
ba-RU	eu-ES	be-BY	bn-IN	bs-Cyrl-BA	bs-Latn-BA	br-FR	bg-BG
ca-ES	zh-HK	zh-MO	zh-CN	zh-SG	zh-TW	co-FR	hr-BA
hr-HR	cs-CZ	da-DK	prs-AF	dv-MV	nl-BE	nl-NL	en-AU
en-BZ	en-CA	en-029	en-IN	en-IE	en-JM	en-MY	en-NZ
en-PH	en-SG	en-ZA	en-TT	en-GB	en-US	en-ZW	et-EE
fo-FO	fil-PH	fi-FI	fr-BE	fr-CA	fr-FR	fr-LU	fr-MC
fr-CH	fy-NL	gl-ES	ka-GE	de-AT	de-DE	de-LI	de-LU
de-CH	el-GR	kl-GL	gu-IN	ha-Latn-NG	he-IL	hi-IN	hu-HU
is-IS	ig-NG	id-ID	iu-Latn-CA	iu-Cans-CA	ga-IE	it-IT	it-CH
ja-JP	kn-IN	kk-KZ	kh-KH	qut-GT	rw-RW	kok-IN	ko-KR
ky-KG	lo-LA	lv-LV	lt-LT	dsb-DE	lb-LU	mk-MK	ms-BN
ms-MY	ml-IN	mt-MT	mi-NZ	arn-CL	mr-IN	moh-CA	mn-Cyrl-MN
mn-Mong-CN	ne-IN	ne-NP	nb-NO	nn-NO	oc-FR	or-IN	ps-AF
fa-IR	pl-PL	pt-BR	pt-PT	pa-IN	quz-BO	quz-EC	quz-PE
ro-RO	rm-CH	ru-RU	smn-FI	smj-NO	smj-SE	se-FI	se-NO
se-SE	sms-FI	sma-NO	sma-SE	sa-IN	sr-Cyrl-BA	sr-Latn-BA	sr-Cyrl-CS
sr-Latn-CS	ns-ZA	tn-ZA	si-LK	sk-SK	sl-SI	es-AR	es-BO
es-CL	es-CO	es-CR	es-DO	es-EC	es-SV	es-GT	es-HN
es-MX	es-NI	es-PA	es-PY	es-PE	es-PR	es-ES	es-ES_tradnl
es-US	es-UY	es-VE	sw-KE	sv-FI	sv-SE	syr-SY	tg-Cyrl-TJ
tmz-Latn-DZ	ta-IN	tt-RU	te-IN	th-TH	bo-BT	bo-CN	tr-TR
tk-TM	ug-CN	uk-UA	wen-DE	tr-IN	ur-PK	uz-Cyrl-UZ	uz-Latn-UZ
vi-VN	cy-GB	wo-SN	xh-ZA	sah-RU	ii-CN	yo-NG	zu-ZA

10.1.3.1 w32_names

```
include std/localeconv.e
public constant w32_names
```

10.1.3.2 w32_name_canonical

```
include std/localeconv.e
public constant w32_name_canonical
```

Canonical locale names for WIN32:

```
Afrikaans_South Africa  Afrikaans_South Africa  Afrikaans_South Africa
Afrikaans_South Africa  Afrikaans_South Africa  Afrikaans_South Africa
```

Afrikaans_South Africa	Afrikaans_South Africa	Afrikaans_South Africa
Afrikaans_South Africa	Afrikaans_South Africa	Afrikaans_South Africa
Afrikaans_South Africa	Afrikaans_South Africa	Afrikaans_South Africa
Afrikaans_South Africa	Afrikaans_South Africa	Afrikaans_South Africa
Afrikaans_South Africa	Afrikaans_South Africa	Afrikaans_South Africa
Afrikaans_South Africa	Afrikaans_South Africa	Afrikaans_South Africa
Basque_Spain	Basque_Spain	Belarusian_Belarus
Belarusian_Belarus	Belarusian_Belarus	Belarusian_Belarus
Belarusian_Belarus	Belarusian_Belarus	Catalan_Spain
Catalan_Spain	Catalan_Spain	Catalan_Spain
Catalan_Spain	Catalan_Spain	Catalan_Spain
Catalan_Spain	Catalan_Spain	Catalan_Spain
Danish_Denmark	Danish_Denmark	Danish_Denmark
Danish_Denmark	Danish_Denmark	English_Australia
English_United States	English_United States	English_United States
English_United States	English_United States	English_United States
English_United States	English_United States	English_United States
English_United States	English_United States	English_United States
English_United States	English_United States	English_United States
English_United States	English_United States	English_United States
Finnish_Finland	French_France	French_France
French_France	French_France	French_France
French_France	French_France	French_France
French_France	French_France	French_France
French_France	French_France	French_France
French_France	French_France	French_France
French_France	French_France	French_France
Hungarian_Hungary	Hungarian_Hungary	Hungarian_Hungary
Hungarian_Hungary	Hungarian_Hungary	Hungarian_Hungary
Hungarian_Hungary	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy

Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Italian_Italy	Italian_Italy
Italian_Italy	Romanian_Romania	Romanian_Romania
Russian_Russia	Russian_Russia	Russian_Russia
Russian_Russia	Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia
Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia
Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia
Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia
Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia	Serbian (Cyrillic)_Serbia
Serbian (Cyrillic)_Serbia	Slovak_Slovakia	Estonian_Estonia
Estonian_Estonia	Estonian_Estonia	Estonian_Estonia
Estonian_Estonia	Estonian_Estonia	Estonian_Estonia
Estonian_Estonia	Estonian_Estonia	Estonian_Estonia
Estonian_Estonia	Estonian_Estonia	Estonian_Estonia
Estonian_Estonia	Estonian_Estonia	Estonian_Estonia
Estonian_Estonia	Estonian_Estonia	Swedish_Sweden
Swedish_Sweden	Swedish_Sweden	Swedish_Sweden
Swedish_Sweden	Swedish_Sweden	Swedish_Sweden
Swedish_Sweden	Swedish_Sweden	Swedish_Sweden
Swedish_Sweden	Swedish_Sweden	Swedish_Sweden
Swedish_Sweden	Swedish_Sweden	Ukrainian_Ukraine
Ukrainian_Ukraine	Ukrainian_Ukraine	Ukrainian_Ukraine
Ukrainian_Ukraine	Ukrainian_Ukraine	Ukrainian_Ukraine
Ukrainian_Ukraine	Ukrainian_Ukraine	Ukrainian_Ukraine
Ukrainian_Ukraine		

10.1.3.3 posix_names

```
include std/localeconv.e
public constant posix_names
```

10.1.3.3.1 POSIX locale names:

af_ZA	sq_AL	gsw_FR	am_ET	ar_DZ	ar_BH	ar_EG	ar_IQ
ar_JO	ar_KW	ar_LB	ar_LY	ar_MA	ar_OM	ar_QA	ar_SA
ar_SY	ar_TN	ar_AE	ar_YE	hy_AM	as_IN	az_Cyrl_AZ	az_Latn_AZ
ba_RU	eu_ES	be_BY	bn_IN	bs_Cyrl_BA	bs_Latn_BA	br_FR	bg_BG
ca_ES	zh_HK	zh_MO	zh_CN	zh_SG	zh_TW	co_FR	hr_BA
hr_HR	cs_CZ	da_DK	prs_AF	dv_MV	nl_BE	nl_NL	en_AU



en_BZ	en_CA	en_029	en_IN	en_IE	en_JM	en_MY	en_NZ
en_PH	en_SG	en_ZA	en_TT	en_GB	en_US	en_ZW	et_EE
fo_FO	fil_PH	fi_FI	fr_BE	fr_CA	fr_FR	fr_LU	fr_MC
fr_CH	fy_NL	gl_ES	ka_GE	de_AT	de_DE	de_LI	de_LU
de_CH	el_GR	kl_GL	gu_IN	ha_Latn_NG	he_IL	hi_IN	hu_HU
is_IS	ig_NG	id_ID	iu_Latn_CA	iu_Cans_CA	ga_IE	it_IT	it_CH
ja_JP	kn_IN	kk_KZ	kh_KH	qut_GT	rw_RW	kok_IN	ko_KR
ky_KG	lo_LA	lv_LV	lt_LT	dsb_DE	lb_LU	mk_MK	ms_BN
ms_MY	ml_IN	mt_MT	mi_NZ	arn_CL	mr_IN	moh_CA	mn_Cyrl_MN
mn_Mong_CN	ne_IN	ne_NP	nb_NO	nn_NO	oc_FR	or_IN	ps_AF
fa_IR	pl_PL	pt_BR	pt_PT	pa_IN	quz_BO	quz_EC	quz_PE
ro_RO	rm_CH	ru_RU	smn_FI	smj_NO	smj_SE	se_FI	se_NO
se_SE	sms_FI	sma_NO	sma_SE	sa_IN	sr_Cyrl_BA	sr_Latn_BA	sr_Cyrl_CS
sr_Latn_CS	ns_ZA	tn_ZA	si_LK	sk_SK	sl_SI	es_AR	es_BO
es_CL	es_CO	es_CR	es_DO	es_EC	es_SV	es_GT	es_HN
es_MX	es_NI	es_PA	es_PY	es_PE	es_PR	es_ES	es_ES_tradnl
es_US	es_UY	es_VE	sw_KE	sv_FI	sv_SE	syr_SY	tg_Cyrl_TJ
tmz_Latn_DZ	ta_IN	tt_RU	te_IN	th_TH	bo_BT	bo_CN	tr_TR
tk_TM	ug_CN	uk_UA	wen_DE	tr_IN	ur_PK	uz_Cyrl_UZ	uz_Latn_UZ
vi_VN	cy_GB	wo_SN	xh_ZA	sah_RU	ii_CN	yo_NG	zu_ZA

10.1.3.4 locale_canonical

```
include std/localeconv.e
public constant locale_canonical
```

10.1.3.5 platform_locale

```
include std/localeconv.e
public constant platform_locale
```

10.1.4 Locale Name Translation

10.1.4.1 canonical

```
include std/localeconv.e
public function canonical(sequence new_locale)
```

Get canonical name for a locale.

10.1.4.1.1 Parameters:

1. `new_locale` : a sequence, the string for the locale.

10.1.4.1.2 Returns:

A **sequence**, either the translated locale on success or `new_locale` on failure.

10.1.4.1.3 See Also:

[get](#), [set](#), [decanonical](#)

10.1.4.2 decanonical

```
include std/localeconv.e
public function decanonical(sequence new_locale)
```

Get the translation of a locale string for current platform.

10.1.4.2.1 Parameters:

1. `new_locale`: a sequence, the string for the locale.

10.1.4.2.2 Returns:

A **sequence**, either the translated locale on success or `new_locale` on failure.

10.1.4.2.3 See Also:

[get](#), [set](#), [canonical](#)

10.1.4.3 canon2win

```
include std/localeconv.e
public function canon2win(sequence new_locale)
```

TODO: document

10.2 Regular Expressions

Introduction

General Use

Option Constants

DEFAULT

CASELESS

MULTILINE

DOTALL

EXTENDED

ANCHORED

DOLLAR_ENDONLY

EXTRA

NOTBOL

NOTEOL

UNGREEDY

NOTEMPTY

UTF8

NO_AUTO_CAPTURE

NO_UTF8_CHECK

AUTO_CALLOUT

PARTIAL

DFA_SHORTEST

DFA_RESTART

FIRSTLINE

DUPNAMES

NEWLINE_CR

NEWLINE_LF

NEWLINE_CRLF

NEWLINE_ANY

NEWLINE_ANYCRLF

BSR_ANYCRLF

BSR_UNICODE

STRING_OFFSETS

Error Constants

ERROR_NOMATCH

ERROR_NULL

ERROR_BADOPTION

ERROR_BADMAGIC

ERROR_UNKNOWN_OPCODE

ERROR_UNKNOWN_NODE

ERROR_NOMEMORY

ERROR_NOSUBSTRING

ERROR_MATCHLIMIT

ERROR_CALLOUT

ERROR_BADUTF8

ERROR_BADUTF8_OFFSET

ERROR_PARTIAL

ERROR_BADPARTIAL
ERROR_INTERNAL
ERROR_BADCOUNT
ERROR_DFA_UITEM
ERROR_DFA_UCOND
ERROR_DFA_UMLIMIT
ERROR_DFA_WSSIZE
ERROR_DFA_RECURSE
ERROR_RECURSIONLIMIT
ERROR_NULLWSLIMIT
ERROR_BADNEWLINE
Create/Destroy
 regex
 new
 error_message
Utility Routines
 escape
 get_ovector_size
Find/Match
 find
 find_all
 has_match
 is_match
 matches
 all_matches
Splitting
 split
 split_limit
Replacement
 find_replace
 find_replace_limit
 find_replace_callback

10.2.1 Introduction

Regular expressions in EUPHORIA are based on the PCRE (Perl Compatible Regular Expressions) library created by Philip Hazel.

This document will detail the EUPHORIA interface to Regular Expressions, not really regular expression syntax. It is a very complex subject that many books have been written on. Here are a few good resources online that can help while learning regular expressions.

- [EUForum Article](#)
- [Perl Regular Expressions Man Page](#)
- [Regular Expression Library](#) (user supplied regular expressions for just about any task).
- [WikiPedia Regular Expression Article](#)

10.2.2 General Use

Many functions take an optional `options` parameter. This parameter can be either a single option constant (see [Option Constants](#), multiple option constants or'ed together into a single atom or a sequence of options, in which the function will take care of ensuring they are or'ed together correctly.

10.2.3 Option Constants

10.2.3.1 DEFAULT

```
include std/regex.e
public constant DEFAULT
```

10.2.3.2 CASELESS

```
include std/regex.e
public constant CASELESS
```

10.2.3.3 MULTILINE

```
include std/regex.e
public constant MULTILINE
```

10.2.3.4 DOTALL

```
include std/regex.e
public constant DOTALL
```

10.2.3.5 EXTENDED

```
include std/regex.e
public constant EXTENDED
```

10.2.3.6 ANCHORED

```
include std/regex.e
public constant ANCHORED
```

10.2.3.7 DOLLAR_ENDONLY

```
include std/regex.e
public constant DOLLAR_ENDONLY
```

10.2.3.8 EXTRA

```
include std/regex.e
public constant EXTRA
```

10.2.3.9 NOTBOL

```
include std/regex.e
public constant NOTBOL
```

10.2.3.10 NOTEOL

```
include std/regex.e
public constant NOTEOL
```

10.2.3.11 UNGREEDY

```
include std/regex.e
public constant UNGREEDY
```

10.2.3.12 NOTEMPTY

```
include std/regex.e
public constant NOTEMPTY
```

10.2.3.13 UTF8

```
include std/regex.e
public constant UTF8
```



10.2.3.14 NO_AUTO_CAPTURE

```
include std/regex.e  
public constant NO_AUTO_CAPTURE
```

10.2.3.15 NO_UTF8_CHECK

```
include std/regex.e  
public constant NO_UTF8_CHECK
```

10.2.3.16 AUTO_CALLOUT

```
include std/regex.e  
public constant AUTO_CALLOUT
```

10.2.3.17 PARTIAL

```
include std/regex.e  
public constant PARTIAL
```

10.2.3.18 DFA_SHORTEST

```
include std/regex.e  
public constant DFA_SHORTEST
```

10.2.3.19 DFA_RESTART

```
include std/regex.e  
public constant DFA_RESTART
```

10.2.3.20 FIRSTLINE

```
include std/regex.e  
public constant FIRSTLINE
```

10.2.3.21 DUPNAMES

```
include std/regex.e
public constant DUPNAMES
```

10.2.3.22 NEWLINE_CR

```
include std/regex.e
public constant NEWLINE_CR
```

10.2.3.23 NEWLINE_LF

```
include std/regex.e
public constant NEWLINE_LF
```

10.2.3.24 NEWLINE_CRLF

```
include std/regex.e
public constant NEWLINE_CRLF
```

10.2.3.25 NEWLINE_ANY

```
include std/regex.e
public constant NEWLINE_ANY
```

10.2.3.26 NEWLINE_ANYCRLF

```
include std/regex.e
public constant NEWLINE_ANYCRLF
```

10.2.3.27 BSR_ANYCRLF

```
include std/regex.e
public constant BSR_ANYCRLF
```


10.2.3.28 BSR_UNICODE

```
include std/regex.e  
public constant BSR_UNICODE
```

10.2.3.29 STRING_OFFSETS

```
include std/regex.e  
public constant STRING_OFFSETS
```

10.2.4 Error Constants

10.2.4.1 ERROR_NOMATCH

```
include std/regex.e  
public constant ERROR_NOMATCH
```

10.2.4.2 ERROR_NULL

```
include std/regex.e  
public constant ERROR_NULL
```

10.2.4.3 ERROR_BADOPTION

```
include std/regex.e  
public constant ERROR_BADOPTION
```

10.2.4.4 ERROR_BADMAGIC

```
include std/regex.e  
public constant ERROR_BADMAGIC
```

10.2.4.5 ERROR_UNKNOWN_OPCODE

```
include std/regex.e  
public constant ERROR_UNKNOWN_OPCODE
```

10.2.4.6 ERROR_UNKNOWN_NODE

```
include std/regex.e
public constant ERROR_UNKNOWN_NODE
```

10.2.4.7 ERROR_NOMEMORY

```
include std/regex.e
public constant ERROR_NOMEMORY
```

10.2.4.8 ERROR_NOSUBSTRING

```
include std/regex.e
public constant ERROR_NOSUBSTRING
```

10.2.4.9 ERROR_MATCHLIMIT

```
include std/regex.e
public constant ERROR_MATCHLIMIT
```

10.2.4.10 ERROR_CALLOUT

```
include std/regex.e
public constant ERROR_CALLOUT
```

10.2.4.11 ERROR_BADUTF8

```
include std/regex.e
public constant ERROR_BADUTF8
```

10.2.4.12 ERROR_BADUTF8_OFFSET

```
include std/regex.e
public constant ERROR_BADUTF8_OFFSET
```

10.2.4.13 ERROR_PARTIAL

```
include std/regex.e
public constant ERROR_PARTIAL
```

10.2.4.14 ERROR_BADPARTIAL

```
include std/regex.e
public constant ERROR_BADPARTIAL
```

10.2.4.15 ERROR_INTERNAL

```
include std/regex.e
public constant ERROR_INTERNAL
```

10.2.4.16 ERROR_BADCOUNT

```
include std/regex.e
public constant ERROR_BADCOUNT
```

10.2.4.17 ERROR_DFA_UITEM

```
include std/regex.e
public constant ERROR_DFA_UITEM
```

10.2.4.18 ERROR_DFA_UCOND

```
include std/regex.e
public constant ERROR_DFA_UCOND
```

10.2.4.19 ERROR_DFA_UMLIMIT

```
include std/regex.e
public constant ERROR_DFA_UMLIMIT
```

10.2.4.20 ERROR_DFA_WSSIZE

```
include std/regex.e
public constant ERROR_DFA_WSSIZE
```

10.2.4.21 ERROR_DFA_RECURSE

```
include std/regex.e
public constant ERROR_DFA_RECURSE
```

10.2.4.22 ERROR_RECURSIONLIMIT

```
include std/regex.e
public constant ERROR_RECURSIONLIMIT
```

10.2.4.23 ERROR_NULLWSLIMIT

```
include std/regex.e
public constant ERROR_NULLWSLIMIT
```

10.2.4.24 ERROR_BADNEWLINE

```
include std/regex.e
public constant ERROR_BADNEWLINE
```

10.2.5 Create/Destroy

10.2.5.1 regex

```
include std/regex.e
public type regex(object o)
```

Regular expression type

10.2.5.2 new

```
include std/regex.e
public function new(sequence pattern, object options = DEFAULT)
```

Return an allocated regular expression

10.2.5.2.1 Parameters:

1. `pattern`: a sequence representing a human readable regular expression
2. `options`: defaults to [DEFAULT](#). See [Option Constants](#).

10.2.5.2.2 Returns:

A **regex**, which other regular expression routines can work on or an atom to indicate an error. If an error, you can call [error_message](#) to get a detailed error message.

10.2.5.2.3 Comments:

This is the only routine that accepts a human readable regular expression. The string is compiled and a [regex](#) is returned. Analyzing and compiling a regular expression is a costly operation and should not be done more than necessary. For instance, if your application looks for an email address among text frequently, you should create the regular expression as a constant accessible to your source code and any files that may use it, thus, the regular expression is analyzed and compiled only once per run of your application.

-- Bad Example

```
while sequence(line) do
    re:regex proper_name = re:new("[A-Z][a-z]+ [A-Z][a-z]+")
    if re:match(line) then
        -- code
    end if
end while
```

-- Good Example

```
constant re_proper_name = re:new("[A-Z][a-z]+ [A-Z][a-z]+")
while sequence(line) do
    if re:match(line) then
        -- code
    end if
end while
```

10.2.5.2.4 Example 1:

```
include regex.e as re
re:regex number = re:new("[0-9]+")
```

10.2.5.2.5 Note:

For simple finds, matches or even simple wildcard matches, the built-in EUPHORIA routines [find](#), [match](#) and [wildcard_match](#) are often times easier to use and a little faster. Regular expressions are faster for complex searching/matching.

10.2.5.2.6 See Also:

[error_message](#), [find](#), [find_all](#)

10.2.5.3 error_message

```
include std/regex.e
public function error_message(object re)
```

If [new](#) () returns an atom, this function will return a text error message as to the reason.

10.2.5.3.1 Parameters:

1. `re`: Regular expression to get the error message from

10.2.5.3.2 Returns:

An atom (0) when no error message exists, otherwise a sequence describing the error.

10.2.5.3.3 Example 1:

```
object r = regex:new("[A-Z[a-z]*")
if atom(r) then
    printf(1, "Regex failed to compile: %s\n", { regex:error_message(r) })
end if
```

10.2.6 Utility Routines

10.2.6.1 escape

```
include std/regex.e
public function escape(sequence s)
```

Escape special regular expression characters that may be entered into a search string from user input.

10.2.6.1.1 Notes:

10.2.6.1.2 Special regex characters are:

. \ + * ? [^] \$ () { } = ! < > | : -

10.2.6.1.3 Parameters:

1. `s`: string sequence to escape

10.2.6.1.4 Returns:

An escaped sequence representing `s`.

10.2.6.1.5 Example 1:

```
sequence search_s = escape("Payroll is $***15.00")
-- search_s = "Payroll is \\$\\*\\*\\*15\\.00"
```

10.2.6.2 get_ovector_size

```
include std/regex.e
public function get_ovector_size(regex ex, integer maxsize = 0)
```

Returns the number of capturing subpatterns (the ovector size) for a regex

10.2.6.2.1 Parameters:

1. `ex`: a regex
2. `maxsize`: optional maximum number of named groups to get data from

10.2.6.2.2 Returns:

An integer

10.2.7 Find/Match

10.2.7.1 find

```
include std/regex.e
public function find(regex re, sequence haystack, integer from = 1, object options = DEFAULT, integer size = 90)
```

Find the first match of `re` in `haystack`. You can optionally start at the position `from`.

10.2.7.1.1 Parameters:

1. `re` : a regex for a subject to be matched against
2. `haystack` : a string in which to searched
3. `from` : an integer setting the starting position to begin searching from. Defaults to 1
4. `options` : defaults to [DEFAULT](#). See [Option Constants](#).
5. `size` : internal (how large an array the C backend should allocate). Defaults to 90, in rare cases this number may need to be increased in order to accomodate complex regex expressions.

10.2.7.1.2 Returns:

An **object**, which is either an atom of 0, meaning nothing found or a sequence of matched pairs. For the explanation of the returned sequence, please see the first example.

10.2.7.1.3 Example 1:

```
r = re:new("([A-Za-z]+) ([0-9]+)") -- John 20 or Jane 45
object result = re:find(r, "John 20")

-- The return value will be:
-- {
--   { 1, 7 }, -- Total match
--   { 1, 4 }, -- First grouping "John" ([A-Za-z]+)
--   { 6, 7 } -- Second grouping "20" ([0-9]+)
-- }
```

10.2.7.2 find_all

```
include std/regex.e
public function find_all(regex re, sequence haystack, integer from = 1, object options = DEFAULT, integer size = 90)
```

Find all occurrences of `re` in `haystack` optionally starting at the sequence position `from`.

10.2.7.2.1 Parameters:

1. `re` : a regex for a subject to be matched against
2. `haystack` : a string in which to searched
3. `from` : an integer setting the starting position to begin searching from. Defaults to 1
4. `options` : defaults to [DEFAULT](#). See [Option Constants](#).

10.2.7.2.2 Returns:

A **sequence**, of matches. Please see [find](#) for a detailed description of the return value.

10.2.7.2.3 Example 1:

```
constant re_number = re:new("[0-9]+")
object matches = re:find_all(re_number, "10 20 30")

-- matches is:
-- {
--     {1, 2},
--     {4, 5},
--     {7, 8}
-- }
```

10.2.7.3 has_match

```
include std/regex.e
public function has_match(regex re, sequence haystack, integer from = 1, object options = DEFAULT)
```

Determine if `re` matches any portion of `haystack`.

10.2.7.3.1 Parameters:

1. `re` : a regex for a subject to be matched against
2. `haystack` : a string in which to searched
3. `from` : an integer setting the starting position to begin searching from. Defaults to 1
4. `options` : defaults to [DEFAULT](#). See [Option Constants](#).

10.2.7.3.2 Returns:

An **atom**, 1 if `re` matches any portion of `haystack` or 0 if not.

10.2.7.4 is_match

```
include std/regex.e
public function is_match(regex re, sequence haystack, integer from = 1, object options = DEFAULT)
```

Determine if the entire `haystack` matches `re`.

10.2.7.4.1 Parameters:

1. `re` : a regex for a subject to be matched against
2. `haystack` : a string in which to searched
3. `from` : an integer setting the starting position to begin searching from. Defaults to 1
4. `options` : defaults to [DEFAULT](#). See [Option Constants](#).

10.2.7.4.2 Returns:

An **atom**, 1 if `re` matches the entire `haystack` or 0 if not.

10.2.7.5 matches

```
include std/regex.e
public function matches(regex re, sequence haystack, integer from = 1, object options = DEFAULT)
```

Get the matched text only.

10.2.7.5.1 Parameters:

1. `re` : a regex for a subject to be matched against
2. `haystack` : a string in which to searched
3. `from` : an integer setting the starting position to begin searching from. Defaults to 1
4. `options` : defaults to [DEFAULT](#). See [Option Constants](#).

10.2.7.5.2 Returns:

Returns a **sequence**, of strings, the first being the entire match and subsequent items being each of the captured groups. The size of the sequence is the number of groups in the expression plus one (for the entire match).

If `options` contains the bit [STRING_OFFSETS](#), then the result is different. For each item, a sequence is returned containing the matched text, the starting index in `haystack` and the ending index in `haystack`.

10.2.7.5.3 Example 1:

```
constant re_name = re:new("([A-Z][a-z]+) ([A-Z][a-z]+)")

object matches = re:matches(re_name, "John Doe and Jane Doe")
-- matches is:
-- {
--   "John Doe", -- full match data
--   "John",     -- first group
--   "Doe"       -- second group
-- }
```

```
matches = re:matches(re_name, "John Doe and Jane Doe", STRING_OFFSETS)
-- matches is:
-- {
--   { "John Doe", 1, 8 }, -- full match data
--   { "John",     1, 4 }, -- first group
--   { "Doe",      6, 8 }  -- second group
-- }
```

10.2.7.5.4 See Also:

[all_matches](#)

10.2.7.6 all_matches

```
include std/regex.e
public function all_matches(regex re, sequence haystack, integer from = 1, object options = DEF
```

Get the text of all matches

10.2.7.6.1 Parameters:

1. `re` : a regex for a subject to be matched against
2. `haystack` : a string in which to searched
3. `from` : an integer setting the starting position to begin searching from. Defaults to 1
4. `options` : options, defaults to [DEFAULT](#). See [Option Constants](#).

10.2.7.6.2 Returns:

Returns a **sequence**, of a sequence of strings, the first being the entire match and subsequent items being each of the captured groups. The size of the sequence is the number of groups in the expression plus one (for the entire match).

If `options` contains the bit [STRING_OFFSETS](#), then the result is different. For each item, a sequence is returned containing the matched text, the starting index in `haystack` and the ending index in `haystack`.

10.2.7.6.3 Example 1:

```
constant re_name = re:new("([A-Z][a-z]+) ([A-Z][a-z]+)")

object matches = re:match_all(re_name, "John Doe and Jane Doe")
-- matches is:
-- {
--   {
--     "John Doe", -- first match
--     "John Doe", -- full match data
--     "John",     -- first group
--     "Doe",      -- second group
--   },
-- }
```



```

--      {                -- second match
--      "Jane Doe", -- full match data
--      "Jane",     -- first group
--      "Doe"       -- second group
--      }
--  }
-- }

matches = re:match_all(re_name, "John Doe and Jane Doe")
-- matches is:
-- {
--   {                -- first match
--     { "John Doe", 1, 8 }, -- full match data
--     { "John",     1, 4 }, -- first group
--     { "Doe",      6, 8 } -- second group
--   },
--   {                -- second match
--     { "Jane Doe", 14, 21 }, -- full match data
--     { "Jane",     14, 17 }, -- first group
--     { "Doe",      19, 21 } -- second group
--   }
-- }--

```

10.2.7.6.4 See Also:

matches[]

10.2.8 Splitting

10.2.8.1 split

```

include std/regex.e
public function split(regex re, sequence text, integer from = 1, object options = DEFAULT)

```

Split a string based on a regex as a delimiter

10.2.8.1.1 Parameters:

1. `re` : a regex which will be used for matching
2. `text` : a string on which search and replace will apply
3. `from` : optional start position
4. `options` : options, defaults to [DEFAULT](#). See [Option Constants](#).

10.2.8.1.2 Returns:

A **sequence**, of string values split at the delimiter.

10.2.8.1.3 Example 1:

```
regex comma_space_re = re:new(`\s,`)
sequence data = re:split(comma_space_re, "euphoria programming, source code, reference data")
-- data is
-- {
--   "euphoria programming",
--   "source code",
--   "reference data"
-- }
```

10.2.8.2 split_limit

```
include std/regex.e
public function split_limit(regex re, sequence text, integer limit = 0, integer from = 1, object options = 0)
```

10.2.9 Replacement

10.2.9.1 find_replace

```
include std/regex.e
public function find_replace(regex ex, sequence text, sequence replacement, integer from = 1, object options = 0)
```

Replaces all matches of a regex with the replacement text.

10.2.9.1.1 Parameters:

1. `re` : a regex which will be used for matching
2. `text` : a string on which search and replace will apply
3. `replacement` : a string, used to replace each of the full matches found
4. `from` : optional start position
5. `options` : options, defaults to [DEFAULT](#)

10.2.9.1.2 Returns:

A **sequence**, the modified `text`.

10.2.9.1.3 Special replacement operators:

- `\` -- Causes the next character to lose its special meaning.
- `\n` ~ -- Inserts a 0x0A (LF) character.
- `\r` -- Inserts a 0x0D (CR) character.
- `\t` -- Inserts a 0x09 (TAB) character.



- `\1` to `\9` -- Recalls stored substrings from registers (`\1`, `\2`, `\3`, to `\9`).
- `\0` -- Recalls entire matched pattern.
- `\u` -- Convert next character to uppercase
- `\l` -- Convert next character to lowercase
- `\U` -- Convert to uppercase till `\E` or `\e`
- `\L` -- Convert to lowercase till `\E` or `\e`
- `\E` or `\e` -- Terminate a `\U` or `\L` conversion

10.2.9.1.4 Example 1:

```
regex r = new(`([A-Za-z]+)\.([A-Za-z]+)`)
sequence details = find_replace(r, "hello.txt", `Filename: \U\l\e Extension: \U\2\e`)
-- details = "Filename: HELLO Extension: TXT"
```

10.2.9.2 find_replace_limit

```
include std/regex.e
public function find_replace_limit(regex ex, sequence text, sequence replacement, integer limit)
```

Replaces up to `limit` matches of `ex` in `text`.

This function is identical to [find_replace](#) except it allows you to limit the number of replacements to perform. Please see the documentation for [find_replace](#) for all the details.

10.2.9.2.1 Parameters:

1. `re` : a regex which will be used for matching
2. `text` : a string on which search and replace will apply
3. `replacement` : a string, used to replace each of the full matches found
4. `limit` : the number of matches to process
5. `from` : optional start position
6. `options` : options, defaults to [DEFAULT](#)

10.2.9.2.2 Returns:

A `sequence`, the modified `text`.

10.2.9.2.3 See Also:

[find_replace](#)

10.2.9.3 find_replace_callback

```
include std/regex.e
public function find_replace_callback(regex ex, sequence text, integer rid, integer limit = 0,
```

Replaces up to `limit` matches of `ex` in `text` with the result of a user defined callback. The callback should take one sequence which will contain a string representing the entire match and also a string for every group within the regular expression.

10.2.9.3.1 Parameters:

1. `re` : a regex which will be used for matching
2. `text` : a string on which search and replace will apply
3. `rid` : routine id to execute for each match
4. `limit` : the number of matches to process
5. `from` : optional start position
6. `options` : options, defaults to [DEFAULT](#)

10.2.9.3.2 Returns:

A **sequence**, the modified `text`.

10.2.9.3.3 Example 1:

```
function my_convert(sequence params)
    switch params[1] do
        case "1" then
            return "one "
        case "2" then
            return "two "
        case else
            return "unknown "
    end switch
end function

regex r = re:new(`\d`)
sequence result = re:find_replace_callback(r, "125", routine_id("my_convert"))
-- result = "one two unknown "
```

10.3 Text Manipulation

Page Contents

[Routines](#)
[sprintf](#)
[sprint](#)
[trim_head](#)

trim_tail
trim
set_encoding_properties
get_encoding_properties
lower
upper
proper
keyvalues
escape
quote
dequote
format
get_text

10.3.1 Routines

10.3.1.1 sprintf

<built-in> `function sprintf(sequence format, object values)`

This is exactly the same as `printf()`, except that the output is returned as a sequence of characters, rather than being sent to a file or device.

10.3.1.1.1 Parameters:

1. `format` : a sequence, the text to print. This text may contain format specifiers.
2. `values` : usually, a sequence of values. It should have as many elements as format specifiers in `format`, as these values will be substituted to the specifiers.

10.3.1.1.2 Returns:

A **sequence**, of printable characters, representing `format` with the values in `values` spliced in.

10.3.1.1.3 Comments:

`printf(fn, st, x)` is equivalent to `puts(fn, sprintf(st, x))`.

Some typical uses of `sprintf()` are:

1. Converting numbers to strings.
2. Creating strings to pass to `system()`.
3. Creating formatted error messages that can be passed to a common error message handler.

10.3.1.1.4 Example 1:

```
s = sprintf("%08d", 12345)
-- s is "00012345"
```

10.3.1.1.5 See Also:

[printf](#), [sprint](#), [format](#)

10.3.1.2 sprint

```
include std/text.e
public function sprint(object x)
```

Returns the representation of any EUPHORIA object as a string of characters.

10.3.1.2.1 Parameters:

1. *x* : Any EUPHORIA object.

10.3.1.2.2 Returns:

A **sequence**, a string representation of *x*.

10.3.1.2.3 Comments:

This is exactly the same as `print(fn, x)`, except that the output is returned as a sequence of characters, rather than being sent to a file or device. *x* can be any EUPHORIA object.

The atoms contained within *x* will be displayed to a maximum of 10 significant digits, just as with [print\(\)](#).

10.3.1.2.4 Example 1:

```
s = sprint(12345)
-- s is "12345"
```

10.3.1.2.5 Example 2:

```
s = sprint({10,20,30}+5)
-- s is "{15,25,35}"
```

10.3.1.2.6 See Also:

[sprintf](#), [printf](#)

10.3.1.3 trim_head

```
include std/text.e
public function trim_head(sequence source, object what = " \t\r\n", integer ret_index = 0)
```

Trim all items in the supplied set from the leftmost (start or head) of a sequence.

10.3.1.3.1 Parameters:

1. `source` : the sequence to trim.
2. `what` : the set of item to trim from `source` (defaults to " \t\r\n").
3. `ret_index` : If zero (the default) returns the trimmed sequence, otherwise it returns the index of the leftmost item **not** in `what`.

10.3.1.3.2 Returns:

A **sequence**, if `ret_index` is zero, which is the trimmed version of `source`

A **integer**, if `ret_index` is not zero, which is index of the leftmost element in `source` that is not in `what`.

10.3.1.3.3 Example 1:

```
object s
s = trim_head("\r\nSentence read from a file\r\n", "\r\n")
-- s is "Sentence read from a file\r\n"
s = trim_head("\r\nSentence read from a file\r\n", "\r\n", TRUE)
-- s is 3
```

10.3.1.3.4 See Also:

[trim_tail](#), [trim](#), [pad_head](#)

10.3.1.4 trim_tail

```
include std/text.e
public function trim_tail(sequence source, object what = " \t\r\n", integer ret_index = 0)
```

Trim all items in the supplied set from the rightmost (end or tail) of a sequence.

10.3.1.4.1 Parameters:

1. `source` : the sequence to trim.
2. `what` : the set of item to trim from `source` (defaults to `"\t\r\n"`).
3. `ret_index` : If zero (the default) returns the trimmed sequence, otherwise it returns the index of the rightmost item **not** in `what`.

10.3.1.4.2 Returns:

A **sequence**, if `ret_index` is zero, which is the trimmed version of `source`

A **integer**, if `ret_index` is not zero, which is index of the rightmost element in `source` that is not in `what`.

10.3.1.4.3 Example 1:

```
object s
s = trim_tail("\r\nSentence read from a file\r\n", "\r\n")
-- s is "\r\nSentence read from a file"
s = trim_tail("\r\nSentence read from a file\r\n", "\r\n", TRUE)
-- s is 27
```

10.3.1.4.4 See Also:

[trim_head](#), [trim](#), [pad_tail](#)

10.3.1.5 trim

```
include std/text.e
public function trim(sequence source, object what = "\t\r\n", integer ret_index = 0)
```

Trim all items in the supplied set from both the left end (head/start) and right end (tail/end) of a sequence.

10.3.1.5.1 Parameters:

1. `source` : the sequence to trim.
2. `what` : the set of item to trim from `source` (defaults to `"\t\r\n"`).
3. `ret_index` : If zero (the default) returns the trimmed sequence, otherwise it returns a 2-element sequence containing the index of the leftmost item and rightmost item **not** in `what`.

10.3.1.5.2 Returns:

A **sequence**, if `ret_index` is zero, which is the trimmed version of `source`

A **2-element sequence**, if `ret_index` is not zero, in the form `{left_index, right_index}`.

10.3.1.5.3 Example 1:

```
object s
s = trim("\r\nSentence read from a file\r\n", "\r\n")
-- s is "Sentence read from a file"
s = trim("\r\nSentence read from a file\r\n", "\r\n", TRUE)
-- s is {3,27}
```

10.3.1.5.4 See Also:

[trim_head](#), [trim_tail](#)

10.3.1.6 set_encoding_properties

```
include std/text.e
public procedure set_encoding_properties(sequence en = "", sequence lc = "", sequence uc = "")
```

Sets the table of lowercase and uppercase characters that is used by [lower](#) and [upper](#)

10.3.1.6.1 Parameters:

1. `en` : The name of the encoding represented by these character sets
2. `lc` : The set of lowercase characters
3. `uc` : The set of upper case characters

10.3.1.6.2 Comments:

- `lc` and `uc` must be the same length.
- If no parameters are given, the default ASCII table is set.

10.3.1.6.3 Example 1:

```
set_encoding_properties( "Elvish", "aeiouy", "AEIOUY")
```

10.3.1.6.4 Example 1:

```
set_encoding_properties( "1251") -- Loads a predefined code page.
```

10.3.1.6.5 See Also:

[lower](#), [upper](#), [get_encoding_properties](#)

10.3.1.7 get_encoding_properties

```
include std/text.e
public function get_encoding_properties()
```

Gets the table of lowercase and uppercase characters that is used by [lower](#) and [upper](#)

10.3.1.7.1 Parameters:

none

10.3.1.7.2 Returns:

A **sequence**, containing three items.
{Encoding_Name, LowerCase_Set, UpperCase_Set}

10.3.1.7.3 Example 1:

```
encode_sets = get_encoding_properties()
```

10.3.1.7.4 See Also:

[lower](#), [upper](#), [set_encoding_properties](#)

10.3.1.8 lower

```
include std/text.e
public function lower(object x)
```

Convert an atom or sequence to lower case.

10.3.1.8.1 Parameters:

1. x : Any EUPHORIA object.

10.3.1.8.2 Returns:

A **sequence**, the lowercase version of x

10.3.1.8.3 Comments:

- For Windows systems, this uses the current code page for conversion
- For non-Windows, this only works on ASCII characters. It alters characters in the 'a'..'z' range. If you need to do case conversion with other encodings use the [set_encoding_properties](#) first.
- x may be a sequence of any shape, all atoms of which will be acted upon.

WARNING, When using ASCII encoding, this can also affects floating point numbers in the range 65 to 90.

10.3.1.8.4 Example 1:

```
s = lower("EUPHORIA")
-- s is "euphoria"

a = lower('B')
-- a is 'b'

s = lower({"EUPHORIA", "Programming"})
-- s is {"euphoria", "programming"}
```

10.3.1.8.5 See Also:

[upper](#), [proper](#), [set_encoding_properties](#), [get_encoding_properties](#)

10.3.1.9 upper

```
include std/text.e
public function upper(object x)
```

Convert an atom or sequence to upper case.

10.3.1.9.1 Parameters:

1. x : Any EUPHORIA object.

10.3.1.9.2 Returns:

A **sequence**, the uppercase version of x

10.3.1.9.3 Comments:

- For Windows systems, this uses the current code page for conversion
- For non-Windows, this only works on ASCII characters. It alters characters in the 'a'..'z' range. If you need to do case conversion with other encodings use the [set_encoding_properties](#) first.

- x may be a sequence of any shape, all atoms of which will be acted upon.

WARNING, When using ASCII encoding, this can also affects floating point numbers in the range 97 to 122.

10.3.1.9.4 Example 1:

```
s = upper("EUPHORIA")
-- s is "EUPHORIA"

a = upper('b')
-- a is 'B'

s = upper({"EUPHORIA", "Programming"})
-- s is {"EUPHORIA", "PROGRAMMING"}
```

10.3.1.9.5 See Also:

[lower](#), [proper](#), [set_encoding_properties](#), [get_encoding_properties](#)

10.3.1.10 proper

```
include std/text.e
public function proper(sequence x)
```

Convert a text sequence to capitalized words.

10.3.1.10.1 Parameters:

1. x : A text sequence.

10.3.1.10.2 Returns:

A **sequence**, the Capitalized Version of x

10.3.1.10.3 Comments:

A text sequence is one in which all elements are either characters or text sequences. This means that if a non-character is found in the input, it is not converted. However this rule only applies to elements on the same level, meaning that sub-sequences could be converted if they are actually text sequences.

10.3.1.10.4 Example 1:

```
s = proper("euphoria programming language")
-- s is "EUPHORIA Programming Language"
s = proper("EUPHORIA PROGRAMMING LANGUAGE")
-- s is "EUPHORIA Programming Language"
s = proper({"EUPHORIA PROGRAMMING", "language", "rapid dEPLOYMENT", "sOfTwArE"})
-- s is {"EUPHORIA Programming", "Language", "Rapid Deployment", "Software"}
s = proper({'a', 'b', 'c'})
-- s is {'A', 'b', 'c'} -- "Abc"
s = proper({'a', 'b', 'c', 3.1472})
-- s is {'a', 'b', 'c', 3.1472} -- Unchanged because it contains a non-character.
s = proper({"abc", 3.1472})
-- s is {"Abc", 3.1472} -- The embedded text sequence is converted.
```

10.3.1.10.5 See Also:

[lower upper](#)

10.3.1.11 keyvalues

```
include std/text.e
public function keyvalues(sequence source, object pair_delim = ";;", object kv_delim = ":", ob
```

Converts a string containing Key/Value pairs into a set of sequences, one per K/V pair.

10.3.1.11.1 Parameters:

1. `source` : a text sequence, containing the representation of the key/values.
2. `pair_delim` : an object containing a list of elements that delimit one key/value pair from the next.
The defaults are semi-colon (;) and comma (,).
3. `kv_delim` : an object containing a list of elements that delimit the key from its value. The defaults are colon (:) and equal (=).
4. `quotes` : an object containing a list of elements that can be used to enclose either keys or values that contain delimiters or whitespace. The defaults are double-quote ("), single-quote (') and back-quote (`).
5. `whitespace` : an object containing a list of elements that are regarded as whitespace characters.
The defaults are space, tab, new-line, and carriage-return.
6. `haskeys` : an integer containing true or false. The default is true. When true, the `kv_delim` values are used to separate keys from values, but when false it is assumed that each 'pair' is actually just a value.

10.3.1.11.2 Returns:

A **sequence**, of pairs. Each pair is in the form {key, value}.

10.3.1.11.3 Comments:

String representations of atoms are not converted, either in the key or value part, but returned as any regular string instead.

If `haskeys` is `true`, but a substring only holds what appears to be a value, the key is synthesized as `p[n]`, where `n` is the number of the pair. See example #2.

By default, pairs can be delimited by either a comma or semi-colon `","` and a key is delimited from its value by either an equal or a colon `"="`. Whitespace between pairs, and between delimiters is ignored.

If you need to have one of the delimiters in the value data, enclose it in quotation marks. You can use any of single, double and back quotes, which also means you can quote quotation marks themselves. See example #3.

It is possible that the value data itself is a nested set of pairs. To do this enclose the value in parentheses. Nested sets can nested to any level. See example #4.

If a sub-list has only data values and not keys, enclose it in either braces or square brackets. See example #5. If you need to have a bracket as the first character in a data value, prefix it with a tilde. Actually a leading tilde will always just be stripped off regardless of what it prefixes. See example #6.

10.3.1.11.4 Example 1:

```
s = keyvalues("foo=bar, qwe=1234, asdf='contains space, comma, and equal(=)')
-- s is { {"foo", "bar"}, {"qwe", "1234"}, {"asdf", "contains space, comma, and equal(=)"} }
```

10.3.1.11.5 Example 2:

```
s = keyvalues("abc fgh=ijk def")
-- s is { {"p[1]", "abc"}, {"fgh", "ijk"}, {"p[3]", "def"} }
```

10.3.1.11.6 Example 3:

```
s = keyvalues("abc=`'quoted'`")
-- s is { {"abc", "'quoted'"} }
```

10.3.1.11.7 Example 4:

```
s = keyvalues("colors=(a=black, b=blue, c=red)")
-- s is { {"colors", {{"a", "black"}, {"b", "blue"}, {"c", "red"}} } }
s = keyvalues("colors=(black=[0,0,0], blue=[0,0,FF], red=[FF,0,0])")
-- s is { {"colors", {{"black", {"0", "0", "0"}}, {"blue", {"0", "0", "FF"}}, {"red", {"FF", "0", "0"}}
```

10.3.1.11.8 Example 5:

```
s = keyvalues("colors=[black, blue, red]")
-- s is { {"colors", { "black", "blue", "red"} } }
```

10.3.1.11.9 Example 6:

```
s = keyvalues("colors=~[black, blue, red]")
-- s is { {"colors", "[black, blue, red]"} } }
-- The following is another way to do the same.
s = keyvalues("colors=`[black, blue, red]`")
-- s is { {"colors", "[black, blue, red]"} } }
```

10.3.1.12 escape

```
include std/text.e
public function escape(sequence s, sequence what = "\"")
```

Escape special characters in a string

10.3.1.12.1 Parameters:

1. s: string to escape
2. what: sequence of characters to escape defaults to escaping a double quote.

10.3.1.12.2 Returns:

An escaped sequence representing s.

10.3.1.12.3 Example 1:

```
sequence s = escape("John \"Mc\" Doe")
puts(1, s)
-- output is: John \"Mc\" Doe
```

10.3.1.12.4 See Also:

[quote](#)

10.3.1.13 quote

```
include std/text.e
public function quote(sequence text_in, object quote_pair = {"\"", "\""}, integer esc = - 1, t_
```

Return a quoted version of the first argument.

10.3.1.13.1 Parameters:

1. `text_in` : The string or set of strings to quote.
2. `quote_pair` : A sequence of two strings. The first string is the opening quote to use, and the second string is the closing quote to use. The default is {"\"", "\""} which means that the output will be enclosed by double-quotation marks.
3. `esc` : A single escape character. If this is not negative (the default), then this is used to 'escape' any embedded quote characters and 'esc' characters already in the `text_in` string.
4. `sp` : A list of zero or more special characters. The `text_in` is only quoted if it contains any of the special characters. The default is "" which means that the `text_in` is always quoted.

10.3.1.13.2 Returns:

A **sequence**, the quoted version of `text_in`.

10.3.1.13.3 Example 1:

```
-- Using the defaults. Output enclosed in double-quotes, no escapes and no specials.
s = quote("The small man")
-- 's' now contains '"the small man"' including the double-quote characters.
```

10.3.1.13.4 Example 2:

```
s = quote("The small man", {"(", ")"})
-- 's' now contains '(the small man)'
```

10.3.1.13.5 Example 3:

```
s = quote("The (small) man", {"(", ")"}, '~' )
-- 's' now contains '(the ~(small~) man)'
```

10.3.1.13.6 Example 4:

```
s = quote("The (small) man", {"(", ")"}, '~', "#" )
-- 's' now contains "the (small) man"
-- because the input did not contain a '#' character.
```

10.3.1.13.7 Example 5:

```
s = quote("The #1 (small) man", {"(", ")"}, '~', "#" )
-- 's' now contains '(the #1 ~(small~) man)'
-- because the input did contain a '#' character.
```

10.3.1.13.8 Example 6:

```
-- input is a set of strings...
s = quote({"a b c", "def", "g hi"},)
-- 's' now contains three quoted strings: '"a b c"', '"def"', and '"g hi"'
```

10.3.1.13.9 See Also:

[escape](#)

10.3.1.14 dequote

```
include std/text.e
public function dequote(object text_in, sequence quote_pairs = {"\"", "\""}, integer esc = -
```

Removes 'quotation' text from the argument.

10.3.1.14.1 Parameters:

1. `text_in` : The string or set of strings to de-quote.
2. `quote_pairs` : A set of one or more sub-sequences of two strings. The first string in each sub-sequence is the opening quote to look for, and the second string is the closing quote. The default is "\"", "\"" which means that the output is 'quoted' if it is enclosed by double-quotation marks.
3. `esc` : A single escape character. If this is not negative (the default), then this is used to 'escape' any embedded occurrences of the quote characters. In which case the 'escape' character is also removed.

10.3.1.14.2 Returns:

A **sequence**, the original text but with 'quote' strings stripped of quotes.

10.3.1.14.3 Example 1:

```
-- Using the defaults.
s = quote("\"The small man\"")
-- 's' now contains "The small man"
```

10.3.1.14.4 Example 2:

```
-- Using the defaults.
s = quote("(The small ?(?) man)", {"[", "]"}, '?')
-- 's' now contains "The small () man"
```

10.3.1.15 format

```
include std/text.e
public function format(sequence format_pattern, object arg_list = {})
```

Formats a set of arguments in to a string based on a supplied pattern.

10.3.1.15.1 Parameters:

1. `format_pattern`: A sequence: the pattern string that contains zero or more tokens.
2. `arg_list`: An object: Zero or more arguments used in token replacement.

10.3.1.15.2 Returns:

A string **sequence**, the original `format_pattern` but with tokens replaced by corresponding arguments.

10.3.1.15.3 Comments:

The `format_pattern` string contains text and argument tokens. The resulting string is the same as the `format` string except that each token is replaced by an item from the argument list.

A token has the form `[<Q>]`, where `<Q>` is are optional qualifier codes.

The qualifier. `<Q>` is a set of zero or more codes that modify the default way that the argument is used to replace the token. The default replacement method is to convert the argument to its shortest string representation and use that to replace the token. This may be modified by the following codes, which can occur in any order.

Qualifier	Usage
N	('N' is an integer) The index of the argument to use
{id}	Uses the argument that begins with "id=" where "id" is an identifier name.
%envvar%	Uses the Environment Symbol 'envvar' as an argument
w	For string arguments, if capitalizes the first letter in each word
u	For string arguments, it converts it to upper case.
l	For string arguments, it converts it to lower case.
<	For numeric arguments, it left justifies it.
>	For string arguments, it right justifies it.
c	Centers the argument.
z	For numbers, it zero fills the left side.
:S	('S' is an integer) The maximum size of the resulting field. Also, if 'S' begins with '0' the field will be zero-filled if the argument is an integer

.N	('N' is an integer) The number of digits after the decimal point
+	For positive numbers, show a leading plus sign
(For negative numbers, enclose them in parentheses
b	For numbers, causes zero to be all blanks
s	If the resulting field would otherwise be zero length, this ensures that at least one space occurs between this token's field
t	After token replacement, the resulting string up to this point is trimmed.
X	Outputs integer arguments using hexadecimal digits.
B	Outputs integer arguments using binary digits.
?	The corresponding argument is a set of two strings. This uses the first string if the previous token's argument is not the value 1 or a zero-length string, otherwise it uses the second string.
[Does not use any argument. Outputs a left-square-bracket symbol
,	Insert thousands separators. The <X> is the character to use. If this is a dot "." then the decimal point is rendered using a comma. Does not apply to zero-filled fields.
X	N.B. if hex or binary output was specified, the separators are every 4 digits otherwise they are every three digits.

Clearly, certain combinations of these qualifier codes do not make sense and in those situations, the rightmost clashing code is used and the others are ignored.

Any tokens in the format that have no corresponding argument are simply removed from the result. Any arguments that are not used in the result are ignored.

Any sequence argument that is not a string will be converted to its *pretty* format before being used in token replacement.

If a token is going to be replaced by a zero-length argument, all white space following the token until the next non-whitespace character is not copied to the result string.

10.3.1.15.4 Examples:

```
format("Cannot open file '[' - code []", {"/usr/temp/work.dat", 32})
-- "Cannot open file '/usr/temp/work.dat' - code 32"

format("Err-[2], Cannot open file '[1]'", {"/usr/temp/work.dat", 32})
-- "Err-32, Cannot open file '/usr/temp/work.dat'"

format("[4w] [3z:2] [6] [5l] [2z:2], [1:4]", {2009,4,21,"DAY","MONTH","of"})
-- "Day 21 of month 04, 2009"

format("The answer is [:6.2]%", {35.22341})
-- "The answer is 35.22%"
```

```

format("The answer is [.6]", {1.2345})
-- "The answer is 1.234500"

format("The answer is [,.2]", {1234.56})
-- "The answer is 1,234.56"

format("The answer is [,.2]", {1234.56})
-- "The answer is 1.234,56"

format("The answer is [,:.2]", {1234.56})
-- "The answer is 1:234.56"

format("[ ] [?]", {5, {"cats", "cat"}})
-- "5 cats"

format("[ ] [?]", {1, {"cats", "cat"}})
-- "1 cat"

format("[<:4]", {"abcdef"})
-- "abcd"

format("[>:4]", {"abcdef"})
-- "cdef"

format("[>:8]", {"abcdef"})
-- " abcdef"

format("seq is [ ]", {{1.2, 5, "abcdef", {3}}})
-- `seq is {1.2,5,"abcdef",{3}}`

format("Today is [{day}], the [{date}]", {"date=10/Oct/2012", "day=Wednesday"})
-- "Today is Wednesday, the 10/Oct/2012"

```

10.3.1.15.5 See Also:

[sprintf](#)

10.3.1.16 get_text

```

include std/text.e
public function get_text(integer MsgNum, sequence LocalQuals = {}, sequence DBBase = "teksto")

```

Get the text associated with the message number in the requested locale.

10.3.1.16.1 Parameters:

1. `MsgNum`: An integer. The message number whose text you are trying to get.
2. `LocalQuals`: A sequence. Zero or more locale codes. Default is {}.
3. `DBBase`: A sequence. The base name for the database files containing the locale text strings. The default is "teksto".

10.3.1.16.2 Returns:

A string **sequence**, the text associated with the message number and locale.
An **integer**, if not associated text can be found.

10.3.1.16.3 Comments:

- This first scans the database(s) linked to the locale codes supplied.
- The database name for each locale takes the format of "<DBBase>_<Locale>.edb" so if the default DBBase is used, and the locales supplied are {"enus", "enau"} the databases scanned are "teksto_enus.edb" and "teksto_enau.edb". The database table name searched is "1" with the key being the message number, and the text is the record data.
- If the message is not found in these databases (or the databases don't exist) a database called "<DBBase>.edb" is searched. Again the table name is "1" but it first looks for keys with the format {<locale>,msgnum} and failing that it looks for keys in the format {"", msgnum}, and if that fails it looks for a key of just the msgnum.

10.4 Unicode

Unicode is not complete. This API will change, even in 4.1 release.

By default, EUPHORIA strings contain UCS-4 code points (characters). **Page Contents**

[isUChar](#)
[utf_8](#)
[utf_16](#)
[utf_32](#)
[BOM_8](#)
[BOM_16be](#)
[BOM_16le](#)
[BOM_32be](#)
[BOM_32le](#)
[get_seqlen](#)
[chars_before](#)
[char_count](#)
[get_char](#)
[encode](#)
[code_length](#)
[validate](#)
[toUTF](#)
[isAlpha](#)
[isUpper](#)
[isLower](#)
[isTitle](#)
[isAlphaNum](#)
[isDigit](#)
[isNumber](#)

isSeparator
isSpace
isLine
isParagraph
isMark
isNonSpacing
isPunctuation
isSymbol
isCurrency
isOther
isControl
isFormat
isSurrogate
isPrivate
isGraph
isPrint
isDirWhiteSpace
isDirLTR
isDirRTL
isDirStrong
isDirWeak
isDirNeutral
isDirSeparator
isBlock
isSegment
isNonBreaking
isMirroring
toLower
toUpper
toTitle
toMirror
getNumericValue

10.4.1 isUChar

```
include std/unicode.e  
public function isUChar(object test_char, integer is_strict = 0)
```

Tests if a value is a valid code point.

10.4.1.1 Parameters:

1. `test_char` : an object, the value to test.
2. `is_strict`: an integer, if non-zero then non-character code points are rejected. The default (zero) means that non-character code points are accepted.

10.4.1.1.1 Returns:

An integer. 1 means that `test_char` is valid and 0 means that it is not.

10.4.1.1.2 Example:

```
for i = 1 to length(s) do
  if not isUChar(s[i]) then
    return 0 -- String is not a utf-32 string.
  end if
end for
```

10.4.1.2 utf_8

```
include std/unicode.e
public enum utf_8
```

10.4.1.3 utf_16

```
include std/unicode.e
public enum utf_16
```

10.4.1.4 utf_32

```
include std/unicode.e
public enum utf_32
```

10.4.1.5 BOM_8

```
include std/unicode.e
public constant BOM_8
```

10.4.1.6 BOM_16be

```
include std/unicode.e
public constant BOM_16be
```

10.4.1.7 BOM_16le

```
include std/unicode.e
public constant BOM_16le
```

10.4.1.8 BOM_32be

```
include std/unicode.e
public constant BOM_32be
```

10.4.1.9 BOM_32le

```
include std/unicode.e
public constant BOM_32le
```

10.4.1.10 get_seqlen

```
include std/unicode.e
public function get_seqlen(sequence input_string, integer from_index, integer encoding_format =
```

Returns the length of an encoded UTF sequence.

10.4.1.10.1 Parameters:

1. `input_string`: a sequence. Contains a UTF encoded string.
2. `from_index`: an integer. The index of a code sequence starting point in `input_string`.
3. `encoding_format`: One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_8`.

10.4.1.10.2 Returns:

An integer. The number of elements in `input_string`, starting from `from_index` that form the code sequence. If 255 (0xFF) is returned, it means that either `from_index` was not at the start of a code sequence, or an invalid `encoding_format` parameter was supplied.

10.4.1.10.3 Example:

```
? get_seqlen("abc", 2) --> 1
```

10.4.1.11 chars_before

```
include std/unicode.e
public function chars_before(sequence input_string, integer from_index, integer encoding_format)
```

Returns the length of an encoded UTF sequence.

10.4.1.11.1 Parameters:

1. `input_string`: a sequence. Contains a UTF encoded string.
2. `from_index`: an integer. The index of a code sequence starting point in `pString`.
3. `encoding_format`: One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_8`.

10.4.1.11.2 Returns:

An integer. The number of UCS characters in `input_string` that occur before `from_index`. However, if `from_index` is invalid it returns -1, and if an invalid UTF sequence is found, it returns -2.

10.4.1.11.3 Example:

```
? chars_before(x"7a C2A9 E6B0B4 F09d849e", 4) --> 2
```

10.4.1.12 char_count

```
include std/unicode.e
public function char_count(sequence input_string, integer encoding_format = utf_8)
```

Returns the number of UCS characters in a UTF string.

10.4.1.12.1 Parameters:

1. `input_string`: a sequence. Contains a UTF encoded string.
2. `encoding_format`: One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_8`.

10.4.1.12.2 Returns:

An integer. The number of UCS characters in `input_string`. However, if an invalid UTF sequence is found it returns -2.

10.4.1.12.3 Example:

```
? char_count(x"7a C2A9 E6B0B4 F09d849e") --> 4
```

10.4.1.13 get_char

```
include std/unicode.e
public function get_char(sequence input_string, integer from_index, integer encoding_format = u
```

Returns the UCS character from a specific location in a UTF string

10.4.1.13.1 Parameters:

1. `input_string`: a sequence. Contains a UTF encoded string.
2. `from_index`: an integer. The index of a code sequence starting point in `input_string`.
3. `encoding_format`: One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_8`.
4. `is_strict`: an integer. If non-zero, then non-character code points are deemed to be invalid characters. The default is zero, which allows non-character code points.

10.4.1.13.2 Returns:

- If `from_index` points to a valid character then a sequence is returned. The first element is the code point value of the encoded USC character in `input_string` at `from_index`, and the second element is the index of the next character in `input_string`.
- If `from_index` does not point to a valid character then an integer is returned.
 - ◆ 1 An element in the coded character is invalid.
 - ◆ 2 In `utf_8` encodings, the leading encoded byte is not valid.
 - ◆ 3 The value of `from_index` is invalid or the length of `input_string` is too short.
 - ◆ 4 In `utf_8` encodings, the non-leading byte is not valid
 - ◆ 5 The UCS character is not a valid one.
 - ◆ 6 The `encoding_format` parameter is not valid
 - ◆ 7 For `utf_16` encodings, the leading surrogate is invalid
 - ◆ 8 For `utf_16` encodings, the trailing surrogate is invalid

10.4.1.13.3 Example:

```
? get_char(u"d834 dd1e", 1, utf_16) --> 0x1D11E
```

10.4.1.14 encode

```
include std/unicode.e
public function encode(atom ucs_char, integer encoding_format = utf_8)
```

Returns the UTF string representation of a UCS character (code point)

10.4.1.14.1 Parameters:

1. `ucs_char`: An atom. The UCS character, also known as a code point, to encode.
2. `encoding_format`: The format at return. One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_8`.

10.4.1.14.2 Returns:

A sequence. The UTF encoding form of the `ucs_char`. Note that if the UCS value supplied is not a valid character, or `encoding_format` is invalid, an empty sequence is returned.

10.4.1.14.3 Example:

```
atom G_Clef = 0x1d11e
? encode(G_Clef, utf_8) --> x"f09d849e"
```

10.4.1.15 code_length

```
include std/unicode.e
public function code_length(atom ucs_char, integer encoding_format = utf_8)
```

Returns the length of the UTF string representation of a UCS character (code point)

10.4.1.15.1 Parameters:

1. `ucs_char`: An atom. The UCS character, also known as a code point.
2. `encoding_format`: The format at return. One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_8`.

10.4.1.15.2 Returns:

An integer. The number of elements to make up the UTF encoded form of `ucs_char`. Note that if the UCS value supplied is not a valid character, or `encoding_format` is invalid, zero is returned.

10.4.1.15.3 Example:

```
atom G_Clef = 0x1d11e
? encode(G_Clef, utf_8) --> 4
```

10.4.1.16 validate

```
include std/unicode.e
public function validate(object input_string, integer encoding_format = utf_8, integer is_strict = 0)
```

Returns the UTF string representation of a UCS character (code point)

10.4.1.16.1 Parameters:

1. `input_string`: A sequence. The UTF encoded text to validate.
2. `encoding_format`: The format of `input_string`. One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_8`.
3. `is_strict`: an integer. If non-zero, then non-character code points are deemed to be invalid characters. The default is zero, which allows non-character code points.

10.4.1.16.2 Returns:

An integer.

- 0 means that the text is valid for the given format.
- `n > 0` is the index into `input_string` where the first invalid element was found.

10.4.1.16.3 Example:

```
? validate(x"7aE289", utf_8) --> 2
```

10.4.1.17 toUTF

```
include std/unicode.e
public function toUTF(sequence input_string, integer source_encoding = utf_32, integer result_encoding = utf_8)
```

Converts the UTF encoding of the given text to another UTF encoding format.

10.4.1.17.1 Parameters:

1. `input_string`: A sequence. The UTF encoded text to convert.
2. `source_encoding`: The format of `input_string`. One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_32`.
3. `result_encoding`: The format of returned result. One of `utf_8`, `utf_16`, or `utf_32`. The default is `utf_8`.

10.4.1.17.2 Returns:

- If `input_string` is not a valid UTF string, this returns the index into it where the first invalid element was found.
- Otherwise it returns a sequence, which is the converted form of `input_string`.

10.4.1.17.3 Example:

```
? toUTF(x"7a C2A9 E6B0B4 F09d849e", utf_8, utf_16) --> u"7a A9 6c34 d834 dd1e"
```

10.4.1.18 isAlpha

```
include std/ucstypes.e
public function isAlpha(atom ucs_char)
```

The code point is a 'letter'.

10.4.1.18.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.18.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.19 isUpper

```
include std/ucstypes.e
public function isUpper(atom ucs_char)
```

The code point is an uppercase 'letter'.

10.4.1.19.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.19.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.20 isLower

```
include std/ucstypes.e
public function isLower(atom ucs_char)
```

The code point is a lowercase 'letter'.

10.4.1.20.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.20.2 Returns:

TRUE when ucs_char is in the set, FALSE otherwise.

10.4.1.21 isTitle

```
include std/ucstypes.e
public function isTitle(atom ucs_char)
```

The code point is a Title case 'letter'.

10.4.1.21.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.21.2 Returns:

TRUE when ucs_char is in the set, FALSE otherwise.

10.4.1.22 isAlphaNum

```
include std/ucstypes.e
public function isAlphaNum(atom ucs_char)
```

The code point is either a 'letter' or a 'digit'.

10.4.1.22.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.22.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.23 isDigit

```
include std/ucstypes.e
public function isDigit(atom ucs_char)
```

The code point is a 'digit'.

10.4.1.23.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.23.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.24 isNumber

```
include std/ucstypes.e
public function isNumber(atom ucs_char)
```

The code point is a 'number'. Some code points represent fractional numbers.

10.4.1.24.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.24.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.25 isSeparator

```
include std/ucstypes.e
public function isSeparator(atom ucs_char)
```

The code point is a word separator.

10.4.1.25.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.25.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.26 isSpace

```
include std/ucstypes.e
public function isSpace(atom ucs_char)
```

The code point is a 'space' character.

10.4.1.26.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.26.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.27 isLine

```
include std/ucstypes.e
public function isLine(atom ucs_char)
```

The code point is a line separator.

10.4.1.27.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.27.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.28 isParagraph

```
include std/ucstypes.e
public function isParagraph(atom ucs_char)
```

The code point is a paragraph separator

10.4.1.28.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.28.2 Returns:

TRUE when ucs_char is in the set, FALSE otherwise.

10.4.1.29 isMark

```
include std/ucstypes.e
public function isMark(atom ucs_char)
```

The code point is a textual 'marker'.

10.4.1.29.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.29.2 Returns:

TRUE when ucs_char is in the set, FALSE otherwise.

10.4.1.30 isNonSpacing

```
include std/ucstypes.e
public function isNonSpacing(atom ucs_char)
```

The code point is a non-spacing letter.

10.4.1.30.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.30.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.31 isPunctuation

```
include std/ucstypes.e
public function isPunctuation(atom ucs_char)
```

The code point is a punctuation mark.

10.4.1.31.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.31.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.32 isSymbol

```
include std/ucstypes.e
public function isSymbol(atom ucs_char)
```

The code point is a 'symbol'.

10.4.1.32.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.32.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.33 isCurrency

```
include std/ucstypes.e
public function isCurrency(atom ucs_char)
```

The code point is a 'symbol'.

10.4.1.33.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.33.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.34 isOther

```
include std/ucstypes.e
public function isOther(atom ucs_char)
```

The code point is a functional item, such as a control character, or private-use.

10.4.1.34.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.34.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.35 isControl

```
include std/ucstypes.e
public function isControl(atom ucs_char)
```

The code point is a control item.

10.4.1.35.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.35.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.36 isFormat

```
include std/ucstypes.e
public function isFormat(atom ucs_char)
```

The code point is a formatting item.

10.4.1.36.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.36.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.37 isSurrogate

```
include std/ucstypes.e
public function isSurrogate(atom ucs_char)
```

The code point is a surrogate code.

10.4.1.37.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.37.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.38 isPrivate

```
include std/ucstypes.e
public function isPrivate(atom ucs_char)
```

The code point is a private-use item.

10.4.1.38.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.38.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.39 isGraph

```
include std/ucstypes.e
public function isGraph(atom ucs_char)
```

The code point is a character with a glyph.

10.4.1.39.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.39.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.40 isPrint

```
include std/ucstypes.e
public function isPrint(atom ucs_char)
```

The code point is a character that is printable.

10.4.1.40.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.40.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.41 isDirWhiteSpace

```
include std/ucstypes.e
public function isDirWhiteSpace(atom ucs_char)
```

The code point is a directional white space character.

10.4.1.41.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.41.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.42 isDirLTR

```
include std/ucstypes.e  
public function isDirLTR(atom ucs_char)
```

The code point is a left-to-right character.

10.4.1.42.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.42.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.43 isDirRTL

```
include std/ucstypes.e  
public function isDirRTL(atom ucs_char)
```

The code point is a right-to-left character.

10.4.1.43.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.43.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.44 isDirStrong

```
include std/ucstypes.e
public function isDirStrong(atom ucs_char)
```

The code point has an exact directional aspect. It must be always be LtR or RtL.

10.4.1.44.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.44.2 Returns:

TRUE when ucs_char is in the set, FALSE otherwise.

10.4.1.45 isDirWeak

```
include std/ucstypes.e
public function isDirWeak(atom ucs_char)
```

The code point has an implied directional aspect, that depends on context.

10.4.1.45.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.45.2 Returns:

TRUE when ucs_char is in the set, FALSE otherwise.

10.4.1.46 isDirNeutral

```
include std/ucstypes.e
public function isDirNeutral(atom ucs_char)
```

The code point has no directional aspect.

10.4.1.46.1 Parameters:

1. ucs_char: An atom. The code point to test.

10.4.1.46.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.47 isDirSeparator

```
include std/ucstypes.e
public function isDirSeparator(atom ucs_char)
```

The code point is a directional separator.

10.4.1.47.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.47.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.48 isBlock

```
include std/ucstypes.e
public function isBlock(atom ucs_char)
```

The code point is a block separator.

10.4.1.48.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.48.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.49 isSegment

```
include std/ucstypes.e
public function isSegment(atom ucs_char)
```

The code point is a segment separator.

10.4.1.49.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.49.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.50 isNonBreaking

```
include std/ucstypes.e
public function isNonBreaking(atom ucs_char)
```

The code point is a non-breaking character.

10.4.1.50.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.50.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.51 isMirroring

```
include std/ucstypes.e
public function isMirroring(atom ucs_char)
```

The code point has a mirror (matching) character, such as parenthesis.

10.4.1.51.1 Parameters:

1. `ucs_char`: An atom. The code point to test.

10.4.1.51.2 Returns:

TRUE when `ucs_char` is in the set, FALSE otherwise.

10.4.1.52 toLower

```
include std/ucstypes.e
public function toLower(object ucs_char)
```

Converts the input to lower case.

10.4.1.52.1 Parameters:

1. `ucs_char`: An object. Either a single code point to convert or a text string.

10.4.1.52.2 Returns:

The converted input.

10.4.1.53 toUpper

```
include std/ucstypes.e
public function toUpper(object ucs_char)
```

Converts the input to upperer case.

10.4.1.53.1 Parameters:

1. `ucs_char`: An object. Either a single code point to convert or a text string.

10.4.1.53.2 Returns:

The converted input.

10.4.1.54 toTitle

```
include std/ucstypes.e
public function toTitle(atom ucs_char)
```

Converts the input to lower case.

10.4.1.54.1 Parameters:

1. `ucs_char`: An atom. The code point to convert.

10.4.1.54.2 Returns:

The converted input.

10.4.1.55 toMirror

```
include std/ucstypes.e
public function toMirror(atom ucs_char)
```

Gets the matching character for the supplied code point.

10.4.1.55.1 Parameters:

1. `ucs_char`: An atom. The code point to match.

10.4.1.55.2 Returns:

If the code point has no matching value, the code point is returned, otherwise the matching (mirror) code point is returned.

10.4.1.56 getNumericValue

```
include std/ucstypes.e
public function getNumericValue(atom ucs_char)
```

Gets the numerical value of the character.

10.4.1.56.1 Parameters:

1. `ucs_char`: An atom. The code point to convert.

10.4.1.56.2 Returns:

An atom: -1 if the code point has no numerical value, -2 if the value is not known, otherwise the numerical value of it.

10.5 Wildcard Matching

Page Contents

[Routines](#)

wildcard_match
wildcard_file

10.5.1 Routines

10.5.1.1 wildcard_match

```
include std/wildcard.e  
public function wildcard_match(sequence pattern, sequence string)
```

Determine whether a string matches a pattern. The pattern may contain * and ? wildcards.

10.5.1.1.1 Parameters:

1. `pattern`: a string, the pattern to match
2. `string`: the string to be matched against

10.5.1.1.2 Returns:

An **integer**, TRUE if `string` matches `pattern`, else FALSE.

10.5.1.1.3 Comments:

Character comparisons are case sensitive. If you want case insensitive comparisons, pass both `pattern` and `string` through [upper\(\)](#), or both through [lower\(\)](#), before calling `wildcard_match()`.

If you want to detect a pattern anywhere within a string, add * to each end of the pattern:

```
i = wildcard_match('*' & pattern & '*', string)
```

There is currently no way to treat * or ? literally in a pattern.

10.5.1.1.4 Example 1:

```
i = wildcard_match("A?B*", "AQBXXYY")  
-- i is 1 (TRUE)
```

10.5.1.1.5 Example 2:

```
i = wildcard_match("*xyz*", "AAAbbbxyz")  
-- i is 1 (TRUE)
```

10.5.1.1.6 Example 3:

```
i = wildcard_match("A*B*C", "a111b222c")
-- i is 0 (FALSE) because upper/lower case doesn't match
```

10.5.1.1.7 Example 4:

```
bin/search.ex
```

10.5.1.1.8 See Also:

[wildcard_file](#), [upper](#), [lower](#), [Regular expressions](#)

10.5.1.2 wildcard_file

```
include std/wildcard.e
public function wildcard_file(sequence pattern, sequence filename)
```

Determine whether a file name matches a wildcard pattern.

10.5.1.2.1 Parameters:

1. `pattern` : a string, the pattern to match
2. `filename` : the string to be matched against

10.5.1.2.2 Returns:

An **integer**, TRUE if `filename` matches `pattern`, else FALSE.

10.5.1.2.3 Comments:

* matches any 0 or more characters, ? matches any single character. On *Unix* the character comparisons are case sensitive. On Windows they are not.

You might use this function to check the output of the [dir\(\)](#) routine for file names that match a pattern supplied by the user of your program.

10.5.1.2.4 Example 1:

```
i = wildcard_file("AB*CD.?", "aB123cD.e")
-- i is set to 1 on Windows, 0 on Linux or FreeBSD
```


10.5.1.2.5 Example 2:

```
i = wildcard_file("AB*CD.?", "abcd.ex")
-- i is set to 0 on all systems,
-- because the file type has 2 letters not 1
```

10.5.1.2.6 Example 3:

```
bin/search.ex
```

10.5.1.2.7 See Also:

[wildcard_match](#), [dir](#)

11 Data Structures

EUPHORIA Database (EDS)

- Database File Format
- Error Status Constants
- Lock Type Constants
- Error Code Constants
- Variables
- Routines
- Managing databases

Prime Numbers

- Routines

Map (hash table)

- Hashing Algorithms
- Operation codes for put
- Types of Maps
- Types
- Routines

Stack

- Constants
- Types
- Routines

Sets

- Types
- Inclusion and belonging.
- Basic set-theoretic operations.
- Maps between sets.
- Reverse mappings
- Products
- Constants
- Operations on sets

11.1 EUPHORIA Database (EDS)

Database File Format

- Header
- Block of table headers
- Table header
- Index block
- Block of key pointers
- Free list

Error Status Constants

- DB_OK
- DB_OPEN_FAIL
- DB_EXISTS_ALREADY

DB_LOCK_FAIL

DB_FATAL_FAIL

Lock Type Constants

DB_LOCK_NO

DB_LOCK_SHARED

DB_LOCK_EXCLUSIVE

Error Code Constants

MISSING_END

NO_DATABASE

BAD_SEEK

NO_TABLE

DUP_TABLE

BAD_RECNO

INSERT_FAILED

LAST_ERROR_CODE

Variables

db_fatal_id

Routines

db_get_errors

db_dump

check_free_list

Managing databases

db_create

db_open

db_select

db_close

db_select_table

Managing tables

db_current_table

db_create_table

db_delete_table

db_clear_table

db_rename_table

db_table_list

Managing Records

db_find_key

db_get_recid

db_insert

db_delete_record

db_replace_recid

db_replace_data

db_table_size

db_record_data

db_fetch_record

db_record_key

db_record_recid

db_compress

db_current

db_cache_clear

11.1.1 Database File Format

11.1.1.1 Header

- byte 0: magic number for this file-type: 77
- byte 1: version number (major)
- byte 2: version number (minor)
- byte 3: 4-byte pointer to block of table headers
- byte 7: number of free blocks
- byte 11: 4-byte pointer to block of free blocks

11.1.1.2 Block of table headers

- -4: allocated size of this block (for possible reallocation)
- 0: number of table headers currently in use
- 4: table header1
- 16: table header2
- 28: etc.

11.1.1.3 Table header

- 0: pointer to the name of this table
- 4: total number of records in this table
- 8: number of blocks of records
- 12: pointer to the index block for this table

There are two levels of pointers. The logical array of key pointers is split up across many physical blocks. A single index block is used to select the correct small block. This allows inserts and deletes to be made without having to shift a large number of key pointers. Only one small block needs to be adjusted. This is particularly helpful when the table contains many thousands of records.

11.1.1.4 Index block

one per table

- -4: allocated size of index block
- 0: number of records in 1st block of key pointers
- 4: pointer to 1st block
- 8: number of records in 2nd " "
- 12: pointer to 2nd block
- 16: etc.

11.1.1.5 Block of key pointers

many per table

- -4: allocated size of this block in bytes
- 0: key pointer 1
- 4: key pointer 2
- 8: etc.

11.1.1.6 Free list

in ascending order of address

- -4: allocated size of block of free blocks
- 0: address of 1st free block
- 4: size of 1st free block
- 8: address of 2nd free block
- 12: size of 2nd free block
- 16: etc.

The key value and the data value for a record are allocated space as needed. A pointer to the data value is stored just before the key value. EUPHORIA objects, key and data, are stored in a compact form.

All allocated blocks have the size of the block in bytes, stored in the four bytes just before the address.

11.1.2 Error Status Constants

11.1.2.1 DB_OK

```
include std/eds.e
public constant DB_OK
```

Database is OK, not error has occurred.

11.1.2.2 DB_OPEN_FAIL

```
include std/eds.e
public constant DB_OPEN_FAIL
```

The database could not be opened.

11.1.2.3 DB_EXISTS_ALREADY

```
include std/eds.e
public constant DB_EXISTS_ALREADY
```

The database could not be created, it already exists.

11.1.2.4 DB_LOCK_FAIL

```
include std/eds.e
public constant DB_LOCK_FAIL
```

A lock could not be gained on the database.

11.1.2.5 DB_FATAL_FAIL

```
include std/eds.e
public constant DB_FATAL_FAIL
```

A fatal error has occurred.

11.1.3 Lock Type Constants

11.1.3.1 DB_LOCK_NO

```
include std/eds.e
public enum DB_LOCK_NO
```

Do not lock the file.

11.1.3.2 DB_LOCK_SHARED

```
include std/eds.e
public enum DB_LOCK_SHARED
```

Open the database with read-only access.

11.1.3.3 DB_LOCK_EXCLUSIVE

```
include std/eds.e
public enum DB_LOCK_EXCLUSIVE
```

Open the database with read and write access.

11.1.4 Error Code Constants

11.1.4.1 MISSING_END

```
include std/eds.e
public enum MISSING_END
```

Missing 0 terminator

11.1.4.2 NO_DATABASE

```
include std/eds.e
public enum NO_DATABASE
```

current_db is not set

11.1.4.3 BAD_SEEK

```
include std/eds.e
public enum BAD_SEEK
```

seek() failed.

11.1.4.4 NO_TABLE

```
include std/eds.e
public enum NO_TABLE
```

no table was found.

11.1.4.5 DUP_TABLE

```
include std/eds.e
public enum DUP_TABLE
```

this table already exists.

11.1.4.6 BAD_RECNO

```
include std/eds.e
public enum BAD_RECNO
```

unknown key_location index was supplied.

11.1.4.7 INSERT_FAILED

```
include std/eds.e
public enum INSERT_FAILED
```

couldn't insert a new record.

11.1.4.8 LAST_ERROR_CODE

```
include std/eds.e
public enum LAST_ERROR_CODE
```

last error code

11.1.5 Variables

11.1.5.1 db_fatal_id

```
include std/eds.e
public integer db_fatal_id db_fatal_id
```

Exception handler

Set this to a valid routine_id value for a procedure that will be called whenever the library detects a serious error. Your procedure will be passed a single text string that describes the error. It may also call [db_get_errors](#) to get more detail about the cause of the error.

11.1.6 Routines

11.1.6.1 db_get_errors

```
include std/eds.e
public function db_get_errors(integer clearing = 1)
```

Fetches the most recent set of errors recorded by the library.

11.1.6.1.1 Parameters:

1. `clearing` : if zero the set of errors is not reset, otherwise it will be cleared out. The default is to clear the set.

11.1.6.1.2 Returns:

A **sequence**, each element is a set of four fields.

1. Error Code.
2. Error Text.
3. Name of library routine that recorded the error.
4. Parameters passed to that routine.

11.1.6.1.3 Comments:

- A number of library routines can detect errors. If the routine is a function, it usually returns an error code. However, procedures that detect an error can't do that. Instead, they record the error details and you can query that after calling the library routine.
- Both functions and procedures that detect errors record the details in the `Last Error Set`, which is fetched by this function.

11.1.6.1.4 Example 1:

```
db_replace_data(recno, new_data)
errs = db_get_errors()
if length(errs) != 0 then
    display_errors(errs)
    abort(1)
end if
```

11.1.6.2 db_dump

```
include std/eds.e
public procedure db_dump(object file_id, integer low_level_too = 0)
```

print the current database in readable form to file fn

11.1.6.2.1 Parameters:

1. `fn` : the destination file for printing the current EUPHORIA database;
2. `low_level_too` : a boolean. If true, a byte-by-byte binary dump is presented as well; otherwise this step is skipped. If omitted, *false* is assumed.

11.1.6.2.2 Errors:

If the current database is not defined, an error will occur.

11.1.6.2.3 Comments:

- All records in all tables are shown.
- If `low_level_too` is non-zero, then a low-level byte-by-byte dump is also shown. The low-level dump will only be meaningful to someone who is familiar with the internal format of a EUPHORIA database.

11.1.6.2.4 Example 1:

```
if db_open("mydata", DB_LOCK_SHARED) != DB_OK then
    puts(2, "Couldn't open the database!\n")
    abort(1)
end if
fn = open("db.txt", "w")
db_dump(fn) -- Simple output
db_dump("lowlvl_db.txt", 1) -- Full low-level dump created.
```

11.1.6.3 check_free_list

```
include std/eds.e
public procedure check_free_list()
```

Detects corruption of the free list in a EUPHORIA database.

11.1.6.3.1 Comments:

This is a debug routine used by RDS to detect corruption of the free list. Users do not normally call this.

11.1.7 Managing databases

11.1.7.1 db_create

```
include std/eds.e
public function db_create(sequence path, integer lock_method = DB_LOCK_NO, integer init_tables
```

Create a new database, given a file path and a lock method.

11.1.7.1.1 Parameters:

1. `path` : a sequence, the path to the file that will contain the database.
2. `lock_method` : an integer specifying which type of access can be granted to the database. The value of `lock_method` can be either `DB_LOCK_NO` (no lock) or `DB_LOCK_EXCLUSIVE` (exclusive lock).
3. `init_tables` : an integer giving the initial number of tables to reserve space for. The default is 5 and the minimum is 1.
4. `init_free` : an integer giving the initial amount of free space pointers to reserve space for. The default is 5 and the minimum is 0.

11.1.7.1.2 Returns:

An **integer**, status code, either `DB_OK` if creation successful or anything else on an error.

11.1.7.1.3 Comments:

On success, the newly created database becomes the **current database** to which all other database operations will apply.

If the path, `s`, does not end in `.edb`, it will be added automatically.

A version number is stored in the database file so future versions of the database software can recognize the format, and possibly read it and deal with it in some way.

If the database already exists, it will not be overwritten. `db_create()` will return `DB_EXISTS_ALREADY`.

11.1.7.1.4 Example 1:

```
if db_create("mydata", DB_LOCK_NO) != DB_OK then
    puts(2, "Couldn't create the database!\n")
    abort(1)
end if
```

11.1.7.1.5 See Also:

[db_open](#), [db_select](#)

11.1.7.2 db_open

```
include std/eds.e
public function db_open(sequence path, integer lock_method = DB_LOCK_NO)
```

Open an existing EUPHORIA database.

11.1.7.2.1 Parameters:

1. `path` : a sequence, the path to the file containing the database
2. `lock_method` : an integer specifying which sort of access can be granted to the database. The types of lock that you can use are:
 1. `DB_LOCK_NO` : (no lock) - The default
 2. `DB_LOCK_SHARED` : (shared lock for read-only access)
 3. `DB_LOCK_EXCLUSIVE` : (for read/write access).

11.1.7.2.2 Returns:

An **integer**, status code, either `DB_OK` if creation successful or anything else on an error.

11.1.7.2.3 The return codes are:

```
public constant
    DB_OK = 0           -- success
    DB_OPEN_FAIL = -1   -- could not open the file
    DB_LOCK_FAIL = -3   -- could not lock the file in the
                        -- manner requested
```

11.1.7.2.4 Comments:

`DB_LOCK_SHARED` is only supported on Unix platforms. It allows you to read the database, but not write anything to it. If you request `DB_LOCK_SHARED` on *WIN32* it will be treated as if you had asked for `DB_LOCK_EXCLUSIVE`.

If the lock fails, your program should wait a few seconds and try again. Another process might be currently accessing the database.

11.1.7.2.5 Example 1:

```
tries = 0
while 1 do
    err = db_open("mydata", DB_LOCK_SHARED)
    if err = DB_OK then
        exit
    elsif err = DB_LOCK_FAIL then
        tries += 1
        if tries > 10 then
            puts(2, "too many tries, giving up\n")
            abort(1)
        else
            sleep(5)
        end if
    else
        puts(2, "Couldn't open the database!\n")
        abort(1)
    end if
end while
```

11.1.7.2.6 See Also:

[db_create](#), [db_select](#)

11.1.7.3 db_select

```
include std/eds.e
public function db_select(sequence path, integer lock_method = - 1)
```

Choose a new, already open, database to be the current database.

11.1.7.3.1 Parameters:

1. `path`: a sequence, the path to the database to be the new current database.
2. `lock_method`: an integer. Optional locking method.

11.1.7.3.2 Returns:

An **integer**, `DB_OK` on success or an error code.

11.1.7.3.3 Comments:

- Subsequent database operations will apply to this database. path is the path of the database file as it was originally opened with `db_open()` or `db_create()`.
- When you create (`db_create`) or open (`db_open`) a database, it automatically becomes the current database. Use `db_select()` when you want to switch back and forth between open databases, perhaps to copy records from one to the other. After selecting a new database, you should select a table within that database using `db_select_table()`.
- If the `lock_method` is omitted and the database has not already been opened, this function will fail. However, if `lock_method` is a valid lock type for `db_open` and the database is not open yet, this function will attempt to open it. It may still fail if the database cannot be opened.

11.1.7.3.4 Example 1:

```
if db_select("employees") != DB_OK then
    puts(2, "Could not select employees database\n")
end if
```

11.1.7.3.5 Example 2:

```
if db_select("customer", DB_LOCK_SHARED) != DB_OK then
    puts(2, "Could not open or select Customer database\n")
end if
```

11.1.7.3.6 See Also:

[db_open](#), [db_select](#)

11.1.7.4 db_close

```
include std/eds.e
public procedure db_close()
```

Unlock and close the current database.

11.1.7.4.1 Comments:

Call this procedure when you are finished with the current database. Any lock will be removed, allowing other processes to access the database file. The current database becomes undefined.

11.1.7.5 db_select_table

```
include std/eds.e
public function db_select_table(sequence name)
```

11.1.7.6 Managing tables

11.1.7.6.1 Parameters:

1. name : a sequence which defines the name of the new current table.

On success, the table with name given by name becomes the current table.

11.1.7.6.2 Returns:

An **integer**, either DB_OK on success or DB_OPEN_FAIL otherwise.

11.1.7.6.3 Errors:

An error occurs if the current database is not defined.

11.1.7.6.4 Comments:

All record-level database operations apply automatically to the current table.

11.1.7.6.5 Example 1:

```
if db_select_table("salary") != DB_OK then
    puts(2, "Couldn't find salary table!\n")
    abort(1)
end if
```

11.1.7.6.6 See Also:

[db_table_list](#)

11.1.7.7 db_current_table

```
include std/eds.e
public function db_current_table()
```

Get name of currently selected table

11.1.7.7.1 Parameters:

1. None.

11.1.7.7.2 Returns:

A **sequence**, the name of the current table. An empty string means that no table is currently selected.

11.1.7.7.3 Example 1:

```
s = db_current_table()
```

11.1.7.7.4 See Also:

[db_select_table](#), [db_table_list](#)

11.1.7.8 db_create_table

```
include std/eds.e
public function db_create_table(sequence name, integer init_records = INIT_RECORDS)
```

Create a new table within the current database.

11.1.7.8.1 Parameters:

1. `name` : a sequence, the name of the new table.
2. `init_records` : The number of records to initially reserve space for. (Default is 50)

11.1.7.8.2 Returns:

An **integer**, either `DB_OK` on success or `DB_EXISTS_ALREADY` on failure.

11.1.7.8.3 Errors:

An error occurs if the current database is not defined.

11.1.7.8.4 Comments:

- The supplied name must not exist already on the current database.
- The table that you create will initially have 0 records. However it will reserve some space for a number of records, which will improve the initial data load for the table.
- It becomes the current table.

11.1.7.8.5 Example 1:

```
if db_create_table("my_new_table") != DB_OK then
    puts(2, "Could not create my_new_table!\n")
end if
```

11.1.7.8.6 See Also:

[db_select_table](#), [db_table_list](#)

11.1.7.9 db_delete_table

```
include std/eds.e
public procedure db_delete_table(sequence name)
```

Delete a table in the current database.

11.1.7.9.1 Parameters:

1. name : a sequence, the name of the table to delete.

11.1.7.9.2 Errors:

An error occurs if the current database is not defined.

11.1.7.9.3 Comments:

If there is no table with the name given by name, then nothing happens. On success, all records are deleted and all space used by the table is freed up. If the table was the current table, the current table becomes undefined.

11.1.7.9.4 See Also:

[db_table_list](#), [db_select_table](#), [db_clear_table](#)

11.1.7.10 db_clear_table

```
include std/eds.e
public procedure db_clear_table(sequence name, integer init_records = INIT_RECORDS)
```

Clears a table of all its records, in the current database.

11.1.7.10.1 Parameters:

1. `name` : a sequence, the name of the table to clear.

11.1.7.10.2 Errors:

An error occurs if the current database is not defined.

11.1.7.10.3 Comments:

If there is no table with the name given by `name`, then nothing happens. On success, all records are deleted and all space used by the table is freed up. If this is the current table, after this operation it will still be the current table.

11.1.7.10.4 See Also:

[db_table_list](#), [db_select_table](#), [db_delete_table](#)

11.1.7.11 db_rename_table

```
include std/eds.e
public procedure db_rename_table(sequence name, sequence new_name)
```

Rename a table in the current database.

11.1.7.11.1 Parameters:

1. `name` : a sequence, the name of the table to rename
2. `new_name` : a sequence, the new name for the table

11.1.7.11.2 Errors:

- An error occurs if the current database is not defined.
- If `name` does not exist on the current database, or if `new_name` does exist on the current database, an error will occur.

11.1.7.11.3 Comments:

The table to be renamed can be the current table, or some other table in the current database.

11.1.7.11.4 See Also:

[db_table_list](#)

11.1.7.12 db_table_list

```
include std/eds.e
public function db_table_list()
```

Lists all tables on the current database.

11.1.7.12.1 Returns:

A **sequence**, of all the table names in the current database. Each element of this sequence is a sequence, the name of a table.

11.1.7.12.2 Errors:

An error occurs if the current database is undefined.

11.1.7.12.3 Example 1:

```
sequence names = db_table_list()
for i = 1 to length(names) do
    puts(1, names[i] & '\n')
end for
```

11.1.7.12.4 See Also:

[db_select_table](#), [db_create_table](#)

11.1.7.13 Managing Records

11.1.7.14 db_find_key

```
include std/eds.e
public function db_find_key(object key, object table_name = current_table_name)
```

Find the record in the current table with supplied key.

11.1.7.14.1 Parameters:

1. `key` : the identifier of the record to be looked up.
2. `table_name` : optional name of table to find key in

11.1.7.14.2 Returns:

An **integer**, either greater or less than zero:

- If above zero, the record identified by `key` was found on the current table, and the returned integer is its record number.
- If less than zero, the record was not found. The returned integer is the opposite of what the record number would have been, had the record been found.
- If equal to zero, an error occurred.

11.1.7.14.3 Errors:

If the current table is not defined, it returns 0.

11.1.7.14.4 Comments:

A fast binary search is used to find the key in the current table. The number of comparisons is proportional to the log of the number of records in the table. The key is unique--a table is more like a dictionary than like a spreadsheet.

You can select a range of records by searching for the first and last key values in the range. If those key values don't exist, you'll at least get a negative value showing where they would be, if they existed. e.g. Suppose you want to know which records have keys greater than "GGG" and less than "MMM". If -5 is returned for key "GGG", it means a record with "GGG" as a key would be inserted as record number 5. -27 for "MMM" means a record with "MMM" as its key would be inserted as record number 27. This quickly tells you that all records, ≥ 5 and < 27 qualify.

11.1.7.14.5 Example 1:

```
rec_num = db_find_key("Millennium")
if rec_num > 0 then
    ? db_record_key(rec_num)
    ? db_record_data(rec_num)
else
    puts(2, "Not found, but if you insert it,\n")

    printf(2, "it will be %#d\n", -rec_num)
end if
```

11.1.7.14.6 See Also:

[db_insert](#), [db_replace_data](#), [db_delete_record](#), [db_get_recid](#)

11.1.7.15 db_get_recid

```
include std/eds.e
public function db_get_recid(object key, object table_name = current_table_name)
```

Returns the unique record identifier (*recid*) value for the record.

11.1.7.15.1 Parameters:

1. *key* : the identifier of the record to be looked up.
2. *table_name* : optional name of table to find key in

11.1.7.15.2 Returns:

An **atom**, either greater or equal to zero:

- If above zero, it is a *recid*.
- If less than zero, the record wasn't found.
- If equal to zero, an error occurred.

11.1.7.15.3 Errors:

If the table is not defined, an error is raised.

11.1.7.15.4 Comments:

A **recid** is a number that uniquely identifies a record in the database. No two records in a database has the same *recid* value. They can be used instead of keys to *quickly* refetch a record, as they avoid the overhead of looking for a matching record key. They can also be used without selecting a table first, as the *recid* is unique to the database and not just a table. However, they only remain valid while a database is open and so long as it doesn't get compressed. Compressing the database will give each record a new *recid* value.

Because it is faster to fetch a record with a *recid* rather than with its key, these are used when you know you have to **refetch** a record.

11.1.7.15.5 Example 1:

```
rec_num = db_get_recid("Millennium")
if rec_num > 0 then
    ? db_record_recid(rec_num) -- fetch key and data.
```

```
else
    puts(2, "Not found\n")
end if
```

11.1.7.15.6 See Also:

[db_insert](#), [db_replace_data](#), [db_delete_record](#), [db_find_key](#)

11.1.7.16 db_insert

```
include std/eds.e
public function db_insert(object key, object data, object table_name = current_table_name)
```

Insert a new record into the current table.

11.1.7.16.1 Parameters:

1. `key` : an object, the record key, which uniquely identifies it inside the current table
2. `data` : an object, associated to key.
3. `table_name` : optional table name to insert record into

11.1.7.16.2 Returns:

An **integer**, either DB_OK on success or an error code on failure.

11.1.7.16.3 Comments:

Within a table, all keys must be unique. `db_insert()` will fail with DB_EXISTS_ALREADY if a record already exists on current table with the same key value.

Both key and data can be any EUPHORIA data objects, atoms or sequences.

11.1.7.16.4 Example 1:

```
if db_insert("Smith", {"Peter", 100, 34.5}) != DB_OK then
    puts(2, "insert failed!\n")
end if
```

11.1.7.16.5 See Also:

[db_delete_record](#)

11.1.7.17 db_delete_record

```
include std/eds.e
public procedure db_delete_record(integer key_location, object table_name = current_table_name)
```

Delete record number `key_location` from the current table.

11.1.7.17.1 Parameter:

1. `key_location` : a positive integer, designating the record to delete.
2. `table_name` : optional table name to delete record from.

11.1.7.17.2 Errors:

If the current table is not defined, or `key_location` is not a valid record index, an error will occur. Valid record indexes are between 1 and the number of records in the table.

11.1.7.17.3 Example 1:

```
db_delete_record(55)
```

11.1.7.17.4 See Also:

[db_find_key](#)

11.1.7.18 db_replace_recid

```
include std/eds.e
public procedure db_replace_recid(integer recid, object data)
```

In the current database, replace the data portion of a record with new data. This can be used to quickly update records that have already been located by calling [db_get_recid](#). This operation is faster than using [db_replace_data](#)

11.1.7.18.1 Parameters:

1. `recid` : an atom, the `recid` of the record to be updated.
2. `data` : an object, the new value of the record.

11.1.7.18.2 Comments:

- `recid` must be fetched using `db_get_recid` first.
- `data` is an EUPHORIA object of any kind, atom or sequence.
- The `recid` does not have to be from the current table.
- This does no error checking. It assumes the database is open and valid.

11.1.7.18.3 Example 1:

```
rid = db_get_recid("Peter")
rec = db_record_recid(rid)
rec[2][3] *= 1.10
db_replace_recid(rid, rec[2])
```

11.1.7.18.4 See Also:

[db_replace_data](#), [db_find_key](#), [db_get_recid](#)

11.1.7.19 db_replace_data

```
include std/eds.e
public procedure db_replace_data(integer key_location, object data, object table_name = current)
```

In the current table, replace the data portion of a record with new data.

11.1.7.19.1 Parameters:

1. `key_location`: an integer, the index of the record the data is to be altered.
2. `data`: an object, the new value associated to the key of the record.
3. `table_name`: optional table name of record to replace data in.

11.1.7.19.2 Comments:

`key_location` must be from 1 to the number of records in the current table. `data` is an EUPHORIA object of any kind, atom or sequence.

11.1.7.19.3 Example 1:

```
db_replace_data(67, {"Peter", 150, 34.5})
```


11.1.7.19.4 See Also:

[db_find_key](#)

11.1.7.20 db_table_size

```
include std/eds.e
public function db_table_size(object table_name = current_table_name)
```

Get the size (number of records) of the default table.

11.1.7.20.1 Parameters:

1. `table_name` : optional table name to get the size of.

Returns An **integer**, the current number of records in the current table. If a value less than zero is returned, it means that an error occurred.

11.1.7.20.2 Errors:

If the current table is undefined, an error will occur.

11.1.7.20.3 Example 1:

```
-- look at all records in the current table
for i = 1 to db_table_size() do
    if db_record_key(i) = 0 then
        puts      (1, "0 key found\n")
        exit
    end if
end for
```

11.1.7.20.4 See Also:

[db_replace_data](#)

11.1.7.21 db_record_data

```
include std/eds.e
public function db_record_data(integer key_location, object table_name = current_table_name)
```

Returns the data in a record queried by position.

11.1.7.21.1 Parameters:

1. `key_location` : the index of the record the data of which is being fetched.
2. `table_name` : optional table name to get record data from.

11.1.7.21.2 Returns:

An **object**, the data portion of requested record.

NOTE This function calls `fatal()` and returns a value of -1 if an error prevented the correct data being returned.

11.1.7.21.3 Comments:

Each record in a EUPHORIA database consists of a key portion and a data portion. Each of these can be any EUPHORIA atom or sequence.

11.1.7.21.4 Errors:

If the current table is not defined, or if the record index is invalid, an error will occur.

11.1.7.21.5 Example 1:

```
puts(1, "The 6th record has data value: ")
? db_record_data(6)
```

11.1.7.21.6 See Also:

[db_find_key](#), [db_replace_data](#)

11.1.7.22 db_fetch_record

```
include std/eds.e
public function db_fetch_record(object key, object table_name = current_table_name)
```

Returns the data for the record with supplied key.

11.1.7.22.1 Parameters:

1. `key` : the identifier of the record to be looked up.
2. `table_name` : optional name of table to find key in

11.1.7.22.2 Returns:

An **integer**,

- If less than zero, the record was not found. The returned integer is the opposite of what the record number would have been, had the record been found.
- If equal to zero, an error occurred. A sequence, the data for the record.

11.1.7.22.3 Errors:

If the current table is not defined, it returns 0.

11.1.7.22.4 Comments:

Each record in a EUPHORIA database consists of a key portion and a data portion. Each of these can be any EUPHORIA atom or sequence. **NOTE** This function does not support records that data consists of a single non-sequence value. In those cases you will need to use [db_find_key](#) and [db_record_data](#).

11.1.7.22.5 Example 1:

```
printf(1, "The record['%s'] has data value:\n", {"foo"})
? db_fetch_record("foo")
```

11.1.7.22.6 See Also:

[db_find_key](#), [db_record_data](#)

11.1.7.23 db_record_key

```
include std/eds.e
public function db_record_key(integer key_location, object table_name = current_table_name)
```

11.1.7.23.1 Parameters:

1. `key_location` : an integer, the index of the record the key is being requested.
2. `table_name` : optional table name to get record key from.

Returns An **object**, the key of the record being queried by index.

NOTE This function calls `fatal()` and returns a value of -1 if an error prevented the correct data being returned.

11.1.7.23.2 Errors:

If the current table is not defined, or if the record index is invalid, an error will occur.

11.1.7.23.3 Comments:

Each record in a EUPHORIA database consists of a key portion and a data portion. Each of these can be any EUPHORIA atom or sequence.

11.1.7.23.4 Example 1:

```
puts(1, "The 6th record has key value: ")  
? db_record_key(6)
```

11.1.7.23.5 See Also:

[db_record_data](#)

11.1.7.24 db_record_recid

```
include std/eds.e  
public function db_record_recid(integer recid)
```

Returns the key and data in a record queried by `recid`.

11.1.7.24.1 Parameters:

1. `recid`: the `recid` of the required record, which has been previously fetched using [db_get_recid](#).

11.1.7.24.2 Returns:

An **sequence**, the first element is the key and the second element is the data portion of requested record.

11.1.7.24.3 Comments:

- This is much faster than calling [db_record_key](#) and [db_record_data](#).
- This does no error checking. It assumes the database is open and valid.
- This function does not need the requested record to be from the current table. The `recid` can refer to a record in any table.

11.1.7.24.4 Example 1:

```
rid = db_get_recid("SomeKey")
? db_record_recid(rid)
```

11.1.7.24.5 See Also:

[db_get_recid](#), [db_replace_recid](#)

11.1.7.25 db_compress

```
include std/eds.e
public function db_compress()
```

Compresses the current database.

11.1.7.25.1 Returns:

An **integer**, either DB_OK on success or an error code on failure.

11.1.7.25.2 Comments:

The current database is copied to a new file such that any blocks of unused space are eliminated. If successful, the return value will be set to DB_OK, and the new compressed database file will retain the same name. The current table will be undefined. As a backup, the original, uncompressed file will be renamed with an extension of .t0 (or .t1, .t2, ..., .t99). In the highly unusual case that the compression is unsuccessful, the database will be left unchanged, and no backup will be made.

When you delete items from a database, you create blocks of free space within the database file. The system keeps track of these blocks and tries to use them for storing new data that you insert. db_compress() will copy the current database without copying these free areas. The size of the database file may therefore be reduced. If the backup filenames reach .t99 you will have to delete some of them.

11.1.7.25.3 Example 1:

```
if db_compress() != DB_OK then
    puts(2, "compress failed!\n")
end if
```

11.1.7.26 db_current

```
include std/eds.e
public function db_current()
```

Get name of currently selected database.

11.1.7.26.1 Parameters:

1. None.

11.1.7.26.2 Returns:

A **sequence**, the name of the current database. An empty string means that no database is currently selected.

11.1.7.26.3 Comments:

The actual name returned is the *path* as supplied to the `db_open` routine.

11.1.7.26.4 Example 1:

```
s = db_current_database()
```

11.1.7.26.5 See Also:

[db_select](#)

11.1.7.27 db_cache_clear

```
include std/eds.e
public procedure db_cache_clear()
```

Forces the database index cache to be cleared.

11.1.7.27.1 Parameters:

None

11.1.7.27.2 Comments:

- This is not normally required to the run. You might run it to set up a predetermined state for performance timing, or to release some memory back to the application.

11.1.7.27.3 Example 1:

```
db_cache_clear() -- Clear the cache.
```

11.1.7.28 db_set_caching

```
include std/eds.e
public function db_set_caching(atom new_setting)
```

Sets the key cache behavior.

Initially, the cache option is turned on. This means that when possible, the keys of a table are kept in RAM rather than read from disk each time `db_select_table()` is called. For most databases, this will improve performance when you have more than one table in it.

11.1.7.28.1 Parameters:

1. `integer` : 0 will turn off caching, 1 will turn it back on.

11.1.7.28.2 Returns:

An **integer**, the previous setting of the option.

11.1.7.28.3 Comments:

When caching is turned off, the current cache contents is totally cleared.

11.1.7.28.4 Example 1:

```
x = db_set_caching(0) -- Turn off key caching.
```

11.2 Prime Numbers

Page Contents

[Routines](#)

[calc_primes](#)
[next_prime](#)
[prime_list](#)

11.2.1 Routines

11.2.1.1 calc_primes

```
include std/primes.e
public function calc_primes(integer max_p, atom time_limit_p = 10)
```

Returns all the prime numbers below some threshold, with a cap on computation time.

11.2.1.1.1 Parameters:

1. `max_p` : an integer, the last prime returned is the next prime after or on this value.
2. `time_out_p` : an atom, the maximum number of seconds that this function can run for. The default is 10 (ten) seconds.

11.2.1.1.2 Returns:

A **sequence**, made of prime numbers in increasing order. The last value is the next prime number that falls on or after the value of `max_p`.

11.2.1.1.3 Comments:

- The returned sequence contains all the prime numbers less than its last element.
- If the function times out, it may not hold all primes below `max_p`, but only the largest ones will be absent. If the last element returned is less than `max_p` then the function timed out.
- To disable the timeout, simply give it a negative value.

11.2.1.1.4 Example 1:

```
? calc_primes(1000, 5)
-- On a very slow computer, you may only get all primes up to say 719.
-- On a faster computer, the last element printed out will be 997.
-- This call will never take longer than 5 seconds.
```

11.2.1.1.5 See Also:

[next_prime](#) [prime_list](#)

11.2.1.2 next_prime

```
include std/primes.e
public function next_prime(integer n, object fail_signal_p = - 1, atom time_out_p = 1)
```

Return the next prime number on or after the supplied number

11.2.1.2.1 Parameters:

1. `n`: an integer, the starting point for the search
2. `fail_signal_p`: an integer, used to signal error. Defaults to -1.

11.2.1.2.2 Returns:

An **integer**, which is prime only if it took less than 1 second to determine the next prime greater or equal to `n`.

11.2.1.2.3 Comments:

The default value of -1 will alert you about an invalid returned value, since a prime not less than `n` is expected. However, you can pass another value for this parameter.

11.2.1.2.4 Example 1:

```
? next_prime(997)
-- On a very slow computer, you might get -997, but 1003 is expected.
```

11.2.1.2.5 See Also:

[calc_primes](#)

11.2.1.3 prime_list

```
include std/primes.e
public function prime_list(integer top_prime_p = 0)
```

Returns a list of prime numbers.

11.2.1.3.1 Parameters:

1. `top_prime_p`: The list will end with the prime less than or equal to this value. If this is zero, the current list calculated primes is returned.

11.2.1.3.2 Returns:

An **sequence**, a list of prime numbers from 2 to `top_prime_p`

11.2.1.3.3 Example 1:

```
sequence pList = prime_list(1000)
-- pList will now contain all the primes from 2 up to the largest less than or
-- equal to 1000.
```

11.2.1.3.4 See Also:

[calc_primes](#), [next_prime](#)

11.3 Map (hash table)

- hash
- Hashing Algorithms
 - HSIEH32
 - FLETCHER32
 - ADLER32
 - MD5
 - SHA256
- Operation codes for put
 - PUT
 - ADD
 - SUBTRACT
 - MULTIPLY
 - DIVIDE
 - APPEND
 - CONCAT
 - LEAVE
- Types of Maps
 - SMALLMAP
- Types
 - map
- Routines
 - calc_hash
 - threshold
 - type_of
 - rehash
 - new
 - new_extra
 - compare
 - has
 - get

```

nested_get
put
nested_put
remove
clear
size
NUM_ENTRIES
statistics
keys
values
pairs
optimize
load_map
SM_TEXT
save_map
copy
new_from_kvpairs
new_from_string
for_each

```

A map is a special array, often called an associative array or dictionary, in which the index to the data can be any EUPHORIA object and not just an integer. These sort of indexes are also called keys. For example we can code things like this...

```

custrec = new() -- Create a new map
put(custrec, "Name", "Joe Blow")
put(custrec, "Address", "555 High Street")
put(custrec, "Phone", 555675632)

```

This creates three elements in the map, and they are indexed by "Name", "Address" and "Phone", meaning that to get the data associated with those keys we can code ...

```

object data = get(custrec, "Phone")
-- data now set to 555675632

```

Note: Only one instance of a given key can exist in a given map, meaning for example, we couldn't have two separate "Name" values in the above *custrec* map.

Maps automatically grow to accommodate all the elements placed into it.

Associative arrays can be implemented in many different ways, depending on what efficiency trade-offs have been made. This implementation allows you to decide if you want a *small* map or a *large* map.

small map

Faster for small numbers of elements. Speed is usually proportional to the number of elements.

large map

Faster for large number of elements. Speed is usually the same regardless of how many elements are in the map. The speed is often slower than a small map.

Note: If the number of elements placed into a *small* map take it over the initial size of the map, it is automatically converted to a *large* map.

11.3.1 hash

```
<built-in> function hash(object source, atom algo)
```

Calculates a hash value from *key* using the algorithm *algo*

11.3.1.1 Parameters:

1. *source* : Any EUPHORIA object
2. *algo* : A code indicating which algorithm to use.
 - ◆ -5 uses Hsieh. Fastest and good dispersion
 - ◆ -4 uses Fletcher. Very fast and good dispersion
 - ◆ -3 uses Adler. Very fast and reasonable dispersion, especially for small strings
 - ◆ -2 uses MD5 (not implemented yet) Slower but very good dispersion. Suitable for signatures.
 - ◆ -1 uses SHA256 (not implemented yet) Slow but excellent dispersion. Suitable for signatures. More secure than MD5.
 - ◆ 0 and above (integers and decimals) and non-integers less than zero use the cyclic variant ($\text{hash} = \text{hash} * \text{algo} + c$). This is a fast and good to excellent dispersion depending on the value of *algo*. Decimals give better dispersion but are slightly slower.

11.3.1.1.1 Returns:

An **integer**, Except for the MD5 and SHA256 algorithms, this is a 32-bit integer.
A **sequence**, MD5 returns a 4-element sequence of integers
SHA256 returns a 8-element sequence of integers.

11.3.1.1.2 Comments:

- For *algo* values from zero to less than 1, that actual value used is ($\text{algo} + 69096$).

11.3.1.1.3 Example 1:

```
x = hash("The quick brown fox jumps over the lazy dog", 0)
-- x is 242399616
x = hash("The quick brown fox jumps over the lazy dog", 99.94)
-- x is 723158
x = hash("The quick brown fox jumps over the lazy dog", -99.94)
-- x is 4175585990
x = hash("The quick brown fox jumps over the lazy dog", -4)
-- x is 467406810
```

11.3.2 Hashing Algorithms

11.3.2.1 HSIEH32

```
include std/map.e
public enum HSIEH32
```

11.3.2.2 FLETCHER32

```
include std/map.e
public enum FLETCHER32
```

11.3.2.3 ADLER32

```
include std/map.e
public enum ADLER32
```

11.3.2.4 MD5

```
include std/map.e
public enum MD5
```

11.3.2.5 SHA256

```
include std/map.e
public enum SHA256
```

11.3.3 Operation codes for put

11.3.3.1 PUT

```
include std/map.e
public enum PUT
```

11.3.3.2 ADD

```
include std/map.e
public enum ADD
```

11.3.3.3 SUBTRACT

```
include std/map.e
public enum SUBTRACT
```

11.3.3.4 MULTIPLY

```
include std/map.e
public enum MULTIPLY
```

11.3.3.5 DIVIDE

```
include std/map.e
public enum DIVIDE
```

11.3.3.6 APPEND

```
include std/map.e
public enum APPEND
```

11.3.3.7 CONCAT

```
include std/map.e
public enum CONCAT
```

11.3.3.8 LEAVE

```
include std/map.e
public enum LEAVE
```

11.3.4 Types of Maps

11.3.4.1 SMALLMAP

```
include std/map.e
public constant SMALLMAP
```

11.3.5 Types

11.3.5.1 map

```
include std/map.e
public type map(object obj_p)
```

Defines the datatype 'map'

11.3.5.1.1 Comments:

Used when declaring a map variable.

11.3.5.1.2 Example:

```
map SymbolTable = new() -- Create a new map to hold the symbol table.
```

11.3.6 Routines

11.3.6.1 calc_hash

```
include std/map.e
public function calc_hash(object key_p, integer max_hash_p = 0)
```

Calculate a Hashing value from the supplied data.

11.3.6.1.1 Parameters:

1. pData : The data for which you want a hash value calculated.
2. max_hash_p : (default = 0) The returned value will be no larger than this value. However, a value of 0 or lower means that it can grow as large as the maximum integer value.

11.3.6.1.2 Returns:

An **integer**, the value of which depends only on the supplied data.

11.3.6.1.3 Comments:

This is used whenever you need a single number to represent the data you supply. It can calculate the number based on all the data you give it, which can be an atom or sequence of any value.

11.3.6.1.4 Example 1:

```
integer h1
h1 = calc_hash( symbol_name )
```

11.3.6.2 threshold

```
include std/map.e
public function threshold(integer new_value_p = 0)
```

Gets or Sets the threshold value that determines at what point a small map converts into a large map structure. Initially this has been set to 50, meaning that maps up to 50 elements use the *small map* structure.

11.3.6.2.1 Parameters:

1. `new_value_p` : If this is greater than zero then it **sets** the threshold value.

11.3.6.2.2 Returns:

An **integer**, the current value (when `new_value_p` is less than 1) or the old value prior to setting it to `new_value_p`.

11.3.6.3 type_of

```
include std/map.e
public function type_of(map the_map_p)
```

Determines the type of the map.

11.3.6.3.1 Parameters:

1. `m` : A map

11.3.6.3.2 Returns:

An **integer**, Either *SMALLMAP* or *LARGEMAP*

11.3.6.4 rehash

```
include std/map.e
public procedure rehash(map the_map_p, integer requested_bucket_size_p = 0)
```


Changes the width, i.e. the number of buckets, of a map. Only effects *large* maps.

11.3.6.4.1 Parameters:

1. `m` : the map to resize
2. `requested_bucket_size_p` : a lower limit for the new size.

11.3.6.4.2 Returns:

A **map**, with the same data in, but more evenly dispatched and hence faster to use.

11.3.6.4.3 Comment:

If `requested_bucket_size_p` is not greater than zero, a new width is automatically derived from the current one.

11.3.6.4.4 See Also:

[statistics](#), [optimize](#)

11.3.6.5 new

```
include std/map.e
public function new(integer initial_size_p = 690)
```

Create a new map data structure

11.3.6.5.1 Parameters:

1. `initial_size_p` : An estimate of how many initial elements will be stored in the map. If this value is less than the [threshold](#) value, the map will initially be a *small* map otherwise it will be a *large* map.

11.3.6.5.2 Returns:

An empty **map**.

11.3.6.5.3 Comments:

A new object of type map is created. The resources allocated for the map will be automatically cleaned up if the reference count of the returned value drops to zero, or if passed in a call to [delete](#).

11.3.6.5.4 Example 1:

```
map m = new()  -- m is now an empty map
map x = new(threshold()) -- Forces a small map to be initialized
x = new()      -- the resources for the map previously stored in x are released automatically
delete( m )    -- the resources for the map are released
```

11.3.6.6 new_extra

```
include std/map.e
public function new_extra(object the_map_p, integer initial_size_p = 690)
```

Returns either the supplied map or a new map.

11.3.6.6.1 Parameters:

1. `the_map_p`: An object, that could be an existing map
2. `initial_size_p`: An estimate of how many initial elements will be stored in a new map.

11.3.6.6.2 Returns:

A **map**, If `m` is an existing map then it is returned otherwise this returns a new empty **map**.

11.3.6.6.3 Comments:

This is used to return a new map if the supplied variable isn't already a map.

11.3.6.6.4 Example 1:

```
map m = new_extra( foo() ) -- If foo() returns a map it is used, otherwise
                           -- a new map is created.
```

11.3.6.7 compare

```
include std/map.e
public function compare(map map_1_p, map map_2_p, integer scope_p = 'd')
```

Compares two maps to test equality.

11.3.6.7.1 Parameters:

1. `map_1_p` : A map
2. `map_2_p` : A map
3. `scope_p` : An integer that specifies what to compare.
 - ◆ 'k' or 'K' to only compare keys.
 - ◆ 'v' or 'V' to only compare values.
 - ◆ 'd' or 'D' to compare both keys and values. This is the default.

11.3.6.7.2 Returns:

An **integer**,

- -1 if they are not equal.
- 0 if they are literally the same map.
- 1 if they contain the same keys and/or values.

11.3.6.7.3 Example 1:

```
map map_1_p = foo()
map map_2_p = bar()
if compare(map_1_p, map_2_p, 'k') >= 0 then
    ... -- two maps have the same keys
```

11.3.6.8 has

```
include std/map.e
public function has(map the_map_p, object the_key_p)
```

Check whether map has a given key.

11.3.6.8.1 Parameters:

1. `the_map_p` : the map to inspect
2. `the_key_p` : an object to be looked up

11.3.6.8.2 Returns:

An **integer**, 0 if not present, 1 if present.

11.3.6.8.3 Example 1:

```
map the_map_p
the_map_p = new()
```

```
put(the_map_p, "name", "John")
? has(the_map_p, "name") -- 1
? has(the_map_p, "age")  -- 0
```

11.3.6.8.4 See Also:

[get](#)

11.3.6.9 get

```
include std/map.e
public function get(map the_map_p, object the_key_p, object default_value_p = 0)
```

Retrieves the value associated to a key in a map.

11.3.6.9.1 Parameters:

1. `the_map_p` : the map to inspect
2. `the_key_p` : an object, the `the_key_p` being looked tp
3. `default_value_p` : an object, a default value returned if `the_key_p` not found. The default is 0.

11.3.6.9.2 Returns:

An **object**, the value that corresponds to `the_key_p` in `the_map_p`. If `the_key_p` is not in `the_map_p`, `default_value_p` is returned instead.

11.3.6.9.3 Example 1:

```
map ages
ages = new()
put(ages, "Andy", 12)
put(ages, "Budi", 13)

integer age
age = get(ages, "Budi", -1)
if age = -1 then
    puts(1, "Age unknown")
else
    printf(1, "The age is %d", age)
end if
```

11.3.6.9.4 See Also:

[has](#)

11.3.6.10 nested_get

```
include std/map.e
public function nested_get(map the_map_p, sequence the_keys_p, object default_value_p = 0)
```

Returns the value that corresponds to the object `the_keys_p` in the nested map `the_map_p`. `the_keys_p` is a sequence of keys. If any key is not in the map, the object `default_value_p` is returned instead.

11.3.6.11 put

```
include std/map.e
public procedure put(map the_map_p, object the_key_p, object the_value_p, integer operation_p =
```

Adds or updates an entry on a map.

11.3.6.11.1 Parameters:

1. `the_map_p` : the map where an entry is being added or updated
2. `the_key_p` : an object, the `the_key_p` to look up
3. `the_value_p` : an object, the value to add, or to use for updating.
4. `operation` : an integer, indicating what is to be done with `the_value_p`. Defaults to PUT.
5. `trigger_p` : an integer. Default is 100. See Comments for details.

11.3.6.11.2 Returns:

The updated **map**.

11.3.6.11.3 Comments:

- The operation parameter can be used to modify the existing value. Valid operations are:

- ◆ PUT -- This is the default, and it replaces any value in there already
- ◆ ADD -- Equivalent to using the += operator
- ◆ SUBTRACT -- Equivalent to using the -= operator
- ◆ MULTIPLY -- Equivalent to using the *= operator
- ◆ DIVIDE -- Equivalent to using the /= operator
- ◆ APPEND -- Appends the value to the existing data
- ◆ CONCAT -- Equivalent to using the &= operator

- ◆ `LEAVE` -- If it already exists, the current value is left unchanged otherwise the new value is added to the map.
- The *trigger* parameter is used when you need to keep the average number of keys in a hash bucket to a specific maximum. The *trigger* value is the maximum allowed. Each time a *put* operation increases the hash table's average bucket size to be more than the *trigger* value the table is expanded by a factor of 3.5 and the keys are rehashed into the enlarged table. This can be a time intensive action so set the value to one that is appropriate to your application.
 - ◆ By keeping the average bucket size to a certain maximum, it can speed up lookup times.
 - ◆ If you set the *trigger* to zero, it will not check to see if the table needs reorganizing. You might do this if you created the original bucket size to an optimal value. See [new](#) on how to do this.

11.3.6.11.4 Example 1:

```
map ages
ages = new()
put (ages, "Andy", 12)
put (ages, "Budi", 13)
put (ages, "Budi", 14)

-- ages now contains 2 entries: "Andy" => 12, "Budi" => 14
```

11.3.6.11.5 See Also:

[remove](#), [has](#), [nested_put](#)

11.3.6.12 nested_put

```
include std/map.e
public procedure nested_put(map the_map_p, sequence the_keys_p, object the_value_p, integer operation_p, integer trigger_p)
```

Adds or updates an entry on a map.

11.3.6.12.1 Parameters:

1. *the_map_p* : the map where an entry is being added or updated
2. *the_keys_p* : a sequence of keys for the nested maps
3. *the_value_p* : an object, the value to add, or to use for updating.
4. *operation_p* : an integer, indicating what is to be done with value. Defaults to PUT.
5. *trigger_p* : an integer. Default is 51. See Comments for details.

11.3.6.12.2 Valid operations are:

- PUT -- This is the default, and it replaces any value in there already
- ADD -- Equivalent to using the += operator
- SUBTRACT -- Equivalent to using the -= operator
- MULTIPLY -- Equivalent to using the *= operator
- DIVIDE -- Equivalent to using the /= operator
- APPEND -- Appends the value to the existing data
- CONCAT -- Equivalent to using the &= operator

11.3.6.12.3 Returns:

The modified **map**.

11.3.6.12.4 Comments:

- If existing entry with the same key is already in the map, the value of the entry is updated.
- The *trigger* parameter is used when you need to keep the average number of keys in a hash bucket to a specific maximum. The *trigger* value is the maximum allowed. Each time a *put* operation increases the hash table's average bucket size to be more than the *trigger* value the table is expanded by a factor 3.5 and the keys are rehashed into the enlarged table. This can be a time intensive action so set the value to one that is appropriate to your application.
 - ◆ By keeping the average bucket size to a certain maximum, it can speed up lookup times.
 - ◆ If you set the *trigger* to zero, it will not check to see if the table needs reorganizing. You might do this if you created the original bucket size to an optimal value. See [new](#) on how to do this.

11.3.6.12.5 Example 1:

```
map city_population
city_population = new()
nested_put(city_population, {"United States", "California", "Los Angeles"}, 3819951 )
nested_put(city_population, {"Canada", "Ontario", "Toronto"}, 2503281 )
```

See also: [put](#)

11.3.6.13 remove

```
include std/map.e
public procedure remove(map the_map_p, object the_key_p)
```

Remove an entry with given key from a map.

11.3.6.13.1 Parameters:

1. `the_map_p` : the map to operate on
2. `key` : an object, the key to remove.

11.3.6.13.2 Returns:

The modified **map**.

11.3.6.13.3 Comments:

- If `key` is not on `the_map_p`, the `the_map_p` is returned unchanged.
- If you need to remove all entries, see [clear](#)

11.3.6.13.4 Example 1:

```
map the_map_p
the_map_p = new()
put(the_map_p, "Amy", 66.9)
remove(the_map_p, "Amy")
-- the_map_p is now an empty map again
```

11.3.6.13.5 See Also:

[clear](#), [has](#)

11.3.6.14 clear

```
include std/map.e
public procedure clear(map the_map_p)
```

Remove all entries in a map.

11.3.6.14.1 Parameters:

1. `the_map_p` : the map to operate on

11.3.6.14.2 Comments:

- This is much faster than removing each entry individually.
- If you need to remove just one entry, see [remove](#)

11.3.6.14.3 Example 1:

```
map the_map_p
the_map_p = new()
put(the_map_p, "Amy", 66.9)
put(the_map_p, "Betty", 67.8)
put(the_map_p, "Claire", 64.1)
...
clear(the_map_p)
-- the_map_p is now an empty map again
```

11.3.6.14.4 See Also:

[remove](#), [has](#)

11.3.6.15 size

```
include std/map.e
public function size(map the_map_p)
```

Return the number of entries in a map.

11.3.6.15.1 Parameters:

`the_map_p` : the map being queried

11.3.6.15.2 Returns:

An **integer**, the number of entries it has.

11.3.6.15.3 Comments:

For an empty map, size will be zero

11.3.6.15.4 Example 1:

```
map the_map_p
put(the_map_p, 1, "a")
put(the_map_p, 2, "b")
? size(the_map_p) -- outputs 2
```

11.3.6.15.5 See Also:

[statistics](#)

11.3.6.16 NUM_ENTRIES

```
include std/map.e
public enum NUM_ENTRIES
```

Retrieves characteristics of a map.

11.3.6.16.1 Parameters:

1. `the_map_p` : the map being queried

11.3.6.16.2 Returns:

A **sequence**, of 7 integers:

- `NUM_ENTRIES` -- number of entries
- `NUM_IN_USE` -- number of buckets in use
- `NUM_BUCKETS` -- number of buckets
- `LARGEST_BUCKET` -- size of largest bucket
- `SMALLEST_BUCKET` -- size of smallest bucket
- `AVERAGE_BUCKET` -- average size for a bucket
- `STDEV_BUCKET` -- standard deviation for the bucket length series

11.3.6.16.3 Example 1:

```
sequence s = statistics(mymap)
printf(1, "The average size of the buckets is %d", s[AVERAGE_BUCKET])
```

11.3.6.17 statistics

```
include std/map.e
public function statistics(map the_map_p)
```

11.3.6.18 keys

```
include std/map.e
public function keys(map the_map_p, integer sorted_result = 0)
```

Return all keys in a map.

Parameters;

1. `the_map_p`: the map being queried
2. `sorted_result`: optional integer. 0 [default] means do not sort the output and 1 means to sort the output before returning.

11.3.6.18.1 Returns:

A **sequence** made of all the keys in the map.

11.3.6.18.2 Comments:

If `sorted_result` is not used, the order of the keys returned is not predicable.

11.3.6.18.3 Example 1:

```
map the_map_p
the_map_p = new()
put(the_map_p, 10, "ten")
put(the_map_p, 20, "twenty")
put(the_map_p, 30, "thirty")
put(the_map_p, 40, "forty")

sequence keys
keys = keys(the_map_p) -- keys might be {20,40,10,30} or some other order
keys = keys(the_map_p, 1) -- keys will be {10,20,30,40}
```

11.3.6.18.4 See Also:

[has](#), [values](#), [pairs](#)

11.3.6.19 values

```
include std/map.e
public function values(map the_map, object keys = 0, object default_values = 0)
```

Return values, without their keys, from a map.

11.3.6.19.1 Parameters:

1. `the_map` : the map being queried
2. `keys` : optional, key list of values to return.
3. `default_values` : optional default values for keys list

11.3.6.19.2 Returns:

A **sequence**, of all values stored in `the_map`.

11.3.6.19.3 Comments:

- The order of the values returned may not be the same as the putting order.
- Duplicate values are not removed.
- You use the `keys` parameter to return a specific set of values from the map. They are returned in the same order as the `keys` parameter. If no `default_values` is given and one is needed, 0 will be used.
- If `default_values` is an atom, it represents the default value for all values in `keys`.
- If `default_values` is a sequence, and its length is less than `keys`, then the last item in `default_values` is used for the rest of the keys.

11.3.6.19.4 Example 1:

```
map the_map_p
the_map_p = new()
put(the_map_p, 10, "ten")
put(the_map_p, 20, "twenty")
put(the_map_p, 30, "thirty")
put(the_map_p, 40, "forty")

sequence values
values = values(the_map_p)
-- values might be {"twenty", "forty", "ten", "thirty"}
-- or some other order
```

11.3.6.19.5 Example 2:

```
map the_map_p
the_map_p = new()
put(the_map_p, 10, "ten")
put(the_map_p, 20, "twenty")
put(the_map_p, 30, "thirty")
put(the_map_p, 40, "forty")

sequence values
values = values(the_map_p, { 10, 50, 30, 9000 })
-- values WILL be { "ten", 0, "thirty", 0 }
values = values(the_map_p, { 10, 50, 30, 9000 }, {-1,-2,-3,-4})
-- values WILL be { "ten", -2, "thirty", -4 }
```

11.3.6.19.6 See Also:

[get](#), [keys](#), [pairs](#)

11.3.6.20 pairs

```
include std/map.e
public function pairs(map the_map_p, integer sorted_result = 0)
```

Return all key/value pairs in a map.

11.3.6.20.1 Parameters:

1. `the_map_p` : the map to get the data from
2. `sorted_result` : optional integer. 0 [default] means do not sort the output and 1 means to sort the output before returning.

11.3.6.20.2 Returns:

A **sequence**, of all key/value pairs stored in `the_map_p`. Each pair is a sub-sequence in the form {key, value}

11.3.6.20.3 Comments:

If `sorted_result` is not used, the order of the values returned is not predicable.

11.3.6.20.4 Example 1:

```
map the_map_p
the_map_p = new()
put(the_map_p, 10, "ten")
put(the_map_p, 20, "twenty")
put(the_map_p, 30, "thirty")
put(the_map_p, 40, "forty")

sequence keyvals
keyvals = pairs(the_map_p) -- might be {{20,"twenty"},{40,"forty"},{10,"ten"},{30,"thirty"}}
keyvals = pairs(the_map_p, 1) -- will be {{10,"ten"},{20,"twenty"},{30,"thirty"},{40,"forty"}}
```

11.3.6.20.5 See Also:

[get](#), [keys](#), [values](#)

11.3.6.21 optimize

```
include std/map.e
public procedure optimize(map the_map_p, integer max_p = 25, atom grow_p = 1.333)
```

Widens a map to increase performance.

11.3.6.21.1 Parameters:

1. `the_map_p` : the map being optimized
2. `max_p` : an integer, the maximum desired size of a bucket. Default is 25. This must be 3 or higher.
3. `grow_p` : an atom, the factor to grow the number of buckets for each iteration of rehashing. Default is 1.333. This must be greater than 1.

11.3.6.21.2 Returns:

The optimized **map**.

11.3.6.21.3 Comments:

This rehashes the map until either the maximum bucket size is less than the desired maximum or the maximum bucket size is less than the largest size statistically expected (mean + 3 standard deviations).

11.3.6.21.4 See Also:

[statistics](#), [rehash](#)

11.3.6.22 load_map

```
include std/map.e
public function load_map(object file_name_p)
```

Loads a map from a file

11.3.6.22.1 Parameters:

1. `file_name_p` : The file to load from. This file may have been created by the [save_map](#) function. This can either be a name of a file or an already opened file handle.

11.3.6.22.2 Returns:

Either a **map**, with all the entries found in `file_name_p`, or **-1** if the file failed to open.

11.3.6.22.3 Comments:

If `file_name_p` is an already opened file handle, this routine will write to that file and not close it. Otherwise, the named file will be created and closed by this routine.

The input file can be either one created by the [save_map](#) function or a manually created/edited text file. See [save_map](#) for details about the required layout of the text file.

11.3.6.22.4 Example 1:

```
object loaded
map AppOptions
sequence SavedMap = "c:\myapp\options.txt"
loaded = load_map(SavedMap)
if equal(loaded, -1) then
    crash("Map '%s' failed to open", SavedMap)
end if
-- By now we know that it was loaded and a new map created,
-- so we can assign it to a 'map' variable.
AppOptions = loaded
if get(AppOptions, "verbose", 1) = 3 then
    ShowInstructions()
end if
```

11.3.6.22.5 See Also:

[new](#), [save_map](#)

11.3.6.23 SM_TEXT

```
include std/map.e
public enum SM_TEXT
```

Saves a map to a file.

11.3.6.23.1 Parameters:

1. *m* : a map.
2. *file_name_p* : Either a sequence, the name of the file to save to, or an open file handle as returned by [open\(\)](#).
3. *type* : an integer. SM_TEXT for a human-readable format (default), SM_RAW for a smaller and faster format, but not human-readable.

11.3.6.23.2 Returns:

An **integer**, the number of keys saved to the file, or -1 if the save failed.

11.3.6.23.3 Comments:

If `file_name_p` is an already opened file handle, this routine will write to that file and not close it. Otherwise, the named file will be created and closed by this routine.

The `SM_TEXT` type saves the map keys and values in a text format which can be read and edited by standard text editor. Each entry in the map is saved as a `KEY/VALUE` pair in the form

```
key = value
```

Note that if the 'key' value is a normal string value, it can be enclosed in double quotes. If it is not thus quoted, the first character of the key determines its EUPHORIA value type. A dash or digit implies an atom, an left-brace implies a sequence, an alphabetic character implies a text string that extends to the next equal '=' symbol, and anything else is ignored.

Note that if a line contains a double-dash, then all text from the double-dash to the end of the line will be ignored. This is so you can optionally add comments to the saved map. Also, any blank lines are ignored too.

All text after the '=' symbol is assumed to be the map item's value data.

The `SM_RAW` type saves the map in an efficient manner. It is generally smaller than the text format and is faster to process, but it is not human readable and standard text editors can not be used to edit it. In this format, the file will

11.3.6.23.4 contain three serialized sequences:

1. Header sequence: {integer:format version, string: date and time of save (YYMMDDhhmmss), sequence: euphoria version {major, minor, revision, patch}}
2. Keys. A list of all the keys
3. Values. A list of the corresponding values for the keys.

11.3.6.23.5 Example 1:

```
map AppOptions
if save_map(AppOptions, "c:\myapp\options.txt") = -1
    Error("Failed to save application options")
end if
if save_map(AppOptions, "c:\myapp\options.dat", SM_RAW) = -1
    Error("Failed to save application options")
end if
```

11.3.6.23.6 See Also:

[load_map](#)

11.3.6.24 save_map

```
include std/map.e
public function save_map(map the_map_, object file_name_p, integer type_ = SM_TEXT)
```

11.3.6.25 copy

```
include std/map.e
public function copy(map source_map, object dest_map = 0, integer put_operation = PUT)
```

Duplicates a map.

11.3.6.25.1 Parameters:

1. source_map : map to copy from
2. dest_map : optional, map to copy to
3. put_operation : optional, operation to use when dest_map## is used. The default is PUT.

11.3.6.25.2 Returns:

If dest_map was not provided, an exact duplicate of source_map otherwise dest_map, which does not have to be empty, is returned with the new values copied from source_map, according to the put_operation value.

11.3.6.25.3 Example 1:

```
map m1 = new()
put(m1, 1, "one")
put(m1, 2, "two")

map m2 = copy(m1)
printf(1, "%s, %s\n", { get(m2, 1), get(m2, 2) })
-- one, two

put(m1, 1, "one hundred")
printf(1, "%s, %s\n", { get(m1, 1), get(m1, 2) })
-- one hundred, two

printf(1, "%s, %s\n", { get(m2, 1), get(m2, 2) })
-- one, two
```

11.3.6.25.4 Example 2:

```
map m1 = new()
map m2 = new()

put(m1, 1, "one")
```

```
put(m1, 2, "two")
put(m2, 3, "three")

copy(m1, m2)

? keys(m2)
-- { 1, 2, 3 }
```

11.3.6.25.5 Example 3:

```
map m1 = new()
map m2 = new()

put(m1, "XY", 1)
put(m1, "AB", 2)
put(m2, "XY", 3)

? pairs(m1)  -- { {"AB", 2}, {"XY", 1} }
? pairs(m2)  -- { {"XY", 3} }

-- Add same keys' values.
copy(m1, m2, ADD)

? pairs(m2)
-- { {"AB", 2}, {"XY", 4} }
```

11.3.6.25.6 See Also:

[put](#)

11.3.6.26 new_from_kvpairs

```
include std/map.e
public function new_from_kvpairs(sequence kv_pairs)
```

Converts a set of Key-Value pairs to a map.

11.3.6.26.1 Parameters:

1. `kv_pairs`: A sequence containing any number of subsequences that have the format {KEY, VALUE}. These are loaded into a new map which is then returned by this function.

11.3.6.26.2 Returns:

A **map**, containing the data from `kv_pairs`

11.3.6.26.3 Example 1:

```
map m1 = new_from_kvpairs( {
    {"application", "EUPHORIA"},
    {"version", "4.0"},
    {"genre", "programming language"},
    {"crc", 0x4F71AE10}
})

v = map:get(m1, "application") --> "EUPHORIA"
```

11.3.6.27 new_from_string

```
include std/map.e
public function new_from_string(sequence kv_string)
```

Converts a set of Key-Value pairs contained in a string to a map.

11.3.6.27.1 Parameters:

1. `kv_string`: A string containing any number of lines that have the format KEY=VALUE. These are loaded into a new map which is then returned by this function.

11.3.6.27.2 Returns:

A **map**, containing the data from `kv_string`

11.3.6.27.3 Comment:

This function actually calls `keyvalues()` to convert the string to key-value pairs, which are then used to create the map.

11.3.6.27.4 Example 1:

Given that a file called "xyz.config" contains the lines ...

```
application = EUPHORIA,
version      = 4.0,
genre        = "programming language",
crc          = 4F71AE10

map m1 = new_from_string( read_file("xyz.config", TEXT_MODE))

printf(1, "%s\n", {map:get(m1, "application")}) --> "EUPHORIA"
printf(1, "%s\n", {map:get(m1, "genre")})       --> "programming language"
printf(1, "%s\n", {map:get(m1, "version")})     --> "4.0"
printf(1, "%s\n", {map:get(m1, "crc")})         --> "4F71AE10"
```

11.3.6.28 for_each

```
include std/map.e
public function for_each(map source_map, integer user_rid, object user_data = 0, integer in_son
```

Calls a user-defined routine for each of the items in a map.

11.3.6.28.1 Parameters:

1. `source_map`: The map containing the data to process
2. `user_rid`: The routine_id of a user defined processing function
3. `user_data`: An object. Optional. This is passed, unchanged to each call of the user defined routine. By default, zero (0) is used.
4. `in_sorted_order`: An integer. Optional. If non-zero the items in the map are processed in ascending key sequence otherwise the order is undefined. By default they are not sorted.

11.3.6.28.2 Returns:

An integer: 0 means that all the items were processed, and anything else is whatever was returned by the user routine to abort the `for_each()` process.

11.3.6.28.3 Comment:

- The user defined routine is a function that must accept four parameters.
 1. Object: an Item Key
 2. Object: an Item Value
 3. Object: The `user_data` value. This is never used by `for_each()` itself, merely passed to the user routine.
 4. Integer: Progress code.
 - ◊ The `abs()` value of the progress code is the ordinal call number. That is 1 means the first call, 2 means the second call, etc ...
 - ◊ If the progress code is negative, it is also the last call to the routine.
 - ◊ If the progress code is zero, it means that the map is empty and thus the item key and value cannot be used.
- The user routine must return 0 to get the next map item. Anything else will cause `for_each()` to stop running, and is returned to whatever called `for_each()`.

11.3.6.28.4 Example 1:

```
include std/map.e
include std/math.e
include std/io.e
function Process_A(object k, object v, object d, integer pc)
```

```
    writeln("[ ] = [ ]", {k, v})
    return 0
end function

function Process_B(object k, object v, object d, integer pc)
    if pc = 0 then
        writeln("The map is empty")
    else
        integer c
        c = abs(pc)
        if c = 1 then
            writeln("---[ ]---", {d}) -- Write the report title.
        end if
        writeln("[ ]: [:15] = [ ]", {c, k, v})
        if pc < 0 then
            writeln(repeat('-', length(d) + 6), {}) -- Write the report end.
        end if
    end if
    return 0
end function

map m1 = new()
map:put(m1, "application", "EUPHORIA")
map:put(m1, "version", "4.0")
map:put(m1, "genre", "programming language")
map:put(m1, "crc", "4F71AE10")

-- Unsorted
map:for_each(m1, routine_id("Process_A"))
-- Sorted
map:for_each(m1, routine_id("Process_B"), "List of Items", 1)
```

The output from the first call could be...

```
application = EUPHORIA
version = 4.0
genre = programming language
crc = 4F71AE10
```

The output from the second call should be...

```
---List of Items---
1: application      = EUPHORIA
2: crc              = 4F71AE10
3: genre            = programming language
4: version          = 4.0
-----
```

11.4 Stack

Page Contents

[Constants](#)

FIFO

Types

stack

Routines

new

is_empty

size

at

push

top

last

pop

peek_top

peek_end

swap

dup

set

clear

11.4.1 Constants

11.4.1.1 FIFO

```
include std/stack.e
public constant FIFO
```

Stack types

- FIFO: like people standing in line: first item in is first item out
- FILO: like for a stack of plates : first item in is last item out

11.4.2 Types

11.4.2.1 stack

```
include std/stack.e
public type stack(object obj_p)
```

A stack is a sequence of objects with some internal data.

11.4.3 Routines

11.4.3.1 new

```
include std/stack.e
public function new(integer typ = FILO)
```

Create a new stack.

11.4.3.1.1 Parameters:

1. `stack_type` : an integer, defining the semantics of the stack. The default is FILO.

11.4.3.1.2 Returns:

An empty **stack**, note that the variable storing the stack must not be an integer. The resources allocated for the stack will be automatically cleaned up if the reference count of the returned value drops to zero, or if passed in a call to [delete](#).

11.4.3.1.3 Comments:

There are two sorts of stacks, designated by the types FILO and FILO:

- A FILO stack is one where the first item to be pushed is popped first. People standing in queue form a FILO stack.
- A FILO stack is one where the item pushed last is popped first. A column of coins is of the FILO kind.

11.4.3.1.4 See Also:

[is_empty](#)

11.4.3.2 is_empty

```
include std/stack.e
public function is_empty(stack sk)
```

Determine whether a stack is empty.

11.4.3.2.1 Parameters:

1. `sk` : the stack being queried.

11.4.3.2.2 Returns:

An **integer**, 1 if the stack is empty, else 0.

11.4.3.2.3 See Also:

[size](#)

11.4.3.3 size

```
include std/stack.e
public function size(stack sk)
```

Returns how many elements a stack has.

11.4.3.3.1 Parameters:

1. `sk` : the stack being queried.

11.4.3.3.2 Returns:

An **integer**, the number of elements in `sk`.

11.4.3.4 at

```
include std/stack.e
public function at(stack sk, integer idx = 1)
```

Fetch a value from the stack without removing it from the stack.

11.4.3.4.1 Parameters:

1. `sk` : the stack being queried
2. `idx` : an integer, the place to inspect. The default is 1 (top item).

11.4.3.4.2 Returns:

An **object**, the `idx`-th item of the stack.

11.4.3.4.3 Errors:

If the supplied value of `idx` does not correspond to an existing element, an error occurs.

11.4.3.4.4 Comments:

- For `FIFO` stacks (queues), the top item is the oldest item in the stack.
- For `FILO` stacks, the top item is the newest item in the stack.

`idx` can be less than 1, in which case it refers relative to the end item. Thus, 0 stands for the end element.

11.4.3.4.5 Example 1:

```
stack sk = new(FILO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
? at(sk, 0) -- 5
? at(sk, -1) -- "abc"
? at(sk, 1) -- 2.3
? at(sk, 2) -- "abc"
```

11.4.3.4.6 Example 2:

```
stack sk = new(FIFO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
? at(sk, 0) -- 2.3
? at(sk, -1) -- "abc"
? at(sk, 1) -- 5
? at(sk, 2) -- "abc"
```

11.4.3.4.7 See Also:

[size](#), [top](#), [peek_top](#), [peek_end](#)

11.4.3.5 push

```
include std/stack.e
public procedure push(stack sk, object value)
```

Adds something to a stack.

11.4.3.5.1 Parameters:

1. `sk` : the stack to augment
2. `value` : an object, the value to push.

11.4.3.5.2 Comments:

`value` appears at the end of `FIFO` stacks and the top of `FILO` stacks. The size of the stack increases by one.

11.4.3.5.3 Example 1:

```
stack sk = new(FIFO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
? top(sk) -- 5
? last(sk) -- 2.3
```

11.4.3.5.4 Example 2:

```
stack sk = new(FILO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
? top(sk) -- 2.3
? last(sk) -- 5
```

11.4.3.5.5 See Also:

[pop](#), [top](#)

11.4.3.6 top

```
include std/stack.e
public function top(stack sk)
```

Retrieve the top element on a stack.

11.4.3.6.1 Parameters:

1. `sk` : the stack to inspect.

11.4.3.6.2 Returns:

An **object**, the top element on a stack.

11.4.3.6.3 Comments:

This call is equivalent to `at (sk, 1)`.

11.4.3.6.4 Example 1:

```
stack sk = new(FILO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
? top(sk) -- 2.3
```

11.4.3.6.5 Example 1:

```
stack sk = new(FIFO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
? top(sk) -- 5
```

11.4.3.6.6 See Also:

[at](#), [pop](#), [peek_top](#), [last](#)

11.4.3.7 last

```
include std/stack.e
public function last(stack sk)
```

Retrieve the end element on a stack.

11.4.3.7.1 Parameters:

1. `sk` : the stack to inspect.

11.4.3.7.2 Returns:

An **object**, the end element on a stack.

11.4.3.7.3 Comments:

This call is equivalent to `at (sk, 0)`.

11.4.3.7.4 Example 1:

```
stack sk = new(FILO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
? last(sk) -- 5
```

11.4.3.7.5 Example 2:

```
stack sk = new(FIFO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
? last(sk) -- 2.3
```

11.4.3.7.6 See Also:

[at](#), [pop](#), [peek_end](#), [top](#)

11.4.3.8 pop

```
include std/stack.e
public function pop(stack sk, integer idx = 1)
```

Removes an object from a stack.

11.4.3.8.1 Parameters:

1. `sk` : the stack to pop
2. `idx` : integer. The *n*-th item to pick from the stack. The default is 1.

11.4.3.8.2 Returns:

An **item**, from the stack, which is also removed from the stack.

11.4.3.8.3 Errors:

- If the stack is empty, an error occurs.
- If the `idx` is greater than the number of items in the stack, an error occurs.

11.4.3.8.4 Comments:

- For FIFO stacks (queues), the top item is the oldest item in the stack.
- For FILO stacks, the top item is the newest item in the stack.

When `idx` is omitted the 'top' of the stack is removed and returned. When `idx` is supplied, it represents the N-th item from the top to be removed and returned. Thus an `idx` of 2 returns the 2nd item from the top, a value of 3 returns the 3rd item from the top, etc ...

11.4.3.8.5 Example 1:

```
stack sk = new(FIFO)
push(sk, 1)
push(sk, 2)
push(sk, 3)
? size(sk) -- 3
? pop(sk) -- 1
? size(sk) -- 2
? pop(sk) -- 2
? size(sk) -- 1
? pop(sk) -- 3
? size(sk) -- 0
? pop(sk) -- *error*
```

11.4.3.8.6 Example 2:

```
stack sk = new(FILO)
push(sk, 1)
push(sk, 2)
push(sk, 3)
? size(sk) -- 3
? pop(sk) -- 3
? size(sk) -- 2
? pop(sk) -- 2
? size(sk) -- 1
? pop(sk) -- 1
? size(sk) -- 0
? pop(sk) -- *error*
```

11.4.3.8.7 Example 3:

```
stack sk = new(FILO)
push(sk, 1)
push(sk, 2)
push(sk, 3)
push(sk, 4)
-- stack contains {1,2,3,4} (oldest to newest)
? size(sk) -- 4
? pop(sk, 2) -- Pluck out the 2nd newest item .. 3
? size(sk) -- 3
-- stack now contains {1,2,4}
```

11.4.3.8.8 Example 4:

```
stack sk = new(FIFO)
push(sk, 1)
push(sk, 2)
push(sk, 3)
push(sk, 4)
-- stack contains {1,2,3,4} (oldest to newest)
? size(sk) -- 4
? pop(sk, 2) -- Pluck out the 2nd oldest item .. 2
? size(sk) -- 3
-- stack now contains {1,3,4}
```

11.4.3.8.9 See Also:

[push](#), [top](#), [is_empty](#)

11.4.3.9 peek_top

```
include std/stack.e
public function peek_top(stack sk, integer idx = 1)
```

Gets an object, relative to the top, from a stack.

11.4.3.9.1 Parameters:

1. `sk` : the stack to get from.
2. `idx` : integer. The n-th item to get from the stack. The default is 1.

11.4.3.9.2 Returns:

An **item**, from the stack, which is **not** removed from the stack.

11.4.3.9.3 Errors:

- If the stack is empty, an error occurs.
- If the `idx` is greater than the number of items in the stack, an error occurs.

11.4.3.9.4 Comments:

This is identical to [pop](#) except that it does not remove the item.

- For `FIFO` stacks (queues), the top item is the oldest item in the stack.
- For `FILO` stacks, the top item is the newest item in the stack.

When `idx` is omitted the 'top' of the stack is returned. When `idx` is supplied, it represents the N-th item from the top to be returned. Thus an `idx` of 2 returns the 2nd item from the top, a value of 3 returns the 3rd item from the top, etc ...

11.4.3.9.5 Example 1:

```
stack sk = new(FIFO)
push(sk, 1)
push(sk, 2)
push(sk, 3)
? peek_top(sk) -- 1
? peek_top(sk, 2) -- 2
? peek_top(sk, 3) -- 3
? peek_top(sk, 4) -- *error*
? peek_top(sk, size(sk)) -- 3 (end item)
```

11.4.3.9.6 Example 2:

```
stack sk = new(FILO)
push(sk, 1)
push(sk, 2)
push(sk, 3)
? peek_top(sk) -- 3
? peek_top(sk, 2) -- 2
? peek_top(sk, 3) -- 1
? peek_top(sk, 4) -- *error*
? peek_top(sk, size(sk)) -- 1 (end item)
```

11.4.3.9.7 See Also:

[pop](#), [top](#), [is_empty](#), [size](#), [peek_end](#)

11.4.3.10 peek_end

```
include std/stack.e
public function peek_end(stack sk, integer idx = 1)
```

Gets an object, relative to the end, from a stack.

11.4.3.10.1 Parameters:

1. `sk` : the stack to get from.
2. `idx` : integer. The n-th item from the end to get from the stack. The default is 1.

11.4.3.10.2 Returns:

An **item**, from the stack, which is **not** removed from the stack.

11.4.3.10.3 Errors:

- If the stack is empty, an error occurs.
- If the `idx` is greater than the number of items in the stack, an error occurs.

11.4.3.10.4 Comments:

- For `FIFO` stacks (queues), the end item is the newest item in the stack.
- For `FILO` stacks, the end item is the oldest item in the stack.

When `idx` is omitted the 'end' of the stack is returned. When `idx` is supplied, it represents the N-th item from the end to be returned. Thus an `idx` of 2 returns the 2nd item from the end, a value of 3 returns the 3rd item from the end, etc ...

11.4.3.10.5 Example 1:

```
stack sk = new(FIFO)
push(sk, 1)
push(sk, 2)
push(sk, 3)
? peek_end(sk) -- 3
? peek_end(sk, 2) -- 2
? peek_end(sk, 3) -- 1
? peek_end(sk, 4) -- *error*
? peek_end(sk, size(sk)) -- 3 (top item)
```

11.4.3.10.6 Example 2:

```
stack sk = new(FILO)
push(sk, 1)
push(sk, 2)
push(sk, 3)
? peek_end(sk) -- 1
? peek_end(sk, 2) -- 2
? peek_end(sk, 3) -- 3
? peek_end(sk, 4) -- *error*
? peek_end(sk, size(sk)) -- 3 (top item)
```

11.4.3.10.7 See Also:

[pop](#), [top](#), [is_empty](#), [size](#), [peek_top](#)

11.4.3.11 swap

```
include std/stack.e
public procedure swap(stack sk)
```

Swap the top two elements of a stack

11.4.3.11.1 Parameters:

1. `sk` : the stack to swap.

11.4.3.11.2 Returns:

A **copy**, of the original **stack**, with the top two elements swapped.

11.4.3.11.3 Comments:

- For `FIFO` stacks (queues), the top item is the oldest item in the stack.
- For `FILO` stacks, the top item is the newest item in the stack.

11.4.3.11.4 Errors:

If the stack has less than two elements, an error occurs.

11.4.3.11.5 Example 1:

```
stack sk = new(FILO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
push(sk, "")
? peek_top(sk, 1)  -- ""
? peek_top(sk, 2)  -- 2.3
swap(sk)
? peek_top(sk, 1)  -- 2.3
? peek_top(sk, 2)  -- ""
```

11.4.3.11.6 Example 2:

```
stack sk = new(FIFO)
push(sk, 5)
push(sk, "abc")
push(sk, 2.3)
push(sk, "")
? peek_top(sk, 1)  -- 5
? peek_top(sk, 2)  -- "abc"
swap(sk)
```

```
? peek_top(sk,1)  -- "abc"
? peek_top(sk,2)  -- 5
```

11.4.3.12 dup

```
include std/stack.e
public procedure dup(stack sk)
```

Repeat the top element of a stack.

11.4.3.12.1 Parameters:

1. `sk` : the stack.

11.4.3.12.2 Side effects:

The value of `top()` is pushed onto the stack, thus the stack size grows by one.

11.4.3.12.3 Comments:

- For FIFO stacks (queues), the top item is the oldest item in the stack.
- For FILO stacks, the top item is the newest item in the stack.

11.4.3.12.4 Errors:

If the stack has no elements, an error occurs.

11.4.3.12.5 Example 1:

```
stack sk = new(FILO)
push(sk,5)
push(sk,"abc")
push(sk, "")
dup(sk)
? peek_top(sk,1)  -- ""
? peek_top(sk,2)  -- "abc"
? size(sk)        -- 3
dup(sk)
? peek_top(sk,1)  -- ""
? peek_top(sk,2)  -- ""
? peek_top(sk,3)  -- "abc"
? size(sk)        -- 4
```

11.4.3.12.6 Example 1:

```
stack sk = new(FIFO)
push(sk, 5)
push(sk, "abc")
push(sk, "")
dup(sk)
? peek_top(sk, 1)  -- 5
? peek_top(sk, 2)  -- "abc"
? size(sk)        -- 3
dup(sk)
? peek_top(sk, 1)  -- 5
? peek_top(sk, 2)  -- 5
? peek_top(sk, 3)  -- "abc"
? size(sk)        -- 4
```

11.4.3.13 set

```
include std/stack.e
public procedure set(stack sk, object val, integer idx = 1)
```

Update a value on the stack

11.4.3.13.1 Parameters:

1. `sk` : the stack being queried
2. `val` : an object, the value to place on the stack
3. `idx` : an integer, the place to inspect. The default is 1 (the top item)

11.4.3.13.2 Errors:

If the supplied value of `idx` does not correspond to an existing element, an error occurs.

11.4.3.13.3 Comments:

- For `FIFO` stacks (queues), the top item is the oldest item in the stack.
- For `FILO` stacks, the top item is the newest item in the stack.

`idx` can be less than 1, in which case it refers to an element relative to the end of the stack. Thus, 0 stands for the end element.

11.4.3.13.4 See Also:

[size](#), [top](#)

11.4.3.14 clear

```
include std/stack.e
public procedure clear(stack sk)
```

Wipe out a stack.

11.4.3.14.1 Parameters:

1. `sk` : the stack to clear.

11.4.3.14.2 Side effect:

The stack contents is emptied.

11.4.3.14.3 See Also:

[new](#), [is_empty](#)

11.5 Sets

Types

- [set](#)
- [map](#)
- [operation](#)

Inclusion and belonging.

- [sequence_to_set](#)
- [cardinal](#)
- [belongs_to](#)
- [add_to](#)
- [remove_from](#)
- [is_subset](#)
- [embedding](#)
- [embed_union](#)
- [subsets](#)

Basic set-theoretic operations.

- [intersection](#)
- [union](#)
- [delta](#)
- [difference](#)
- [product](#)

Maps between sets.

- [define_map](#)
- [sequences_to_map](#)
- [image](#)

```

range
direct_map
restrict
change_target
combine_maps
compose_map
diagram_commutates
is_injective
is_surjective
is_bijective
Reverse mappings
  fiber_over
  reverse_map
  section
Products
  product_map
  amalgamated_sum
  fiber_product
Constants
  SIDE_NONE
Operations on sets
  define_operation
  is_symmetric
  is_associative
  all_left_units
  all_right_units
  has_unit
  is_unit
  has_inverse
  distributes_over

```

The sets.e module defines a type for sets and provides basic tools for handling them.

Other modules may be built upon them, for instance graph handling or simple topology, finite groups etc.

11.5.1 Notes:

- A *set* is an ordered sequence in ascending order, not more, not less
- A *map* from setA to setB is a sequence the length of setA whose elements are indexes into setB, followed by {length(setA),length(setB)}.
- An *operation* of $E \times F \implies G$ is a two dimensional sequence of elements of G, indexed by $E \times F$, and the triple {card(E),card(F),card(G)}.

11.5.2 Types

11.5.2.1 set

```
include std/sets.e
public type set(object s)
```

A set is a sequence in which each item is greater than the previous item.

11.5.2.1.1 See Also:

[compare](#)

11.5.2.2 map

```
include std/sets.e
public type map(object s)
```

Returns 1 if a sequence of integers is a valid map descriptor, else 0.

11.5.2.2.1 Comments:

A map is a sequence of indexes. Each index is between 1 and the maximum allowed for the particular map.

Actually, what is being called a `map` is a class of maps, as the elements of the input sequence, except for the last two, are ordinals rather than set elements. A map contains the information required to map as expected the elements of a set, given by index, to another set, where the images are indexes again. Technically, those are maps of the category of finite sets quotiented by equality of cardinal.

The objects that `map.e` handle are completely unrelated to these.

11.5.2.2.2 Example 1:

```
sequence s0 = {2, 3, 4, 1, 4, 2, 6, 4}
? map(s0)    -- prints out 1.
```

11.5.2.2.3 See also:

[define_map](#), [fiber_over](#), [restrict](#), [direct_map](#), [\[\[:reverse_map\]](#), [is_injective](#), [is_surjective](#), [is_bijective](#)

11.5.2.3 operation

```
include std/sets.e
public type operation(object s)
```

Returns 1 if the data represents a map from the product of two sets to a third one.

11.5.2.3.1 Comments:

An operation from $F \times G$ to H is defined as a sequence of mappings from G to H , plus the cardinals of the sets F , G and H . If the input data is consistent with this description, 1 is returned, else 0.

11.5.2.3.2 Example 1:

```
sequence s = {{2, 3}, {3, 1}, {1, 2}, {2, 3}, {3, 1}}, {5,2,3}
-- s represents the addition modulo 3 from {0, 1, 2, 3, 4} x {1, 2} to {0, 1, 2}
? operation(s)    -- prints out 1.
```

11.5.3 Inclusion and belonging.

11.5.3.1 sequence_to_set

```
include std/sets.e
public function sequence_to_set(sequence s)
```

Makes a set out of a sequence by sorting it and removing duplicate elements.

11.5.3.1.1 Parameters:

1. s : the sequence to transform.

11.5.3.1.2 Returns:

A **set**, which is the ordered list of distinct elements in s .

11.5.3.1.3 Example 1:

```
sequence s0 s0={1,3,7,5,7,4,1}
set s1 s1=sequence_to_set(s0)    -- s1 is now {1,3,4,5,7}
```

11.5.3.1.4 See also:

[set](#)

11.5.3.2 cardinal

```
include std/sets.e
public function cardinal(set S)
```

Return the cardinal of a set

11.5.3.2.1 Parameters:

1. S : the set being queried.

11.5.3.2.2 Returns:

An **integer**, the count of elements in S.

11.5.3.2.3 See Also:

[set](#)

11.5.3.3 belongs_to

```
include std/sets.e
public function belongs_to(object x, set s)
```

Decide whether an object is in a set.

11.5.3.3.1 Parameters:

1. x : the object inquired about
2. S : the set being queried

11.5.3.3.2 Returns:

An **integer**, 1 if x is in S, else 0.

11.5.3.3.3 Example 1:

```
set s0 s0={1,3,5,7}
?belongs_to(2,s)    -- prints out 0
```


11.5.3.3.4 See Also:

[is_subset](#) , [intersection](#), [difference](#)

11.5.3.4 add_to

```
include std/sets.e
public function add_to(object x, set S)
```

Add an object to a set.

11.5.3.4.1 Parameters:

1. *x* : the object to add
2. *S* : the set to augment

11.5.3.4.2 Returns:

A **set**, which is a **copy** of *S*, with the addition of *x* if it was not there already.

11.5.3.4.3 Example 1:

```
set s0 s0={1,3,5,7}
s0=add_to(2,s)      -- s0 is now {1,2,3,5,7}
```

11.5.3.4.4 See Also:

[remove_from](#), [belongs_to](#), [union](#)

11.5.3.5 remove_from

```
include std/sets.e
public function remove_from(object x, set s)
```

Remove an object from a set.

11.5.3.5.1 Parameters:

1. *x* : the object to add
2. *S* : the set to remove from

11.5.3.5.2 Returns:

A **set**, which is a **copy** of `S`, with `x` removed if it was there.

11.5.3.5.3 Example 1:

```
set s0 s0={1,2,3,5,7}
s0=remove_from(2,s0)  -- s0 is now {1,3,5,7}
```

11.5.3.5.4 See Also:

[remove_from](#), [belongs_to](#), [union](#)

11.5.3.6 is_subset

```
include std/sets.e
public function is_subset(set small, set large)
```

Checks whether a set is a subset of another.

11.5.3.6.1 Parameters:

1. `small` : the set to test
2. `large` : the supposedly larger set.

11.5.3.6.2 Returns:

An **integer**, 1 if `small` is a subset of `large`, else 0.

11.5.3.6.3 Example 1:

```
set s0 s0={1,3,5,7}
? is_subset({3,5},s0)  -- prints out 1
```

11.5.3.6.4 See Also:

[subsets](#), [belongs_to](#), [difference](#), [embedding](#), [embed_union](#)

11.5.3.7 embedding

```
include std/sets.e
public function embedding(set small, set large)
```

Returns the set of indexes of the elements of a set in a larger set, or 0 if not applicable

11.5.3.7.1 Parameters:

1. `small` : the set to embed
2. `large` : the supposedly larger set

11.5.3.7.2 Returns:

A **set**, of indexes if `small` [is_subset\(\)](#) `large`, else 0. Each element is the index in `large` of the corresponding element of `small`. Its length is `length(small)` and the values range from 1 to `length(large)`.

11.5.3.7.3 Example 1:

```
set s0 s0={1,3,5,7}
set s s=embedding({3,5},s0)  -- s is now {2,3}
```

11.5.3.7.4 See Also:

[subsets](#), [belongs_to](#), [difference](#), [is_subset](#)

11.5.3.8 embed_union

```
include std/sets.e
public function embed_union(set s1, set s2)
```

Returns the embedding of a set into its union with another.

11.5.3.8.1 Parameters:

1. `S1` : the set to embed
2. `S2` : the other set

11.5.3.8.2 Returns:

A **set**, of indexes representing `S1` inside `union(S1, S2)`. Its length is `length(S1)`, and the values range from 1 to `length(S1) + length(S2)`.

11.5.3.8.3 Example 1:

```
set s1 = {2, 5, 7}, s2 = {1, 3, 4}
sequence s = embed_union(s1,s2) -- s is now {2, 5, 6}
```

11.5.3.8.4 See Also:

[embedding](#), [union](#)

11.5.3.9 subsets

```
include std/sets.e
public function subsets(set s)
```

Returns the list of all subsets of the input set.

11.5.3.9.1 Parameters:

1. `s` : the set to enumerate the subsets of.

11.5.3.9.2 Returns:

A **sequence**, containing all the subsets of the input set.

11.5.3.9.3 Comments:

`s` must not have more than 29 elements, as the length of the output sequence is `power(2, length(s))`, which rapidly grows out of integer range. The order in which the subsets are output is implementation dependent.

11.5.3.9.4 Example 1:

```
set s0 s0={1,3,5,7}
s0=subsets(s0) -- s0 is now:
{{}, {1}, {3}, {5}, {7}, {1,3}, {1,5}, {1,7}, {3,5}, {3,7}, {5,7}, {1,3,5}, {1,3,7}, {1,5,7}, {3,5,7}, {1,3,5,7}}
```

11.5.3.9.5 See Also:

[is_subset](#)

11.5.4 Basic set-theoretic operations.

11.5.4.1 intersection

```
include std/sets.e
public function intersection(set S1, set S2)
```

Returns the set of elements belonging to both s1 and s2.

11.5.4.1.1 Parameters:

1. S1 : One of the sets to intersect
2. S2 : the other set.

11.5.4.1.2 Returns:

A **set**, made of all elements belonging to both S1 and S2.

11.5.4.1.3 Example 1:

```
set s0,s1,s2
s1={1,3,5,7} s2={-1,2,3,7,11}
s0=intersection(s1,s2)  -- s0 is now {3,7}.
```

11.5.4.1.4 See Also:

[is_subset](#), [subsets](#), [belongs_to](#)

11.5.4.2 union

```
include std/sets.e
public function union(set S1, set S2)
```

Returns the set of elements belonging to any of two sets.

11.5.4.2.1 Parameters:

1. S1: one of the sets to merge
2. S2: the other set.

11.5.4.2.2 Returns:

The **set** of all elements belonging to S1 or S2, and possibly to both.

11.5.4.2.3 Example 1:

```
set s0,s1,s2
s1={1,3,5,7} s2={-1,2,3,7,11}
s0=union(s1,s2)    -- s0 is now {-1,1,2,3,5,7,11}.
```

11.5.4.2.4 See Also:

[is_subset](#), [subsets](#), [belongs_to](#)

11.5.4.3 delta

```
include std/sets.e
public function delta(set s1, set s2)
```

Returns the set of elements belonging to either of two sets.

11.5.4.3.1 Parameters:

1. s1 : One of the sets to take a symmetrical difference with
2. s2 : the other set.

11.5.4.3.2 Returns:

The **set**, of all elements belonging to either s1 or s2.

11.5.4.3.3 Example 1:

```
set s0,s1,s2
s1={1,3,5,7} s2={-1,2,3,7,11}
s0=delta(s1,s2)    -- s0 is now {-1,1,2,5,11}.
```

11.5.4.3.4 See Also:

[intersection](#), [union](#), [difference](#)

11.5.4.4 difference

```
include std/sets.e
public function difference(set base, set removed)
```

Returns the set of elements belonging to some set and not to another.

11.5.4.4.1 Parameters:

1. `base` : the set from which a difference is to be taken
2. `removed` : the set of elements to remove from `base`.

11.5.4.4.2 Returns:

The **set**, of elements belonging to `base` but not to `removed`.

11.5.4.4.3 Example 1:

```
set s0,s1,s2
s1={1,3,5,7} s2={-1,2,3,7,11}
s0=difference(s1,s2)    -- s0 is now {1,5}.
```

11.5.4.4.4 See Also:

[remove_from](#), [is_subset](#), [delta](#)

11.5.4.5 product

```
include std/sets.e
public function product(set S1, set S2)
```

Returns the set of all pairs made of an element of a set and an element of another set.

11.5.4.5.1 Parameters:

1. `S1` : The set where the first coordinate lives
2. `S2` : The set where the second coordinate lives

11.5.4.5.2 Returns:

The **set**, of all pairs made of an element of `S1` and an element of `S2`.

11.5.4.5.3 Example 1:

```
set s0,s1,s2
s1 = {1, 3, 5, 7} s2 = {-1, 3}
s0 = product(s1, s2)  -- s0 is now {{1, -1}, {1, 3}, {3, -1}, {3, 3}, {5, -1}, {5, 3}, {7, -1}}
```

11.5.4.5.4 See Also:

[product_map](#), [amalgamated_sum](#), [fiber_product](#)

11.5.5 Maps between sets.

11.5.5.1 define_map

```
include std/sets.e
public function define_map(sequence mapping, set target)
```

Returns a map which sends each element of its source set to the corresponding one in a list.

11.5.5.1.1 Parameters:

1. `mapping` : the sequence mapped to
2. `target` : the target set that contains the elements mapping refers to by index

11.5.5.1.2 Returns:

The requested **map**, descriptor.

11.5.5.1.3 Example 1:

```
sequence s0 = {2, 3, 4, 1, 4, 2}
set s1 = {-1, 1, 2, 3, 4}
map f = define_map(s0,s1)
-- As a sequence, f is {3, 4, 5, 2, 5, 3, 6, 5}
```

11.5.5.1.4 See Also:

[map](#), [sequences_to_map](#), [direct_map](#)

11.5.5.2 sequences_to_map

```
include std/sets.e
public function sequences_to_map(sequence mapped, sequence mapped_to, integer mode)
```

Returns a map which sends each element of some sequence to the corresponding one in another sequence.

11.5.5.2.1 Parameters:

1. mapped : the source sequence
2. mapped_to : the sequence it must map to.
3. mode : an integer, nonzero to also return the minimal sets the result map maps.

11.5.5.2.2 Returns:

A sequence,

- If mode is 0, a map which maps mapped to mapped_to, between the smallest possible sets.
- If mode is not zero, the sequence has length 3. The first element is the map above. The other two elements are the sets derived from the input sequences.

11.5.5.2.3 Comments:

Elements in excess in mapped_to are discarded.

If an element is repeated in mapped, only the mapping of the last occurrence is retained.

11.5.5.2.4 Example 1:

```
sequence s0, s1
s0 = {2, 3, 4, 1, 4}
s1 = {"aba", "aac", 3, "def"}

map f f:=sequences_to_map(s0,s1)
-- As a sequence, f is {3,2,1,4,4,4}
```

11.5.5.2.5 See Also:

[map](#), [define_map](#)

11.5.5.3 image

```
include std/sets.e
public function image(map f, object x, set input, set output)
```

If an object is in some input set, returns how it is mapped to a set.

11.5.5.3.1 Parameters:

1. f : the map to apply
2. x : the object to apply f to
3. `input` : the source set
4. `output` : the target set.

11.5.5.3.2 Returns:

An **object**, $f(x)$ if it can be reckoned.

11.5.5.3.3 Errors:

x must belong to `input` for $f(x)$ to be computed. f must not map to sets larger than `output`; otherwise, it cannot be defined from `input` to `output`.

11.5.5.3.4 Example 1:

```
map f={3,1,2,2,4,3}
set s1,s2
s1={"Albert","Beatrix","Conrad","Doris"} s2={13,17,19}
object x x=image(f,"Conrad",s1,s2)
-- x is now 17.
```

11.5.5.3.5 See Also:

[direct_map](#)

11.5.5.4 range

```
include std/sets.e
public function range(map f, set s)
```

Returns the set of all values taken by a map in some output set.

11.5.5.4.1 Parameters:

1. f : the map to inspect
2. `set` : the output set

11.5.5.4.2 Returns:

The **set**, of all $f(x)$.

11.5.5.4.3 Example 1:

```
map f = {3, 2, 5, 2, 4, 6}
set s = {"Albert", "Beatrix", "Conrad", "Doris", "Eugene", "Fabiola"}
set s1 = range(f, s)
-- s1 is now {"Beatrix", "Conrad", "Eugene"}
```

11.5.5.4.4 See Also:

[direct_map](#), [image](#)

11.5.5.5 direct_map

```
include std/sets.e
public function direct_map(map f, set s1, sequence s0, set s2)
```

Returns the image of a list by a map, given the input and output sets.

11.5.5.5.1 Parameters:

1. f : the map to apply
2. `input` : the source set
3. `elements` : the sequence to map
4. `output` : the target set.

11.5.5.5.2 Returns:

A **sequence**, of elements of `output` obtained by applying f to the corresponding element of `input`.

11.5.5.5.3 Errors:

This function errors out if f cannot map `input` to `output`.

11.5.5.5.4 Comments:

If `elements` has items which are not on `input`, they are ignored. Items may appear in any order any number of times.

11.5.5.5.5 Example:

```
sequence s0 s0={2,3,4,1,4}
set t1,t2
t1={1,2,2.5,3,4} t2={11,13,17,19,23,29}
map f f={3,1,4,5,3,5,5}
sequence s2 s2=direct_map(f,t1,s0,t2)
-- s2 is now {11,29,17,17,17}.
```

11.5.5.5.6 See Also:

[reverse_map](#)

11.5.5.6 restrict

```
include std/sets.e
public function restrict(map f, set source, set restriction)
```

Restricts *f* to the intersection of an input set and another set

11.5.5.6.1 Parameters:

1. *f* : the map to restrict
2. *source* : the initial source set for *f*
3. *restriction* : the set which will help forming a restricted source set.

11.5.5.6.2 Returns:

A **map**, defined on `difference(source, restriction)` which agrees with *f*.

11.5.5.6.3 Example 1:

```
set s1 s1={1,3,5,7,9,11,13,17,19,23}
map f f=[3,7,1,4,5,2,7,1,6,2,10,7]
set s0 s0={3,11,13,19,29}
map f0 f0=restrict(f,s1,s0)
f0 is now: {7,2,7,6,4,7}
```

11.5.5.6.4 See Also:

[is_subset](#), [direct_map](#), [difference](#)

11.5.5.7 change_target

```
include std/sets.e
public function change_target(map f, set old_target, set new_target)
```

Converts a map by changing its output set.

11.5.5.7.1 Parameters:

1. `f` : the map to retarget
2. `old_target` : the initial target set for `f`
3. `new_target` : the new target set.

11.5.5.7.2 Returns:

A **map**, which agrees with `f` and has values in `new_target` instead of `old_target`, or "" if `f` hits something outside `new_target`.

11.5.5.7.3 Example 1:

```
set s1,s2
s1={1,3,5,7,9,11} s2={1,3,7,11,17,19,23}
map f f={2,1,4,6,2,6,6,6}
map f0 f0=change_target(f,s1,s2)
f0 is now: {2,1,3,4,2,4,6,7}
```

11.5.5.7.4 See Also:

[restrict](#), [direct_map](#)

11.5.5.8 combine_maps

```
include std/sets.e
public function combine_maps(map f1, set source1, set target1, map f2, set source2, set target2)
```

Combines two maps into one defined from the union of source sets to the union of target sets.

11.5.5.8.1 Parameters:

1. `f1` : the first map
2. `source1` : its source set
3. `target1` : its target set
4. `f2` : the second map
5. `source2` : its source set

6. target2 : its target set

11.5.5.8.2 Returns:

A **map**, from `union(source1, source2)` to `union(target1, target2)` which agrees with `f1` and `f2`, or "" if `f1` and `f2` disagree at any point of intersection(`s11, s21`).

11.5.5.8.3 Errors:

If `f1` and `f2` are both defined for some point, they must have the same value at this point..

11.5.5.8.4 Example 1:

```
set s11,s12,s21,s22
s11={2,3,5,7,11,13,17,19} s21={7,13,19,23,29}
s12={-1,0,1,4} s22={-2,0,1,2,6}
map f1,f2
f1={2,1,3,3,2,3,1,2,8,4} f2={3,3,2,4,5,5,5}
map f f=combine_maps(f1,s11,s12,f2,s21,s22)
-- f is now: {3,2,4,4,3,4,2,3,5,7,10,7}.
```

11.5.5.8.5 See Also:

[restrict](#), [direct_map](#)

11.5.5.9 compose_map

```
include std/sets.e
public function compose_map(map f1, map f2)
```

Creates a new map using elements from `f2`, mapped against `f1`

11.5.5.9.1 Parameters:

1. `f1` : the map containing indexes into `f2`
2. `f2` : the map containing elements used to build the resulting map.

11.5.5.9.2 Returns:

A **map**, `f` defined by $f(x) = f2(f1(x))$ for all `x`

11.5.5.9.3 Comments:

Each element in `f1` is an index into the elements of `f2`. So if `f1` contains `{3,2,1}` the result map contains the 3rd, 2nd and 1st element from `f2` in that order.

11.5.5.9.4 Errors:

Every element of `f1` must be a valid index into `f2`.

11.5.5.9.5 Example 1:

```
map f1,f2,f
f1={2,3,1,1,2,5,3}
f2={4,8,1,2,6,7,6,9}
f=compose_map(f1,f2)
-- f is now: {8,1,4,4,8,5,9}
```

11.5.5.9.6 See Also:

[diagram_commutates](#)

11.5.5.10 diagram_commutates

```
include std/sets.e
public function diagram_commutates(sequence f12a, sequence f12b, sequence f2a3, sequence f2b3)
```

Decide whether taking two different paths along a square map diagrams results in the same map.

11.5.5.10.1 Parameters:

1. `from_base_path_1` : the outgoing map along path 1
2. `from_base_path_2` : the outgoing map along path 2
3. `to_target_path_1` : the incoming map along path 1
4. `to_target_path_2` : the incoming map along path 2

11.5.5.10.2 Returns:

An **integer**, either 1 if `to_target_path_1` o `from_base_path_1` = `to_target_path_2` o `from_base_path_2`.

11.5.5.10.3 Example 1:

```
map f12a,f12b,f2a3,f2b3
f12a={2,3,1,1,2,5,3}
f2a3={4,8,1,2,6,7,6,9}
```

```
f12b={2,4,2,3,1,5,4}
f2b3={8,8,4,1,3,5,8}
?diagram_commutates(f12a,f12b,f2a3,f2b3)  -- prints out 0
```

11.5.5.10.4 See Also:

[compose_map](#)

11.5.5.11 is_injective

```
include std/sets.e
public function is_injective(map f)
```

Determines whether there is a point in an output set hit twice or more by a map.

11.5.5.11.1 Parameters:

1. f : the map being queried.

11.5.5.11.2 Returns:

An **integer**, 0 if f ever maps two points to the same element, else 1.

11.5.5.11.3 Example 1:

```
map f f={2,3,1,1,2,5,3}
?is_injective(f)  -- prints out 0
```

11.5.5.11.4 See Also:

[is_surjective](#), [is_bijective](#), [reverse_map](#), [fiber_over](#)

11.5.5.12 is_surjective

```
include std/sets.e
public function is_surjective(map f)
```

Determine whether all points in the output set are hit by a map.

11.5.5.12.1 Parameters:

1. f : the map to test.

11.5.5.12.2 Returns:

An **integer**, 0 if f ever misses some point in the target set, else 1.

11.5.5.12.3 Example 1:

```
map f f={2,3,1,1,2,5,3}
?is_surjective(f)  -- prints out 1
```

11.5.5.12.4 See Also:

[is_surjective](#), [is_bijective](#), [direct_map](#), [section](#)

11.5.5.13 is_bijective

```
include std/sets.e
public function is_bijective(map f)
```

Determine whether a map is one-to-one.

11.5.5.13.1 Parameters:

1. f : the map to test.

11.5.5.13.2 Returns:

An **integer**, 1 if f is one-to-one, else 0.

11.5.5.13.3 Example 1:

```
map f f={2,3,1,1,2,5,3}
? is_surjective(f)  -- prints out 0
```

11.5.5.13.4 See Also:

[is_surjective](#), [is_bijective](#), [direct_map](#), [has_inverse](#)

11.5.6 Reverse mappings

11.5.6.1 fiber_over

```
include std/sets.e
public function fiber_over(map f, set source, set target)
```

Given a map between two sets, returns {list of antecedents of elements in target, effective target}.

11.5.6.1.1 Parameters:

1. f : the inspected map
2. `source` : the source set
3. `target` : the target set.

11.5.6.1.2 Returns:

A **sequence**, which is empty on failure. On success, it has two elements:

- A sequence of sets; each of these sets is included in `source` and is mapped to a single point by f .
- A set, the points in `target` hit by f .

11.5.6.1.3 Comments:

The listed sets, which are reverse images of points in `target`, are called *fibers* of f over points, specially if they are isomorphic to one another for some extra algebraic or topological structure.

The fibers are enumerated in the same order as the points in the effective target, i.e. the points in `target` f hits.

11.5.6.1.4 Example 1:

```
set s1,s2
s1={5,7,9,11} s2={13,17,19,23,29}
map f f={2,1,4,1,4,5}
sequence s s=fiber_over(f,s1,s2)
-- s is now {{7,11},{5},{9}},{13,17,23}}.
```

11.5.6.1.5 See Also:

[reverse_map](#), [fiber_product](#)

11.5.6.2 reverse_map

```
include std/sets.e
public function reverse_map(map f, set s1, sequence s0, set s2)
```

Given a map between two sets, returns the smallest subset whose image contains the set of elements in a list.

11.5.6.2.1 Parameters:

1. `f` : the map relative to which reverse images are to be taken
2. `source` : the source set
3. `elements` : the list of elements in target to lift to source
4. `target` : the target set

11.5.6.2.2 Returns:

A **set**, which is included in `source` and contains all antecedents of elements in `elements` by `f`.

11.5.6.2.3 Comments:

Elements which `f` does not hit are ignored.

11.5.6.2.4 Example 1:

```
set s1,s2
s1={5,7,9,11} s2={13,17,19,23,29}
sequence s0 s0={23,13,17,23}
map f f={5,3,1,3,4,5}
set s s=reverse_map(f,s1,s0,s2)
s is now {9}.
```

11.5.6.2.5 See Also:

[direct_map](#), [fiber_over](#)

11.5.6.3 section

```
include std/sets.e
public function section(map f)
```

Return a right, and left is possible, inverse of a map over its [range](#).

11.5.6.3.1 Parameters:

1. f : the map to invert.

11.5.6.3.2 Returns:

A **map**, g such that $f(g(y)) = y$ whenever y is hit by f . and If f is injective, it also holds that $g(f(x)) = x$.

11.5.6.3.3 Example 1:

```
map f = {2, 3, 1, 1, 2, 5, 3}, g = section(f)
-- g is now {3,1,2,3,5}.
```

11.5.6.3.4 See Also:

[reverse_map](#), [is_injective](#)

11.5.7 Products

11.5.7.1 product_map

```
include std/sets.e
public function product_map(map f1, map f2)
```

Builds a map to a product from a map to each of its components.

11.5.7.1.1 Parameters:

1. $f1$: the map going to the first component
2. $f2$: the map going to the second component

11.5.7.1.2 Returns:

A **map**, $f = f1 \times f2$ defined by $f(x, y) = \{f1(x), f2(y)\}$ wherever this makes sense.

11.5.7.1.3 Example 1:

```
set s s={1,3,5,7}
map f f={3,1,4,1,4,4}
map f1 f1=product(f,f)
-- f1 is {11,9,12,9,3,1,4,1,15,13,16,13,3,1,4,1,16,16}.
```

11.5.7.1.4 See Also:

[product](#), [amalgamated_sum](#), [fiber_product](#)

11.5.7.2 amalgamated_sum

```
include std/sets.e
public function amalgamated_sum(set first, set second, set base, map base_to_1, map base_to_2)
```

Returns all pairs in a product that come from applying two maps to the same element in a base set.

11.5.7.2.1 Parameters:

1. `first` : one of the sets to involved in the sum
2. `second` : the other set
3. `base` : the base set
4. `base_to_1` : the map from base to first
5. `base_to_2` : the map from base to second

11.5.7.2.2 Returns:

A **set**, of pairs obtained by applying `f01Xf02` to `s0`.

11.5.7.2.3 Example 1:

```
set s0,s1,s2
s0={1,2,3} s1={5,7,9,11} s2={13,17,19}
map f01,f02
f01={2,4,1,3,4} f02={2,2,1,3,3}
set s s=amalgamated_product(s1,s2,s0,f01,f02)
-- s is now {{7,17},{11,17},{5,13}}.
```

11.5.7.2.4 See Also:

[product](#), [product_map](#), [fiber_product](#)

11.5.7.3 fiber_product

```
include std/sets.e
public function fiber_product(set first, set second, set base, map from_1_to_base, map from_2_to_base)
```

Returns the set of all pairs in a product on which two given componentwise maps agree.

11.5.7.3.1 Parameters:

1. `first` : the first product component
2. `second` : the second product component
3. `base` : the base set the fiber product is built on
4. `from_1_to_base` : the map from `first` to `base`.
5. `from_2_to_base` : the map from `second` to `base`.

11.5.7.3.2 Returns:

The **set**, of pairs whose coordinates are mapped consistently to `base` by `from_1_to_base` and `from_2_to_base` respectively.

11.5.7.3.3 Example 1:

```
set s0,s1,s2
s0={1,2,3} s1={5,7,9,11} s2={13,17,19,23,29}
map f10,f20
f10={2,1,2,1,4,3} f20={1,3,3,2,3,5,3}
set s=s_fiber_product(s1,s2,s0,f10,f20)
-- s is now {{5,23},{7,13},{9,23},{11,13}}.
```

11.5.7.3.4 See Also:

[reverse_map](#), [{ :amalgamated_sum }](#), [fiber_over](#)

11.5.8 Constants

11.5.8.1 SIDE_NONE

```
include std/sets.e
public enum SIDE_NONE
```

The following constants denote orientation of distributivity or unitarity:

- `SIDE_NONE` -- no units, or no distributivity
- `SIDE_LEFT` -- property is requested or verified on the left side
- `SIDE_RIGHT` -- property is requested or verified on the right side
- `SIDE_BOTH` -- property is requested or verified on both sides.

11.5.9 Operations on sets

11.5.9.1 define_operation

```
include std/sets.e
public function define_operation(sequence left_actions)
```

Returns an operation that splits by left action into the supplied mappings.

11.5.9.1.1 Parameters:

1. `left_actions` : a sequence of maps, the left actions of each element in the left hand set.

11.5.9.1.2 Returns:

An operation F , realizing the conditions above, with minimal cardinal values, or "" if the maps are not defined on the same set.

11.5.9.1.3 Errors:

`left_actions` must be a rectangular matrix.

11.5.9.1.4 Comments:

If F is the result, and is defined from $E1 \times E2$ to E , then each left action is a map from $E2$ to E , the "left multiplication" by an element of $E1$.

11.5.9.1.5 Example 1:

```
sequence s = {{2, 3, 2, 3}, {3, 1, 2, 5}, {1, 2, 2, 2}, {2, 3, 2, 4}, {3, 1, 2, 3}}
operation F = define_operation(s)
-- F is now {{2,3},{3,1},{1,2},{2,3},{3,1}},{5,2,3}
? operation(s)  -- prints out 1.
```

11.5.9.1.6 See Also:

[operation](#)

11.5.9.2 is_symmetric

```
include std/sets.e
public function is_symmetric(operation f)
```

Determine whether $f(x,y)$ always equals $f(y,x)$.

11.5.9.2.1 Parameters:

1. f : the operation to test.

11.5.9.2.2 Returns:

An **integer**, 1 if exchanging operands makes sense and has no effect, else 0.

11.5.9.2.3 Example 1:

```
operation f f={{1,2,3},{2,3,4},{3,4,5}},{3,3,5}}
-- f is the addition from {0,1,2}x{0,1,2} to {0,1,2,3,4}.
? is_symmetric(f)    -- prints out 1.
```

11.5.9.2.4 See Also:

[operation](#), [has_unit](#)

11.5.9.3 is_associative

```
include std/sets.e
public function is_associative(operation f)
```

Determine whether the identity $f(f(x,y),z)=f(x,f(y,z))$ always makes sense and holds.

11.5.9.3.1 Parameters:

1. f : the operation to test.

11.5.9.3.2 Returns:

An **integer**, 1 if f is an internal operation on a set and is associative, else 0.

11.5.9.3.3 Comments:

Being associative is equivalent to not depending on parentheses for defining iterated execution.

11.5.9.3.4 Example 1:

```
operation f f={{1, 2, 3}, {2, 3, 1}, {3, 1, 2}}, {3, 3, 3}}
-- f is the addition modulo 3 from {0, 1, 2} x {0, 1, 2} to {0, 1, 2}.
? is_symmetric(f)    -- prints out 1.
```

11.5.9.3.5 See Also:

[operation](#), [has_unit](#)

11.5.9.4 all_left_units

```
include std/sets.e
public function all_left_units(operation f)
```

Finds all left units for an operation.

11.5.9.4.1 Parameters:

1. f : the operation to test.

11.5.9.4.2 Returns:

A possibly empty **sequence**, listing all x such that $f(x, .)$ is the identity map.

11.5.9.4.3 Example 1:

```
operation f f={{1,2,3}, {1,2,3}, {3,1,2}}, {3,3,3}}
sequence s s=all_left_units(f)
s is now {1,2}.
```

11.5.9.4.4 See Also:

[all_right_units](#), [is_unit](#), [has_unit](#)

11.5.9.5 all_right_units

```
include std/sets.e
public function all_right_units(operation f)
```

Finds all right units for an operation.

11.5.9.5.1 Parameters:

1. f : the operation to test.

11.5.9.5.2 Returns:

A possibly empty **sequence**, of all y such that $f(., y)$ is the identity map..

11.5.9.5.3 Example 1:

```
operation f f={{1,2,3},{1,2,3},{3,1,2}},{3,3,3}}
sequence s s=all_right_units(f)
s is now empty.
```

11.5.9.5.4 See Also:

[all_left_units](#), [is_unit](#), [has_unit](#)

11.5.9.6 has_unit

```
include std/sets.e
public function has_unit(operation f, integer flags = SIDE_BOTH)
```

Returns an unit of a given kind for an operation if there is any, else 0.

11.5.9.6.1 Parameters:

1. f : the operation to test.
2. $flags$: an integer, which says whether one or two sided units are looked for. Defaults to `SIDE_BOTH`.

11.5.9.6.2 Returns:

An **integer**, if f has a unit of the requested type, it is returned. Otherwise, 0 is returned..

11.5.9.6.3 Comments:

If there is a two sided inverse, it is unique.

Only the two lower bits of `flags` matter. They must be `SIDE_LEFT` to check for left units, `SIDE_RIGHT` for right units. Otherwise, two sided units are determined.

11.5.9.6.4 Example 1:

```
operation f f={{1,2,3},{2,3,1},{3,1,2}},{3,3,3}}
? has_unit(f)  -- prints out 1.
```

11.5.9.6.5 See Also:

[all_left_units](#), [all_right_units](#), [is_unit](#)

11.5.9.7 is_unit

```
include std/sets.e
public function is_unit(integer x, operation f)
```

Determines if an element is a (one sided) unit for an operation.

11.5.9.7.1 Parameters:

1. `x` : an integer, the element to test
2. `f` : the operation involved

11.5.9.7.2 Returns:

An **integer**, either `SIDE_NONE`, `SIDE_LEFT`, `SIDE_RIGHT` or `SIDE_BOTH`.

11.5.9.7.3 Example 1:

```
operation f f={{1, 2, 3}, {1, 2, 3}, {3, 1, 2}}, {3, 3,3 }}
? is_left_unit(3, f)  -- prints out 0.
```

11.5.9.7.4 See Also:

[all_left_units](#), [has_unit](#)

11.5.9.8 has_inverse

```
include std/sets.e
public function has_inverse(integer x, operation f)
```

Returns the bilateral inverse of an element by a operation if it exists and the operation has a unit.

11.5.9.8.1 Parameters:

1. x : the element to test
2. f : the operation involved.

11.5.9.8.2 Returns:

If f , has a bilateral unit e and there is a (necessarily unique) y such that $f(x, y) = e$, y is returned. Otherwise, 0 is returned..

11.5.9.8.3 Example 1:

```
operation f f={{1, 2, 3}, {2, 3, 1}, {3, 1, 2}}, {3, 3, 3}}
? has_inverse(3, f) -- prints out 2.
```

11.5.9.8.4 See Also:

[has_unit](#)

11.5.9.9 distributes_over

```
include std/sets.e
public function distributes_over(operation product, operation sum, integer transpose = 0)
```

Determine whether a product map distributes over a sum

11.5.9.9.1 Parameters:

1. $product$: the operation that may be distributive over sum
2. sum : the operations over which $product$ might distribute
3. $transpose$: an integer, nonzero if $product$ is a right operation. Defaults to 0.

11.5.9.9.2 Returns:

An **integer**, either of

- SIDE_NONE -- product does not distribute either way over sum
- SIDE_LEFT -- product distributes over sum on the left only
- SIDE_RIGHT -- product distributes over sum on the right only
- SIDE_BOTH -- product distributes over sum o(both ways)

11.5.9.3 Example 1:

```
operation sum          sum=    {{ {1,2,3}, {2,3,1}, {3,1,2} }, {3,3,3} }
operation product      product={{ {1,1,1}, {1,2,3}, {1,3,2} }, {3,3,3} }
?distributes_right(product,sum,0)  -- prints out 1.
```

12 Networking Routines

Core Sockets

- Error Information
- Socket Type Constants
- Select Accessor Constants
- Shutdown Options
- Socket Options
- Send Flags
- Server and Client sides
- Client side only
- Server side only
- UDP only
- Information

Common Internet Routines

- IP Address Handling
- URL Parsing

DNS

- Constants
- General Routines

HTTP

- Constants
- Header management
- Web interface

URL handling

- Parsing
- URL encoding and decoding

12.1 Core Sockets

Error Information

- error_code
- OK
- ERR_ACCESS
- ERR_ADDRINUSE
- ERR_ADDRNOTAVAIL
- ERR_AFNOSUPPORT
- ERR_AGAIN
- ERR_ALREADY
- ERR_CONNABORTED
- ERR_CONNREFUSED
- ERR_CONNRESET
- ERR_DESTADDRREQ
- ERR_FAULT
- ERR_HOSTUNREACH

ERR_INPROGRESS
ERR_INTR
ERR_INVAL
ERR_IO
ERR_ISCONN
ERR_ISDIR
ERR_LOOP
ERR_MFILE
ERR_MSGSIZE
ERR_NAMETOOLONG
ERR_NETDOWN
ERR_NETRESET
ERR_NETUNREACH
ERR_NFILE
ERR_NOBUFS
ERR_NOENT
ERR_NOTCONN
ERR_NOTDIR
ERR_NOTINITIALISED
ERR_NOTSOCK
ERR_OPNOTSUPP
ERR_PROTONOSUPPORT
ERR_PROTOTYPE
ERR_ROFS
ERR_SHUTDOWN
ERR_SOCKETNOSUPPORT
ERR_TIMEDOUT
ERR_WOULDBLOCK

Socket Type Constants

AF_UNSPEC
AF_UNIX
AF_INET
AF_INET6
AF_APPLETALK
AF_BTH
SOCK_STREAM
SOCK_DGRAM
SOCK_RAW
SOCK_RDM
SOCK_SEQPACKET

Select Accessor Constants

SELECT_SOCKET
SELECT_IS_READABLE
SELECT_IS_WRITABLE
SELECT_IS_ERROR

Shutdown Options

SD_SEND
SD_RECEIVE
SD_BOTH

Socket Options

SOL_SOCKET
SO_DEBUG
SO_ACCEPTCONN
SO_REUSEADDR
SO_KEEPAIVE
SO_DONTROUTE
SO_BROADCAST
SO_USELOOPBACK
SO_LINGER
SO_DONTLINGER
SO_OOBINLINE
SO_REUSEPORT
SO_SNDBUF
SO_RCVBUF
SO_SNDLOWAT
SO_RCVLOWAT
SO_SNDTIMEO
SO_RCVTIMEO
SO_ERROR
SO_TYPE
SO_CONNDATA
SO_CONNOPT
SO_DISCDATA
SO_DISCOPT
SO_CONNDATALEN
SO_CONNOPTLEN
SO_DISCDATALEN
SO_DISCOPTLEN
SO_OPENTYPE
SO_MAXDG
SO_MAXPATHDG
SO_SYNCHRONOUS_ALERT
SO_SYNCHRONOUS_NONALERT

Send Flags

MSG_OOB
MSG_PEEK
MSG_DONTROUTE
MSG_TRYHARD
MSG_CTRUNC
MSG_PROXY
MSG_TRUNC
MSG_DONTWAIT
MSG_EOR
MSG_WAITALL
MSG_FIN
MSG_SYN
MSG_CONFIRM
MSG_RST



- MSG_ERRQUEUE
- MSG_NOSIGNAL
- MSG_MORE
- Server and Client sides
 - SOCKET_SOCKET
 - SOCKET_SOCKADDR_IN
 - socket
 - create
 - close
 - shutdown
 - select
 - send
 - receive
 - get_option
 - set_option
- Client side only
 - connect
- Server side only
 - bind
 - listen
 - accept
- UDP only
 - send_to
 - receive_from
- Information
 - service_by_name
 - service_by_port

12.1.1 Error Information

12.1.1.1 error_code

```
include std/socket.e
public function error_code()
```

Get the error code.

12.1.1.1.1 Returns:

Integer **OK** on no error, otherwise any one of the `ERR_` constants to follow.

12.1.1.2 OK

```
include std/socket.e
public constant OK
```

No error occurred.

12.1.1.3 ERR_ACCESS

```
include std/socket.e
public constant ERR_ACCESS
```

Permission has been denied. This can happen when using a `send_to` call on a broadcast address without setting the socket option `SO_BROADCAST`. Another, possibly more common, reason is you have tried to bind an address that is already exclusively bound by another application.

May occur on a Unix Domain Socket when the socket directory or file could not be accessed due to security.

12.1.1.4 ERR_ADDRINUSE

```
include std/socket.e
public constant ERR_ADDRINUSE
```

Address is already in use.

12.1.1.5 ERR_ADDRNOTAVAIL

```
include std/socket.e
public constant ERR_ADDRNOTAVAIL
```

The specified address is not a valid local IP address on this computer.

12.1.1.6 ERR_AFNOSUPPORT

```
include std/socket.e
public constant ERR_AFNOSUPPORT
```

Address family not supported by the protocol family.

12.1.1.7 ERR_AGAIN

```
include std/socket.e
public constant ERR_AGAIN
```

Kernel resources to complete the request are temporarily unavailable.

12.1.1.8 ERR_ALREADY

```
include std/socket.e
public constant ERR_ALREADY
```

Operation is already in progress.

12.1.1.9 ERR_CONNABORTED

```
include std/socket.e
public constant ERR_CONNABORTED
```

Software has caused a connection to be aborted.

12.1.1.10 ERR_CONNREFUSED

```
include std/socket.e
public constant ERR_CONNREFUSED
```

Connection was refused.

12.1.1.11 ERR_CONNRESET

```
include std/socket.e
public constant ERR_CONNRESET
```

An incoming connection was supplied however it was terminated by the remote peer.

12.1.1.12 ERR_DESTADDRREQ

```
include std/socket.e
public constant ERR_DESTADDRREQ
```

Destination address required.

12.1.1.13 ERR_FAULT

```
include std/socket.e
public constant ERR_FAULT
```

Address creation has failed internally.

12.1.1.14 ERR_HOSTUNREACH

```
include std/socket.e
public constant ERR_HOSTUNREACH
```

No route to the host specified could be found.

12.1.1.15 ERR_INPROGRESS

```
include std/socket.e
public constant ERR_INPROGRESS
```

A blocking call is inprogress.

12.1.1.16 ERR_INTR

```
include std/socket.e
public constant ERR_INTR
```

A blocking call was cancelled or interrupted.

12.1.1.17 ERR_INVALID

```
include std/socket.e
public constant ERR_INVALID
```

An invalid sequence of command calls were made, for instance trying to accept before an actual listen was called.

12.1.1.18 ERR_IO

```
include std/socket.e
public constant ERR_IO
```

An I/O error occurred while making the directory entry or allocating the inode. (Unix Domain Socket).

12.1.1.19 ERR_ISCONN

```
include std/socket.e  
public constant ERR_ISCONN
```

Socket is already connected.

12.1.1.20 ERR_ISDIR

```
include std/socket.e  
public constant ERR_ISDIR
```

An empty pathname was specified. (Unix Domain Socket).

12.1.1.21 ERR_LOOP

```
include std/socket.e  
public constant ERR_LOOP
```

Too many symbolic links were encountered. (Unix Domain Socket).

12.1.1.22 ERR_MFILE

```
include std/socket.e  
public constant ERR_MFILE
```

The queue is not empty upon routine call.

12.1.1.23 ERR_MSGSIZE

```
include std/socket.e  
public constant ERR_MSGSIZE
```

Message is too long for buffer size. This would indicate an internal error to EUPHORIA as EUPHORIA sets a dynamic buffer size.

12.1.1.24 ERR_NAMETOOLONG

```
include std/socket.e  
public constant ERR_NAMETOOLONG
```

Component of the path name exceeded 255 characters or the entire path exceeded 1023 characters. (Unix Domain Socket).

12.1.1.25 ERR_NETDOWN

```
include std/socket.e  
public constant ERR_NETDOWN
```

The network subsystem is down or has failed

12.1.1.26 ERR_NETRESET

```
include std/socket.e  
public constant ERR_NETRESET
```

Network has dropped it's connection on reset.

12.1.1.27 ERR_NETUNREACH

```
include std/socket.e  
public constant ERR_NETUNREACH
```

Network is unreachable.

12.1.1.28 ERR_NFILE

```
include std/socket.e  
public constant ERR_NFILE
```

Not a file. (Unix Domain Sockets).

12.1.1.29 ERR_NOBUFS

```
include std/socket.e  
public constant ERR_NOBUFS
```

No buffer space is available.

12.1.1.30 ERR_NOENT

```
include std/socket.e  
public constant ERR_NOENT
```

Named socket does not exist. (Unix Domain Socket).

12.1.1.31 ERR_NOTCONN

```
include std/socket.e  
public constant ERR_NOTCONN
```

Socket is not connected.

12.1.1.32 ERR_NOTDIR

```
include std/socket.e  
public constant ERR_NOTDIR
```

Component of the path prefix is not a directory. (Unix Domain Socket).

12.1.1.33 ERR_NOTINITIALISED

```
include std/socket.e  
public constant ERR_NOTINITIALISED
```

Socket system is not initialized (Windows only)

12.1.1.34 ERR_NOTSOCK

```
include std/socket.e  
public constant ERR_NOTSOCK
```

The descriptor is not a socket.

12.1.1.35 ERR_OPNOTSUPP

```
include std/socket.e  
public constant ERR_OPNOTSUPP
```

Operation is not supported on this type of socket.

12.1.1.36 ERR_PROTONOSUPPORT

```
include std/socket.e  
public constant ERR_PROTONOSUPPORT
```

Protocol not supported.

12.1.1.37 ERR_PROTOTYPE

```
include std/socket.e  
public constant ERR_PROTOTYPE
```

Protocol is the wrong type for the socket.

12.1.1.38 ERR_ROFS

```
include std/socket.e  
public constant ERR_ROFS
```

The name would reside on a read-only file system. (Unix Domain Socket).

12.1.1.39 ERR_SHUTDOWN

```
include std/socket.e  
public constant ERR_SHUTDOWN
```

The socket has been shutdown. Possibly a send/receive call after a shutdown took place.

12.1.1.40 ERR_SOCKETNOSUPPORT

```
include std/socket.e  
public constant ERR_SOCKETNOSUPPORT
```

Socket type is not supported.

12.1.1.41 ERR_TIMEDOUT

```
include std/socket.e
public constant ERR_TIMEDOUT
```

Connection has timed out.

12.1.1.42 ERR_WOULDBLOCK

```
include std/socket.e
public constant ERR_WOULDBLOCK
```

The operation would block on a socket marked as non-blocking.

12.1.2 Socket Type Constants

These values are passed as the `family` and `sock_type` parameters of the [create](#) function.

12.1.2.1 AF_UNSPEC

```
include std/socket.e
public constant AF_UNSPEC
```

Address family is unspecified

12.1.2.2 AF_UNIX

```
include std/socket.e
public constant AF_UNIX
```

Local communications

12.1.2.3 AF_INET

```
include std/socket.e
public constant AF_INET
```

IPv4 Internet protocols

12.1.2.4 AF_INET6

```
include std/socket.e
public constant AF_INET6
```

IPv6 Internet protocols

12.1.2.5 AF_APPLETALK

```
include std/socket.e
public constant AF_APPLETALK
```

Appletalk

12.1.2.6 AF_BTH

```
include std/socket.e
public constant AF_BTH
```

Bluetooth

12.1.2.7 SOCK_STREAM

```
include std/socket.e
public constant SOCK_STREAM
```

Provides sequenced, reliable, two-way, connection-based byte streams. An out-of-band data transmission mechanism may be supported.

12.1.2.8 SOCK_DGRAM

```
include std/socket.e
public constant SOCK_DGRAM
```

Supports datagrams (connectionless, unreliable messages of a fixed maximum length).

12.1.2.9 SOCK_RAW

```
include std/socket.e
public constant SOCK_RAW
```

Provides raw network protocol access.

12.1.2.10 SOCK_RDM

```
include std/socket.e
public constant SOCK_RDM
```

Provides a reliable datagram layer that does not guarantee ordering.

12.1.2.11 SOCK_SEQPACKET

```
include std/socket.e
public constant SOCK_SEQPACKET
```

Obsolete and should not be used in new programs

12.1.3 Select Accessor Constants

Use with the result of [select](#).

12.1.3.1 SELECT_SOCKET

```
include std/socket.e
public enum SELECT_SOCKET
```

The socket

12.1.3.2 SELECT_IS_READABLE

```
include std/socket.e
public enum SELECT_IS_READABLE
```

Boolean (1/0) value indicating the readability.

12.1.3.3 SELECT_IS_WRITABLE

```
include std/socket.e
public enum SELECT_IS_WRITABLE
```

Boolean (1/0) value indicating the writeability.

12.1.3.4 SELECT_IS_ERROR

```
include std/socket.e
public enum SELECT_IS_ERROR
```

Boolean (1/0) value indicating the error state.

12.1.4 Shutdown Options

Pass one of the following to the `method` parameter of [shutdown](#).

12.1.4.1 SD_SEND

```
include std/socket.e
public constant SD_SEND
```

Shutdown the send operations.

12.1.4.2 SD_RECEIVE

```
include std/socket.e
public constant SD_RECEIVE
```

Shutdown the receive operations.

12.1.4.3 SD_BOTH

```
include std/socket.e
public constant SD_BOTH
```

Shutdown both send and receive operations.

12.1.5 Socket Options

Pass to the `optname` parameter of the functions [get_option](#) and [set_option](#).

These options are highly OS specific and are normally not needed for most socket communication. They are provided here for your convenience. If you should need to set socket options, please refer to your OS

reference material.

There may be other values that your OS defines and some defined here are not supported on all operating systems.

12.1.5.1 SOL_SOCKET

```
include std/socket.e
public constant SOL_SOCKET
```

12.1.5.2 SO_DEBUG

```
include std/socket.e
public constant SO_DEBUG
```

12.1.5.3 SO_ACCEPTCONN

```
include std/socket.e
public constant SO_ACCEPTCONN
```

12.1.5.4 SO_REUSEADDR

```
include std/socket.e
public constant SO_REUSEADDR
```

12.1.5.5 SO_KEEPALIVE

```
include std/socket.e
public constant SO_KEEPALIVE
```

12.1.5.6 SO_DONTROUTE

```
include std/socket.e
public constant SO_DONTROUTE
```

12.1.5.7 SO_BROADCAST

```
include std/socket.e
public constant SO_BROADCAST
```

12.1.5.8 SO_USELOOPBACK

```
include std/socket.e
public constant SO_USELOOPBACK
```

12.1.5.9 SO_LINGER

```
include std/socket.e
public constant SO_LINGER
```

12.1.5.10 SO_DONTLINGER

```
include std/socket.e
public constant SO_DONTLINGER
```

12.1.5.11 SO_OOBINLINE

```
include std/socket.e
public constant SO_OOBINLINE
```

12.1.5.12 SO_REUSEPORT

```
include std/socket.e
public constant SO_REUSEPORT
```

12.1.5.13 SO_SNDBUF

```
include std/socket.e
public constant SO_SNDBUF
```



12.1.5.14 SO_RCVBUF

```
include std/socket.e  
public constant SO_RCVBUF
```

12.1.5.15 SO_SNDLOWAT

```
include std/socket.e  
public constant SO_SNDLOWAT
```

12.1.5.16 SO_RCVLOWAT

```
include std/socket.e  
public constant SO_RCVLOWAT
```

12.1.5.17 SO_SNDTIMEO

```
include std/socket.e  
public constant SO_SNDTIMEO
```

12.1.5.18 SO_RCVTIMEO

```
include std/socket.e  
public constant SO_RCVTIMEO
```

12.1.5.19 SO_ERROR

```
include std/socket.e  
public constant SO_ERROR
```

12.1.5.20 SO_TYPE

```
include std/socket.e  
public constant SO_TYPE
```

12.1.5.21 SO_CONNDATA

```
include std/socket.e
public constant SO_CONNDATA
```

12.1.5.22 SO_CONNOPT

```
include std/socket.e
public constant SO_CONNOPT
```

12.1.5.23 SO_DISCDATA

```
include std/socket.e
public constant SO_DISCDATA
```

12.1.5.24 SO_DISCOPT

```
include std/socket.e
public constant SO_DISCOPT
```

12.1.5.25 SO_CONNDATALEN

```
include std/socket.e
public constant SO_CONNDATALEN
```

12.1.5.26 SO_CONNOPTLEN

```
include std/socket.e
public constant SO_CONNOPTLEN
```

12.1.5.27 SO_DISCDATALEN

```
include std/socket.e
public constant SO_DISCDATALEN
```


12.1.5.28 SO_DISCOPTLEN

```
include std/socket.e
public constant SO_DISCOPTLEN
```

12.1.5.29 SO_OPENTYPE

```
include std/socket.e
public constant SO_OPENTYPE
```

12.1.5.30 SO_MAXDG

```
include std/socket.e
public constant SO_MAXDG
```

12.1.5.31 SO_MAXPATHDG

```
include std/socket.e
public constant SO_MAXPATHDG
```

12.1.5.32 SO_SYNCHRONOUS_ALERT

```
include std/socket.e
public constant SO_SYNCHRONOUS_ALERT
```

12.1.5.33 SO_SYNCHRONOUS_NONALERT

```
include std/socket.e
public constant SO_SYNCHRONOUS_NONALERT
```

12.1.6 Send Flags

Pass to the `flags` parameter of [send](#) and [receive](#)

12.1.6.1 MSG_OOB

```
include std/socket.e
public constant MSG_OOB
```

Sends out-of-band data on sockets that support this notion (e.g., of type [SOCK_STREAM](#)); the underlying protocol must also support out-of-band data.

12.1.6.2 MSG_PEEK

```
include std/socket.e
public constant MSG_PEEK
```

This flag causes the receive operation to return data from the beginning of the receive queue without removing that data from the queue. Thus, a subsequent receive call will return the same data.

12.1.6.3 MSG_DONTROUTE

```
include std/socket.e
public constant MSG_DONTROUTE
```

Don't use a gateway to send out the packet, only send to hosts on directly connected networks. This is usually used only by diagnostic or routing programs. This is only defined for protocol families that route; packet sockets don't.

12.1.6.4 MSG_TRYHARD

```
include std/socket.e
public constant MSG_TRYHARD
```

12.1.6.5 MSG_CTRUNC

```
include std/socket.e
public constant MSG_CTRUNC
```

indicates that some control data were discarded due to lack of space in the buffer for ancillary data.

12.1.6.6 MSG_PROXY

```
include std/socket.e
public constant MSG_PROXY
```

12.1.6.7 MSG_TRUNC

```
include std/socket.e
public constant MSG_TRUNC
```

indicates that the trailing portion of a datagram was discarded because the datagram was larger than the buffer supplied.

12.1.6.8 MSG_DONTWAIT

```
include std/socket.e
public constant MSG_DONTWAIT
```

Enables non-blocking operation; if the operation would block, EAGAIN or EWOULDBLOCK is returned.

12.1.6.9 MSG_EOR

```
include std/socket.e
public constant MSG_EOR
```

Terminates a record (when this notion is supported, as for sockets of type [SOCK_SEQPACKET](#)).

12.1.6.10 MSG_WAITALL

```
include std/socket.e
public constant MSG_WAITALL
```

This flag requests that the operation block until the full request is satisfied. However, the call may still return less data than requested if a signal is caught, an error or disconnect occurs, or the next data to be received is of a different type than that returned.

12.1.6.11 MSG_FIN

```
include std/socket.e
public constant MSG_FIN
```

12.1.6.12 MSG_SYN

```
include std/socket.e
public constant MSG_SYN
```

12.1.6.13 MSG_CONFIRM

```
include std/socket.e
public constant MSG_CONFIRM
```

Tell the link layer that forward progress happened: you got a successful reply from the other side. If the link layer doesn't get this it will regularly reprobe the neighbor (e.g., via a unicast ARP). Only valid on [SOCK_DGRAM](#) and [SOCK_RAW](#) sockets and currently only implemented for IPv4 and IPv6.

12.1.6.14 MSG_RST

```
include std/socket.e
public constant MSG_RST
```

12.1.6.15 MSG_ERRQUEUE

```
include std/socket.e
public constant MSG_ERRQUEUE
```

indicates that no data was received but an extended error from the socket error queue.

12.1.6.16 MSG_NOSIGNAL

```
include std/socket.e
public constant MSG_NOSIGNAL
```

Requests not to send SIGPIPE on errors on stream oriented sockets when the other end breaks the connection. The EPIPE error is still returned.

12.1.6.17 MSG_MORE

```
include std/socket.e
public constant MSG_MORE
```

The caller has more data to send. This flag is used with TCP sockets to obtain the same effect as the TCP_CORK socket option, with the difference that this flag can be set on a per-call basis.

12.1.7 Server and Client sides

12.1.7.1 SOCKET_SOCKET

```
include std/socket.e
export enum SOCKET_SOCKET
```

Accessor index for socket handle of a socket type

12.1.7.2 SOCKET_SOCKADDR_IN

```
include std/socket.e
export enum SOCKET_SOCKADDR_IN
```

Accessor index for the sockaddr_in pointer of a socket type

12.1.7.3 socket

```
include std/socket.e
public type socket(object o)
```

Socket type

12.1.7.4 create

```
include std/socket.e
public function create(integer family, integer sock_type, integer protocol)
```

Create a new socket

12.1.7.4.1 Parameters:

1. family: an integer
2. sock_type: an integer, the type of socket to create
3. protocol: an integer, the communication protocol being used

family options:

- [AF_UNIX](#)
- [AF_INET](#)
- [AF_INET6](#)

- [AF_APPLETALK](#)
- [AF_BTH](#)

sock_type options:

- [SOCK_STREAM](#)
- [SOCK_DGRAM](#)
- [SOCK_RAW](#)
- [SOCK_RDM](#)
- [SOCK_SEQPACKET](#)

12.1.7.4.2 Returns:

An **object**, -1 on failure, else a supposedly valid socket id.

12.1.7.4.3 Example 1:

```
socket = create(AF_INET, SOCK_STREAM, 0)
```

12.1.7.5 close

```
include std/socket.e  
public function close(socket sock)
```

Closes a socket.

12.1.7.5.1 Parameters:

1. sock: the socket to close

12.1.7.5.2 Returns:

An **integer**, 0 on success and -1 on error.

12.1.7.5.3 Comments:

It may take several minutes for the OS to declare the socket as closed.

12.1.7.6 shutdown

```
include std/socket.e
public function shutdown(socket sock, atom method = SD_BOTH)
```

Partially or fully close a socket.

12.1.7.6.1 Parameters:

1. `sock` : the socket to shutdown
2. `method` : the method used to close the socket

12.1.7.6.2 Returns:

An **integer**, 0 on success and -1 on error.

12.1.7.6.3 Comments:

Three constants are defined that can be sent to `method`:

- **SD_SEND** - shutdown the send operations.
- **SD_RECEIVE** - shutdown the receive operations.
- **SD_BOTH** - shutdown both send and receive operations.

It may take several minutes for the OS to declare the socket as closed.

12.1.7.7 select

```
include std/socket.e
public function select(object sockets_read, object sockets_write, object sockets_err, integer t
```

Determine the read, write and error status of one or more sockets.

Using `select`, you can check to see if a socket has data waiting and is read to be read, if a socket can be written to and if a socket has an error status.

`select` allows for fine-grained control over your sockets, allow you to specify that a given socket only be checked for reading or for only reading and writing, etc.

12.1.7.7.1 Parameters:

1. `sockets_read` : either one socket or a sequence of sockets to check for reading.
2. `sockets_write` : either one socket or a sequence of sockets to check for writing.
3. `sockets_err` : either one socket or a sequence of sockets to check for errors.

4. `timeout` : maximum time to wait to determine a sockets status, seconds part
5. `timeout_micro` : maximum time to wait to determine a sockets status, microsecond part

12.1.7.7.2 Returns:

A **sequence**, of the same size of all unique sockets containing { `socket`, `read_status`, `write_status`, `error_status` } for each socket passed 2 to the function. Note that the sockets returned are not guaranteed to be in any particular order.

12.1.7.8 send

```
include std/socket.e
public function send(socket sock, sequence data, atom flags = 0)
```

Send TCP data to a socket connected remotely.

12.1.7.8.1 Parameters:

1. `sock` : the socket to send data to
2. `data` : a sequence of atoms, what to send
3. `flags` : flags (see [Send Flags](#))

12.1.7.8.2 Returns:

An **integer**, the number of characters sent, or -1 for an error.

12.1.7.9 receive

```
include std/socket.e
public function receive(socket sock, atom flags = 0)
```

Receive data from a bound socket.

12.1.7.9.1 Parameters:

1. `sock` : the socket to get data from
2. `flags` : flags (see [Send Flags](#))

12.1.7.9.2 Returns:

A **sequence**, either a full string of data on success, or an atom indicating the error code.

12.1.7.9.3 Comments:

This function will not return until data is actually received on the socket, unless the flags parameter contains [MSG_DONTWAIT](#).

[MSG_DONTWAIT](#) only works on Linux kernels 2.4 and above. To be cross-platform you should use [select](#) to determine if a socket is readable, i.e. has data waiting.

12.1.7.10 get_option

```
include std/socket.e
public function get_option(socket sock, integer level, integer optname)
```

Get options for a socket.

12.1.7.10.1 Parameters:

1. `sock` : the socket
2. `level` : an integer, the option level
3. `optname` : requested option (See [Socket Options](#))

12.1.7.10.2 Returns:

An **object**, either:

- On error, {"ERROR",error_code}.
- On success, either an atom or a sequence containing the option value, depending on the option.

12.1.7.10.3 Comments:

Primarily for use in multicast or more advanced socket applications. Level is the option level, and option_name is the option for which values are being sought. Level is usually [SOL_SOCKET](#).

12.1.7.10.4 Returns:

An **atom**, On error, an atom indicating the error code.

A **sequence** or **atom**, On success, either an atom or a sequence containing the option value.

12.1.7.10.5 See also:

[get_option](#)

12.1.7.11 set_option

```
include std/socket.e
public function set_option(socket sock, integer level, integer optname, object val)
```

Set options for a socket.

12.1.7.11.1 Parameters:

1. `sock` : an atom, the socket id
2. `level` : an integer, the option level
3. `optname` : requested option (See [Socket Options](#))
4. `val` : an object, the new value for the option

12.1.7.11.2 Returns:

An **integer**, 0 on success, -1 on error.

12.1.7.11.3 Comments:

Primarily for use in multicast or more advanced socket applications. Level is the option level, and option_name is the option for which values are being set. Level is usually [SOL_SOCKET](#).

12.1.7.11.4 See Also:

[get_option](#)

12.1.8 Client side only

12.1.8.1 connect

```
include std/socket.e
public function connect(socket sock, sequence address, integer port = - 1)
```

Establish an outgoing connection to a remote computer. Only works with TCP sockets.

12.1.8.1.1 Parameters:

1. `sock` : the socket
2. `address` : ip address to connect, optionally with `:PORT` at the end
3. `port` : port number

12.1.8.1.2 Returns:

An **integer**, 0 for success and -1 on failure.

12.1.8.1.3 Comments:

`address` can contain a port number. If it does not, it has to be supplied to the `port` parameter.

12.1.8.1.4 Example 1:

```
success = connect(sock, "11.1.1.1") -- uses default port 80
success = connect(sock, "11.1.1.1:110") -- uses port 110
success = connect(sock, "11.1.1.1", 345) -- uses port 345
```

12.1.9 Server side only

12.1.9.1 bind

```
include std/socket.e
public function bind(socket sock, sequence address, integer port = - 1)
```

Joins a socket to a specific local internet address and port so later calls only need to provide the socket.

12.1.9.1.1 Parameters:

1. `sock` : the socket
2. `address` : the address to bind the socket to
3. `port` : optional, if not specified you must include `:PORT` in the address parameter.

12.1.9.1.2 Returns:

An **integer**, 0 on success and -1 on failure.

12.1.9.1.3 Example 1:

```
-- Bind to all interfaces on the default port 80.
success = bind(socket, "0.0.0.0")
-- Bind to all interfaces on port 8080.
success = bind(socket, "0.0.0.0:8080")
-- Bind only to the 243.17.33.19 interface on port 345.
success = bind(socket, "243.17.33.19", 345)
```

12.1.9.2 listen

```
include std/socket.e
public function listen(socket sock, integer backlog)
```

Start monitoring a connection. Only works with TCP sockets.

12.1.9.2.1 Parameters:

1. `sock` : the socket
2. `backlog` : the number of connection requests that can be kept waiting before the OS refuses to hear any more.

12.1.9.2.2 Returns:

An **integer**, 0 on success and an error code on failure.

12.1.9.2.3 Comments:

Once the socket is created and bound, this will indicate to the operating system that you are ready to being listening for connections.

The value of `backlog` is strongly dependent on both the hardware and the amount of time it takes the program to process each connection request.

This function must be executed after `bind()`.

12.1.9.3 accept

```
include std/socket.e
public function accept(socket sock)
```

Produces a new socket for an incoming connection.

12.1.9.3.1 Parameters:

1. `sock`: the server socket

12.1.9.3.2 Returns:

An **atom**, on error

A **sequence**, {`socket client`, `sequence client_ip_address`} on success.

12.1.9.3.3 Comments:

Using this function allows communication to occur on a "side channel" while the main server socket remains available for new connections.

`accept()` must be called after `bind()` and `listen()`.

12.1.10 UDP only

12.1.10.1 `send_to`

```
include std/socket.e
public function send_to(socket sock, sequence data, sequence address, integer port = - 1, atom
```

Send a UDP packet to a given socket

12.1.10.1.1 Parameters:

1. `sock`: the server socket
2. `data`: the data to be sent
3. `ip`: the ip where the data is to be sent to (ip:port) is acceptable
4. `port`: the port where the data is to be sent on (if not supplied with the ip)
5. `flags`: flags (see [Send Flags](#))

12.1.10.1.2 Returns:

An `integer` status code.

12.1.10.1.3 See Also:

[receive_from](#)

12.1.10.2 receive_from

```
include std/socket.e
public function receive_from(socket sock, atom flags = 0)
```

Receive a UDP packet from a given socket

12.1.10.2.1 Parameters:

1. sock: the server socket
2. flags : flags (see [Send Flags](#))

12.1.10.2.2 Returns:

A sequence containing { client_ip, client_port, data } or an atom error code.

12.1.10.2.3 See Also:

[send_to](#)

12.1.11 Information

12.1.11.1 service_by_name

```
include std/socket.e
public function service_by_name(sequence name, object protocol = 0)
```

Get service information by name.

12.1.11.1.1 Parameters:

1. name : service name.
2. protocol : protocol. Default is not to search by protocol.

12.1.11.1.2 Returns:

A **sequence**, containing { official protocol name, protocol, port number } or an atom indicating the error code.

12.1.11.1.3 Example 1:

```
object result = getservbyname("http")
-- result = { "http", "tcp", 80 }
```

12.1.11.1.4 See Also:

[service_by_port](#)

12.1.11.2 service_by_port

```
include std/socket.e
public function service_by_port(integer port, object protocol = 0)
```

Get service information by port number.

12.1.11.2.1 Parameters:

1. port : port number.
2. protocol : protocol. Default is not to search by protocol.

12.1.11.2.2 Returns:

A **sequence**, containing { official protocol name, protocol, port number } or an atom indicating the error code.

12.1.11.2.3 Example 1:

```
object result = getservbyport(80)
-- result = { "http", "tcp", 80 }
```

12.1.11.2.4 See Also:

[service_by_name](#)

12.2 Common Internet Routines

12.2.1 IP Address Handling

12.2.1.1 is_inetaddr

```
include std/net/common.e
public function is_inetaddr(sequence address)
```

Checks if x is an IP address in the form (#.#.#.#[:#])

12.2.1.1.1 Parameters:

1. address : the address to check

12.2.1.1.2 Returns:

An **integer**, 1 if x is an inetaddr, 0 if it is not

12.2.1.1.3 Comments:

Some ip validation algorithms do not allow 0.0.0.0. We do here because many times you will want to bind to 0.0.0.0. However, you cannot connect to 0.0.0.0 of course.

With sockets, normally binding to 0.0.0.0 means bind to all interfaces that the computer has.

12.2.1.2 parse_ip_address

```
include std/net/common.e
public function parse_ip_address(sequence address, integer port = - 1)
```

Converts a text "address:port" into {"address", port} format.

12.2.1.2.1 Parameters:

1. address : ip address to connect, optionally with :PORT at the end
2. port : optional, if not specified you may include :PORT in the address parameter otherwise the default port 80 is used.

12.2.1.2.2 Comments:

If port is supplied, it overrides any ":PORT" value in the input address.

12.2.1.2.3 Returns:

A **sequence**, of two elements: "address" and integer port number.

12.2.1.2.4 Example 1:

```
addr = parse_ip_address("11.1.1.1") --> {"11.1.1.1", 80} -- default port
addr = parse_ip_address("11.1.1.1:110") --> {"11.1.1.1", 110}
addr = parse_ip_address("11.1.1.1", 345) --> {"11.1.1.1", 345}
```

12.2.2 URL Parsing

12.2.2.1 URL_ENTIRE

```
include std/net/common.e
public constant URL_ENTIRE
```

12.2.2.2 URL_PROTOCOL

```
include std/net/common.e
public constant URL_PROTOCOL
```

12.2.2.3 URL_HTTP_DOMAIN

```
include std/net/common.e
public constant URL_HTTP_DOMAIN
```

12.2.2.4 URL_HTTP_PATH

```
include std/net/common.e
public constant URL_HTTP_PATH
```

12.2.2.5 URL_HTTP_QUERY

```
include std/net/common.e
public constant URL_HTTP_QUERY
```

12.2.2.6 URL_MAIL_ADDRESS

```
include std/net/common.e
public constant URL_MAIL_ADDRESS
```

12.2.2.7 URL_MAIL_USER

```
include std/net/common.e
public constant URL_MAIL_USER
```

12.2.2.8 URL_MAIL_DOMAIN

```
include std/net/common.e
public constant URL_MAIL_DOMAIN
```

12.2.2.9 URL_MAIL_QUERY

```
include std/net/common.e
public constant URL_MAIL_QUERY
```

12.2.2.10 parse_url

```
include std/net/common.e
public function parse_url(sequence url)
```

Parse a common URL. Currently supported URLs are http(s), ftp(s), gopher(s) and mailto.

12.2.2.10.1 Parameters:

1. url : url to be parsed

12.2.2.10.2 Returns:

A **sequence**, containing the URL details. You should use the `URL_` constants to access the values of the returned sequence. You should first check the protocol ([URL_PROTOCOL](#)) that was returned as the data contained in the return value can be of different lengths.

On a parse error, -1 will be returned.

12.2.2.10.3 Example 1:

```
object url_data = parse_url("http://john.com/index.html?name=jeff")
-- url_data = {
--   "http://john.com/index.html?name=jeff", -- URL_ENTIRE
--   "http", -- URL_PROTOCOL
--   "john.com", -- URL_DOMAIN
--   "/index.html", -- URL_PATH
--   "?name=jeff" -- URL_QUERY
-- }

url_data = parse_url("mailto:john@mail.doe.com?subject=Hello%20John%20Doe")
-- url_data = {
--   "mailto:john@mail.doe.com?subject=Hello%20John%20Doe",
--   "mailto",
--   "john@mail.doe.com",
--   "john",
--   "mail.doe.com",
--   "?subject=Hello%20John%20Doe"
-- }
```

12.3 DNS

Based on EuNet project, version 1.3.2 at SourceForge.

Constants

```
ADDR_FLAGS
ADDR_FAMILY
ADDR_TYPE
ADDR_PROTOCOL
ADDR_ADDRESS
HOST_OFFICIAL_NAME
HOST_ALIASES
HOST_IPS
HOST_TYPE
DNS_QUERY_STANDARD
DNS_QUERY_ACCEPT_TRUNCATED_RESPONSE
DNS_QUERY_USE_TCP_ONLY
DNS_QUERY_NO_RECURSION
DNS_QUERY_BYPASS_CACHE
DNS_QUERY_NO_WIRE_QUERY
DNS_QUERY_NO_LOCAL_NAME
DNS_QUERY_NO_HOSTS_FILE
DNS_QUERY_NO_NETBT
DNS_QUERY_WIRE_ONLY
DNS_QUERY_RETURN_MESSAGE
DNS_QUERY_TREAT_AS_FQDN
DNS_QUERY_DONT_RESET_TTL_VALUES
DNS_QUERY_RESERVED
NS_C_IN
```

```
NS_C_ANY
NS_KT_RSA
NS_KT_DH
NS_KT_DSA
NS_KT_PRIVATE
NS_T_A
NS_T_NS
NS_T_PTR
NS_T_MX
NS_T_AAAA
NS_T_A6
NS_T_ANY
General Routines
  host_by_name
  host_by_addr
```

12.3.1 Constants

12.3.1.1 ADDR_FLAGS

```
include std/net/dns.e
public enum ADDR_FLAGS
```

12.3.1.2 ADDR_FAMILY

```
include std/net/dns.e
public enum ADDR_FAMILY
```

12.3.1.3 ADDR_TYPE

```
include std/net/dns.e
public enum ADDR_TYPE
```

12.3.1.4 ADDR_PROTOCOL

```
include std/net/dns.e
public enum ADDR_PROTOCOL
```

12.3.1.5 ADDR_ADDRESS

```
include std/net/dns.e
public enum ADDR_ADDRESS
```

12.3.1.6 HOST_OFFICIAL_NAME

```
include std/net/dns.e
public enum HOST_OFFICIAL_NAME
```

12.3.1.7 HOST_ALIASES

```
include std/net/dns.e
public enum HOST_ALIASES
```

12.3.1.8 HOST_IPS

```
include std/net/dns.e
public enum HOST_IPS
```

12.3.1.9 HOST_TYPE

```
include std/net/dns.e
public enum HOST_TYPE
```

12.3.1.10 DNS_QUERY_STANDARD

```
include std/net/dns.e
public constant DNS_QUERY_STANDARD
```

12.3.1.11 DNS_QUERY_ACCEPT_TRUNCATED_RESPONSE

```
include std/net/dns.e
public constant DNS_QUERY_ACCEPT_TRUNCATED_RESPONSE
```

12.3.1.12 DNS_QUERY_USE_TCP_ONLY

```
include std/net/dns.e
public constant DNS_QUERY_USE_TCP_ONLY
```

12.3.1.13 DNS_QUERY_NO_RECURSION

```
include std/net/dns.e
public constant DNS_QUERY_NO_RECURSION
```

12.3.1.14 DNS_QUERY_BYPASS_CACHE

```
include std/net/dns.e
public constant DNS_QUERY_BYPASS_CACHE
```

12.3.1.15 DNS_QUERY_NO_WIRE_QUERY

```
include std/net/dns.e
public constant DNS_QUERY_NO_WIRE_QUERY
```

12.3.1.16 DNS_QUERY_NO_LOCAL_NAME

```
include std/net/dns.e
public constant DNS_QUERY_NO_LOCAL_NAME
```

12.3.1.17 DNS_QUERY_NO_HOSTS_FILE

```
include std/net/dns.e
public constant DNS_QUERY_NO_HOSTS_FILE
```

12.3.1.18 DNS_QUERY_NO_NETBT

```
include std/net/dns.e
public constant DNS_QUERY_NO_NETBT
```

12.3.1.19 DNS_QUERY_WIRE_ONLY

```
include std/net/dns.e
public constant DNS_QUERY_WIRE_ONLY
```

12.3.1.20 DNS_QUERY_RETURN_MESSAGE

```
include std/net/dns.e
public constant DNS_QUERY_RETURN_MESSAGE
```

12.3.1.21 DNS_QUERY_TREAT_AS_FQDN

```
include std/net/dns.e
public constant DNS_QUERY_TREAT_AS_FQDN
```

12.3.1.22 DNS_QUERY_DONT_RESET_TTL_VALUES

```
include std/net/dns.e
public constant DNS_QUERY_DONT_RESET_TTL_VALUES
```

12.3.1.23 DNS_QUERY_RESERVED

```
include std/net/dns.e
public constant DNS_QUERY_RESERVED
```

12.3.1.24 NS_C_IN

```
include std/net/dns.e
public constant NS_C_IN
```

12.3.1.25 NS_C_ANY

```
include std/net/dns.e
public constant NS_C_ANY
```

12.3.1.26 NS_KT_RSA

```
include std/net/dns.e
public constant NS_KT_RSA
```

12.3.1.27 NS_KT_DH

```
include std/net/dns.e
public constant NS_KT_DH
```

12.3.1.28 NS_KT_DSA

```
include std/net/dns.e
public constant NS_KT_DSA
```

12.3.1.29 NS_KT_PRIVATE

```
include std/net/dns.e
public constant NS_KT_PRIVATE
```

12.3.1.30 NS_T_A

```
include std/net/dns.e
public constant NS_T_A
```

12.3.1.31 NS_T_NS

```
include std/net/dns.e
public constant NS_T_NS
```

12.3.1.32 NS_T_PTR

```
include std/net/dns.e
public constant NS_T_PTR
```


12.3.1.33 NS_T_MX

```
include std/net/dns.e
public constant NS_T_MX
```

12.3.1.34 NS_T_AAAA

```
include std/net/dns.e
public constant NS_T_AAAA
```

12.3.1.35 NS_T_A6

```
include std/net/dns.e
public constant NS_T_A6
```

12.3.1.36 NS_T_ANY

```
include std/net/dns.e
public constant NS_T_ANY
```

12.3.2 General Routines

12.3.2.1 host_by_name

```
include std/net/dns.e
public function host_by_name(sequence name)
```

Get the host information by name.

12.3.2.1.1 Parameters:

1. name : host name

12.3.2.1.2 Returns:

A sequence, containing

```
{
    official name,
    { alias1, alias2, ... },
}
```



```
{ ip1, ip2, ... },  
address_type  
}
```

12.3.2.1.3 Example 1:

```
object data = host_by_name("www.google.com")  
-- data = {  
--   "www.l.google.com",  
--   {  
--     "www.google.com"  
--   },  
--   {  
--     "74.125.93.104",  
--     "74.125.93.147",  
--     ...  
--   },  
--   2  
-- }
```

12.3.2.2 host_by_addr

```
include std/net/dns.e  
public function host_by_addr(sequence address)
```

Get the host information by address.

12.3.2.2.1 Parameters:

1. address : host address

12.3.2.2.2 Returns:

A sequence, containing

```
{  
  official name,  
  { alias1, alias2, ... },  
  { ip1, ip2, ... },  
  address_type  
}
```

12.3.2.2.3 Example 1:

```
object data = host_by_addr("74.125.93.147")  
-- data = {  
--   "www.l.google.com",  
--   {  
--     {  
--       "www.google.com",  
--       {  
--         "74.125.93.104",  
--         "74.125.93.147",  
--         ...  
--       },  
--       2  
--     }  
--   }  
-- }
```

```
--      "www.google.com"
--    },
--    {
--      "74.125.93.104",
--      "74.125.93.147",
--      ...
--    },
--    2
--  }
```

12.4 HTTP

Constants

- HTTP_HEADER_HTTPVERSION
- HTTP_HEADER_GET
- HTTP_HEADER_HOST
- HTTP_HEADER_REFERER
- HTTP_HEADER_USERAGENT
- HTTP_HEADER_ACCEPT
- HTTP_HEADER_ACCEPTCHARSET
- HTTP_HEADER_ACCEPTENCODING
- HTTP_HEADER_ACCEPTLANGUAGE
- HTTP_HEADER_ACCEPTRANGES
- HTTP_HEADER_AUTHORIZATION
- HTTP_HEADER_DATE
- HTTP_HEADER_IFMODIFIEDSINCE
- HTTP_HEADER_POST
- HTTP_HEADER_POSTDATA
- HTTP_HEADER_CONTENTTYPE
- HTTP_HEADER_CONTENTLENGTH
- HTTP_HEADER_FROM
- HTTP_HEADER_KEEPALIVE
- HTTP_HEADER_CACHECONTROL
- HTTP_HEADER_CONNECTION

Header management

- get_sendheader
- set_sendheader_default
- set_sendheader
- set_sendheader_useragent_msie
- parse_rcvheader
- get_rcvheader

Web interface

- get_http
- get_http_use_cookie
- get_url

12.4.1 Constants

12.4.1.1 HTTP_HEADER_HTTPVERSION

```
include std/net/http.e
public constant HTTP_HEADER_HTTPVERSION
```

12.4.1.2 HTTP_HEADER_GET

```
include std/net/http.e
public constant HTTP_HEADER_GET
```

12.4.1.3 HTTP_HEADER_HOST

```
include std/net/http.e
public constant HTTP_HEADER_HOST
```

12.4.1.4 HTTP_HEADER_REFERER

```
include std/net/http.e
public constant HTTP_HEADER_REFERER
```

12.4.1.5 HTTP_HEADER_USERAGENT

```
include std/net/http.e
public constant HTTP_HEADER_USERAGENT
```

12.4.1.6 HTTP_HEADER_ACCEPT

```
include std/net/http.e
public constant HTTP_HEADER_ACCEPT
```

12.4.1.7 HTTP_HEADER_ACCEPTCHARSET

```
include std/net/http.e
public constant HTTP_HEADER_ACCEPTCHARSET
```

12.4.1.8 HTTP_HEADER_ACCEPTENCODING

```
include std/net/http.e
public constant HTTP_HEADER_ACCEPTENCODING
```

12.4.1.9 HTTP_HEADER_ACCEPTLANGUAGE

```
include std/net/http.e
public constant HTTP_HEADER_ACCEPTLANGUAGE
```

12.4.1.10 HTTP_HEADER_ACCEPTRANGES

```
include std/net/http.e
public constant HTTP_HEADER_ACCEPTRANGES
```

12.4.1.11 HTTP_HEADER_AUTHORIZATION

```
include std/net/http.e
public constant HTTP_HEADER_AUTHORIZATION
```

12.4.1.12 HTTP_HEADER_DATE

```
include std/net/http.e
public constant HTTP_HEADER_DATE
```

12.4.1.13 HTTP_HEADER_IFMODIFIEDSINCE

```
include std/net/http.e
public constant HTTP_HEADER_IFMODIFIEDSINCE
```

12.4.1.14 HTTP_HEADER_POST

```
include std/net/http.e
public constant HTTP_HEADER_POST
```

12.4.1.15 HTTP_HEADER_POSTDATA

```
include std/net/http.e
public constant HTTP_HEADER_POSTDATA
```

12.4.1.16 HTTP_HEADER_CONTENTTYPE

```
include std/net/http.e
public constant HTTP_HEADER_CONTENTTYPE
```

12.4.1.17 HTTP_HEADER_CONTENTLENGTH

```
include std/net/http.e
public constant HTTP_HEADER_CONTENTLENGTH
```

12.4.1.18 HTTP_HEADER_FROM

```
include std/net/http.e
public constant HTTP_HEADER_FROM
```

12.4.1.19 HTTP_HEADER_KEEPALIVE

```
include std/net/http.e
public constant HTTP_HEADER_KEEPALIVE
```

12.4.1.20 HTTP_HEADER_CACHECONTROL

```
include std/net/http.e
public constant HTTP_HEADER_CACHECONTROL
```

12.4.1.21 HTTP_HEADER_CONNECTION

```
include std/net/http.e
public constant HTTP_HEADER_CONNECTION
```

12.4.2 Header management

12.4.2.1 get_sendheader

```
include std/net/http.e
public function get_sendheader(object field)
```

Retrieve either the whole sendheader sequence, or just a single field.

12.4.2.1.1 Parameters:

1. `field`: an object indicating which part is being requested, see Comments section.

12.4.2.1.2 Returns:

An **object**, either:

- -1 if the field cannot be found,
- `{{"label", "delimiter", "value"}, ...}` for the whole sendheader sequence
- a three-element sequence in the form `{"label", "delimiter", "value"}` when only a single field is selected.

12.4.2.1.3 Comments:

`field` can be either an `HTTP_HEADER_xxx` access constant, the number 0 to retrieve the whole sendheader sequence, or a string matching one of the header field labels. The string is not case sensitive.

12.4.2.2 set_sendheader_default

```
include std/net/http.e
public procedure set_sendheader_default()
```

Sets header elements to default values. The default User Agent is Opera (currently the most standards compliant). Before setting any header option individually, programs must call this procedure.

12.4.2.2.1 See Also:

[get_sendheader](#), [set_sendheader](#), [set_sendheader_useragent_msie](#)

12.4.2.3 set_sendheader

```
include std/net/http.e
public procedure set_sendheader(object whatheader, sequence whatdata)
```

Set an individual header field.

12.4.2.3.1 Parameters:

1. `whatheader` : an object, either an explicit name string or a `HTTP_HEADER_XXX` constant
2. `whatdata` : a string, the associated data

12.4.2.3.2 Comments:

If the requested field is not one of the default header fields, the field **MUST** be set by string. This will increase the length of the header overall.

12.4.2.3.3 Example 1:

```
set_sendheader("Referer", "search.yahoo.com")
```

12.4.2.3.4 See Also:

[get_sendheader](#)

12.4.2.4 set_sendheader_useragent_msie

```
include std/net/http.e
public procedure set_sendheader_useragent_msie()
```

Inform listener that user agent is Microsoft (R) Internet Explorer (TM).

12.4.2.4.1 Comments:

This is a convenience procedure to tell a website that a Microsoft Internet Explorer (TM) browser is requesting data. Because some websites format their response differently (or simply refuse data) depending on the browser, this procedure provides a quick means around that. For example, see:

<http://www.missporters.org/podium/nonsupport.aspx>

12.4.2.5 parse_rcvheader

```
include std/net/http.e
public procedure parse_rcvheader(sequence header)
```

Populates the internal sequence rcvheader from the flat string header.

12.4.2.5.1 Parameters:

1. header : a string, the header data

12.4.2.5.2 Comments:

This must be called prior to calling [get_rcvheader\(\)](#).

12.4.2.6 get_rcvheader

```
include std/net/http.e
public function get_rcvheader(object field)
```

Return the value of a named field in the received http header as returned by the most recent call to [get_http](#).

12.4.2.6.1 Parameters:

1. field : an object, either a string holding a field name (case insensitive), 0 to return the whole header, or a numerical index.

12.4.2.6.2 Returns:

An object,

- -1 on error
- a sequence in the form, {field name, field value} on success.

12.4.3 Web interface

12.4.3.1 get_http

```
include std/net/http.e
public function get_http(sequence inet_addr, sequence hostname, sequence file, integer timeout)
```

Returns data from an http internet site.

12.4.3.1.1 Parameters:

1. `inet_addr` : a sequence holding an address
2. `hostname` : a string, the name for the host
3. `file` : a file name to transmit

12.4.3.1.2 Returns:

A **sequence**, empty sequence on error, of length 2 on success, like {sequence header, sequence data}.

12.4.3.2 `get_http_use_cookie`

```
include std/net/http.e
public function get_http_use_cookie(sequence inet_addr, sequence hostname, sequence file)
```

Works the same as `get_url()`, but maintains an internal state register based on cookies received.

12.4.3.2.1 Warning:

This function is not yet implemented.

12.4.3.2.2 Parameters:

1. `inet_addr` : a sequence holding an address
2. `hostname` : a string, the name for the host
3. `file` : a file name to transmit

12.4.3.2.3 Returns:

A **sequence**, {header, body} on success, or an empty sequence on error.

12.4.3.2.4 Example 1:

```
addrinfo = getaddrinfo("www.yahoo.com", "http", 0)
if atom(addrinfo) or length(addrinfo) < 1 or
   length(addrinfo[1]) < 5 then
   puts(1, "Uh, oh")
   return {}
else
   inet_addr = addrinfo[1][5]
```

```
end if
data = get_http_use_cookie(inet_addr,"www.yahoo.com","")
```

12.4.3.2.5 See also:

[get_url](#)

12.4.3.3 get_url

```
include std/net/http.e
public function get_url(sequence url, sequence post_data = "")
```

Returns data from an http internet site.

12.4.3.3.1 Parameters:

1. url: URL to access
2. post_data: Optional post data

12.4.3.3.2 Returns:

A **sequence** {header, body} on success, or an empty sequence on error.

12.4.3.3.3 Comments:

If post_data is empty, then a normal GET request is done. If post_data is non-empty then get_url will perform a POST request and supply post_data during the request.

12.4.3.3.4 Example 1:

```
url = "http://banners.wunderground.com/weathersticker/mini" &
      "Weather2_metric_cond/language/www/US/PA/Philadelphia.gif"

temp = get_url(url)
if length(temp)>=2 and length(temp[2])>0 then
    tempfp = open(TEMPDIR&"current_weather.gif","wb")
    puts(tempfp,temp[2])
    close(tempfp)
end if
```

12.5 URL handling

12.5.1 Parsing

12.5.1.1 parse_querystring

```
include std/net/url.e
public function parse_querystring(object query_string)
```

Parse a query string into a map

12.5.1.1.1 Parameters:

1. query_string: Query string to parse

12.5.1.1.2 Returns:

map containing the key/value pairs

12.5.1.1.3 Example 1:

```
map qs = parse_querystring("name=John&age=18")
printf(1, "%s is %s years old\n", { map:get(qs, "name"), map:get(qs, "age") })
```

12.5.1.2 URL_PROTOCOL

```
include std/net/url.e
public enum URL_PROTOCOL
```

12.5.1.3 URL_HOSTNAME

```
include std/net/url.e
public enum URL_HOSTNAME
```

12.5.1.4 URL_PORT

```
include std/net/url.e
public enum URL_PORT
```

12.5.1.5 URL_PATH

```
include std/net/url.e
public enum URL_PATH
```

12.5.1.6 URL_USER

```
include std/net/url.e
public enum URL_USER
```

12.5.1.7 URL_PASSWORD

```
include std/net/url.e
public enum URL_PASSWORD
```

12.5.1.8 URL_QUERY_STRING

```
include std/net/url.e
public enum URL_QUERY_STRING
```

12.5.1.9 parse

```
include std/net/url.e
public function parse(sequence url, integer querystring_also = 0)
```

Parse a URL returning its various elements.

12.5.1.9.1 Parameters:

1. url: URL to parse
2. querystring_also: Parse the query string into a map also?

12.5.1.9.2 Returns:

A multi-element sequence containing:

1. protocol
2. host name
3. port
4. path
5. user name

6. password
7. query string

Or, zero if the URL could not be parsed.

12.5.1.9.3 Notes:

If the host name, port, path, username, password or query string are not part of the URL they will be returned as an integer value of zero.

12.5.1.9.4 Example 1:

```
sequence parsed = parse("http://user:pass@www.debian.org:80/index.html?name=John&age=39")
-- parsed is
-- {
--     "http",
--     "www.debian.org",
--     80,
--     "/index.html",
--     "user",
--     "pass",
--     "name=John&age=39"
-- }
```

12.5.2 URL encoding and decoding

12.5.2.1 encode

```
include std/net/url.e
public function encode(sequence what, sequence spacecode = "+")
```

Converts all non-alphanumeric characters in a string to their percent-sign hexadecimal representation, or plus sign for spaces.

12.5.2.1.1 Parameters:

1. what : the string to encode
2. spacecode : what to insert in place of a space

12.5.2.1.2 Returns:

A **sequence**, the encoded string.

12.5.2.1.3 Comments:

spacecode defaults to + as it is more correct, however, some sites want %20 as the space encoding.

12.5.2.1.4 Example 1:

```
puts(1, encode("Fred & Ethel"))
-- Prints "Fred+%26+Ethel"
```

12.5.2.1.5 See Also:

[decode](#)

12.5.2.2 decode

```
include std/net/url.e
public function decode(sequence what)
```

Convert all encoded entities to their decoded counter parts

12.5.2.2.1 Parameters:

1. what: what value to decode

12.5.2.2.2 Returns:

A decoded sequence

12.5.2.2.3 Example 1:

```
puts(1, decode("Fred+%26+Ethel"))
-- Prints "Fred & Ethel"
```

12.5.2.2.4 See Also:

[encode](#)

13 Low Level Routines

Dynamic Linking to external code

 C Type Constants

 External EUPHORIA Type Constants

 Constants

 Routines

Errors and Warnings

 Routines

Pseudo Memory

Indirect Routine Calling

 SAFE mode

 Data Execute mode

 Accessing EUPHORIA coded routines

 Accessing EUPHORIA internals

Types supporting Memory

 Allocating and Writing to memory:

 Memory disposal

Memory Management - Low-Level

 Usage Notes

 Memory allocation

 Reading from, Writing to, and Calling into Memory

 Safe memory access

 safe.e

 Microsoft Windows Memory Protection Constants

 Standard Library Memory Protection Constants

13.1 Dynamic Linking to external code

 C Type Constants

 C_CHAR

 C_BYTE

 C_UCHAR

 C_UBYTE

 C_SHORT

 C_WORD

 C_USHORT

 C_INT

 C_BOOL

 C_UINT

 C_SIZE_T

 C_LONG

 C_ULONG

 C_POINTER

 C_HANDLE

- C_HWND
- C_DWORD
- C_WPARAM
- C_LPARAM
- C_HRESULT
- C_FLOAT
- C_DOUBLE
- C_DWORDLONG
- External EUPHORIA Type Constants
 - E_INTEGER
 - E_ATOM
 - E_SEQUENCE
 - E_OBJECT
- Constants
 - NULL
- Routines
 - open_dll
 - define_c_var
 - define_c_proc
 - define_c_func
 - c_func
 - c_proc
 - call_back

13.1.1 C Type Constants

These C type constants are used when defining external C functions in a shared library file.

13.1.1.1 Example 1:

See [define_c_proc](#)

13.1.1.1.1 See Also:

[define_c_proc](#), [define_c_func](#), [define_c_var](#)

13.1.1.2 C_CHAR

```
include std/dll.e
public constant C_CHAR
```

char 8-bits

13.1.1.3 C_BYTE

```
include std/dll.e  
public constant C_BYTE
```

byte 8-bits

13.1.1.4 C_UCHAR

```
include std/dll.e  
public constant C_UCHAR
```

unsigned char 8-bits

13.1.1.5 C_UBYTE

```
include std/dll.e  
public constant C_UBYTE
```

ubyte 8-bits

13.1.1.6 C_SHORT

```
include std/dll.e  
public constant C_SHORT
```

short 16-bits

13.1.1.7 C_WORD

```
include std/dll.e  
public constant C_WORD
```

word 16-bits

13.1.1.8 C_USHORT

```
include std/dll.e  
public constant C_USHORT
```

unsigned short 16-bits

13.1.1.9 C_INT

```
include std/dll.e  
public constant C_INT
```

int 32-bits

13.1.1.10 C_BOOL

```
include std/dll.e  
public constant C_BOOL
```

bool 32-bits

13.1.1.11 C_UINT

```
include std/dll.e  
public constant C_UINT
```

unsigned int 32-bits

13.1.1.12 C_SIZE_T

```
include std/dll.e  
public constant C_SIZE_T
```

size_t 32-bits

13.1.1.13 C_LONG

```
include std/dll.e  
public constant C_LONG
```

long 32-bits

13.1.1.14 C_ULONG

```
include std/dll.e  
public constant C_ULONG
```

unsigned long 32-bits

13.1.1.15 C_POINTER

```
include std/dll.e  
public constant C_POINTER
```

any valid pointer 32-bits

13.1.1.16 C_HANDLE

```
include std/dll.e  
public constant C_HANDLE
```

handle 32-bits

13.1.1.17 C_HWND

```
include std/dll.e  
public constant C_HWND
```

hwnd 32-bits

13.1.1.18 C_DWORD

```
include std/dll.e  
public constant C_DWORD
```

dword 32-bits

13.1.1.19 C_WPARAM

```
include std/dll.e  
public constant C_WPARAM
```

wparam 32-bits

13.1.1.20 C_LPARAM

```
include std/dll.e  
public constant C_LPARAM
```

lparam 32-bits

13.1.1.21 C_HRESULT

```
include std/dll.e  
public constant C_HRESULT
```

hresult 32-bits

13.1.1.22 C_FLOAT

```
include std/dll.e  
public constant C_FLOAT
```

float 32-bits

13.1.1.23 C_DOUBLE

```
include std/dll.e  
public constant C_DOUBLE
```

double 64-bits

13.1.1.24 C_DWORDLONG

```
include std/dll.e  
public constant C_DWORDLONG
```

dwordlong 64-bits

13.1.2 External EUPHORIA Type Constants

These are used for arguments to and the return value from a EUPHORIA shared library file (.dll, .so or .dylib).

13.1.2.1 E_INTEGER

```
include std/dll.e  
public constant E_INTEGER
```

integer

13.1.2.2 E_ATOM

```
include std/dll.e
public constant E_ATOM
```

atom

13.1.2.3 E_SEQUENCE

```
include std/dll.e
public constant E_SEQUENCE
```

sequence

13.1.2.4 E_OBJECT

```
include std/dll.e
public constant E_OBJECT
```

object

13.1.3 Constants

13.1.3.1 NULL

```
include std/dll.e
public constant NULL
```

C's NULL pointer

13.1.4 Routines

13.1.4.1 open_dll

```
include std/dll.e
public function open_dll(sequence file_name)
```

Open a Windows dynamic link library (.dll) file, or a *Unix* shared library (.so) file.

13.1.4.1.1 Parameters:

1. `file_name` : a sequence, the name of the shared library to open or a sequence of filename's to try to open.

13.1.4.1.2 Returns:

An **atom**, actually a 32-bit address. 0 is returned if the .dll can't be found.

13.1.4.1.3 Errors:

The length of `file_name` (or any filename contained therein) should not exceed 1,024 characters.

13.1.4.1.4 Comments:

`file_name` can be a relative or an absolute file name. Most operating systems will use the normal search path for locating non-relative files.

`file_name` can be a list of file names to try. On different Linux platforms especially, the filename will not always be the same. For instance, you may wish to try opening `libmylib.so`, `libmylib.so.1`, `libmylib.so.1.0`, `libmylib.so.1.0.0`. If given a sequence of file names to try, the first successful library loaded will be returned. If no library could be loaded, 0 will be returned after exhausting the entire list of file names.

The value returned by `open_dll()` can be passed to `define_c_proc()`, `define_c_func()`, or `define_c_var()`.

You can open the same .dll or .so file multiple times. No extra memory is used and you'll get the same number returned each time.

EUPHORIA will close the .dll/.so for you automatically at the end of execution.

13.1.4.1.5 Example 1:

```
atom user32
user32 = open_dll("user32.dll")
if user32 = 0 then
    puts(1, "Couldn't open user32.dll!\n")
end if
```

13.1.4.1.6 Example 2:

```
atom mysql_lib
mysql_lib = open_dll({"libmysqlclient.so", "libmysqlclient.so.15", "libmysqlclient.so.15.0"})
if mysql_lib = 0 then
    puts(1, "Couldn't find the mysql client library\n")
end if
```

13.1.4.1.7 See Also:

[define_c_func](#), [define_c_proc](#), [define_c_var](#), [c_func](#), [c_proc](#)

13.1.4.2 define_c_var

```
include std/dll.e
public function define_c_var(atom lib, sequence variable_name)
```

Gets the address of a symbol in a shared library or in RAM.

13.1.4.2.1 Parameters:

1. `lib` : an atom, the address of a Linux or FreeBSD shared library, or Windows .dll, as returned by `open_dll()`.
2. `variable_name` : a sequence, the name of a public C variable defined within the library.

13.1.4.2.2 Returns:

An **atom**, the memory address of `variable_name`.

13.1.4.2.3 Comments:

Once you have the address of a C variable, and you know its type, you can use `peek()` and `poke()` to read or write the value of the variable. You can in the same way obtain the address of a C function and pass it to any external routine that requires a callback address.

13.1.4.2.4 Example:

see `euphoria/demo/linux/mylib.ex`

13.1.4.2.5 See Also:

[c_proc](#), [define_c_func](#), [c_func](#), [open_dll](#)

13.1.4.3 define_c_proc

```
include std/dll.e
public function define_c_proc(object lib, object routine_name, sequence arg_types)
```

Define the characteristics of either a C function, or a machine-code routine that you wish to call as a

procedure from your EUPHORIA program.

13.1.4.3.1 Parameters:

1. `lib` : an object, either an entry point returned as an atom by `open_dll()`, or "" to denote a routine the RAM address is known.
2. `routine_name` : an object, either the name of a procedure in a shared object or the machine address of the procedure.
3. `argtypes` : a sequence of type constants.

13.1.4.3.2 Returns:

A small **integer**, known as a routine id, will be returned.

13.1.4.3.3 Errors:

The length of `name` should not exceed 1,024 characters.

13.1.4.3.4 Comments:

Use the returned routine id as the first argument to `c_proc()` when you wish to call the routine from EUPHORIA.

A returned value of -1 indicates that the procedure could not be found or linked to.

On Windows, you can add a '+' character as a prefix to the procedure name. This tells EUPHORIA that the function uses the `cdecl` calling convention. By default, EUPHORIA assumes that C routines accept the `stdcall` convention.

When defining a machine code routine, `lib` must be the empty sequence, "" or {}, and `routine_name` indicates the address of the machine code routine. You can poke the bytes of machine code into a block of memory reserved using `allocate()`. On Windows, the machine code routine is normally expected to follow the `stdcall` calling convention, but if you wish to use the `cdecl` convention instead, you can code {'+', address} instead of address.

`argtypes` is made of type constants, which describe the C types of arguments to the procedure. They may be used to define machine code parameters as well.

The C function that you define could be one created by the EUPHORIA To C Translator, in which case you can pass EUPHORIA data to it, and receive EUPHORIA data back. A list of EUPHORIA types is shown above.

You can pass any C integer type or pointer type. You can also pass a EUPHORIA atom as a C double or float.

Parameter types which use 4 bytes or less are all passed the same way, so it is not necessary to be exact.

Currently, there is no way to pass a C structure by value. You can only pass a pointer to a structure. However, you can pass a 64 bit integer by pretending to pass two C_LONG instead. When calling the routine, pass low doubleword first, then high doubleword.

The C function can return a value but it will be ignored. If you want to use the value returned by the C function, you must instead define it with [define_c_func\(\)](#) and call it with [c_func\(\)](#).

13.1.4.3.5 Example 1:

```
atom user32
integer ShowWindow

-- open user32.dll - it contains the ShowWindow C function
user32 = open_dll("user32.dll")

-- It has 2 parameters that are both C int.
ShowWindow = define_c_proc(user32, "ShowWindow", {C_INT, C_INT})
-- If ShowWindow used the cdecl convention,
-- we would have coded "+ShowWindow" here

if ShowWindow = -1 then
  puts(1, "ShowWindow not found!\n")
end if
```

13.1.4.3.6 See Also:

[c_proc](#), [define_c_func](#), [c_func](#), [open_dll](#)

13.1.4.4 define_c_func

```
include std/dll.e
public function define_c_func(object lib, object routine_name, sequence arg_types, atom return_
```

Define the characteristics of either a C function, or a machine-code routine that returns a value.

13.1.4.4.1 Parameters:

1. `lib` : an object, either an entry point returned as an atom by [open_dll\(\)](#), or "" to denote a routine the RAM address is known.
2. `routine_name` : an object, either the name of a procedure in a shared object or the machine address of the procedure.
3. `argtypes` : a sequence of type constants.
4. `return_type` : an atom, indicating what type the function will return.

13.1.4.4.2 Returns:

A small **integer**, known as a routine id, will be returned.

13.1.4.4.3 Errors:

The length of `name` should not exceed 1,024 characters.

13.1.4.4.4 Comments:

Use the returned routine id as the first argument to `c_proc()` when you wish to call the routine from EUPHORIA.

A returned value of -1 indicates that the procedure could not be found or linked to.

On Windows, you can add a '+' character as a prefix to the function name. This indicates to EUPHORIA that the function uses the `cdecl` calling convention. By default, EUPHORIA assumes that C routines accept the `stdcall` convention.

When defining a machine code routine, `x1` must be the empty sequence, `""` or `{ }`, and `x2` indicates the address of the machine code routine. You can poke the bytes of machine code into a block of memory reserved using `allocate()`. On Windows, the machine code routine is normally expected to follow the `stdcall` calling convention, but if you wish to use the `cdecl` convention instead, you can code `{ '+', address }` instead of `address` for `x2`.

The C function that you define could be one created by the EUPHORIA To C Translator, in which case you can pass EUPHORIA data to it, and receive EUPHORIA data back. A list of EUPHORIA types is contained in `dll.e`:

- `E_INTEGER = #06000004`
- `E_ATOM = #07000004`
- `E_SEQUENCE = #08000004`
- `E_OBJECT = #09000004`

You can pass or return any C integer type or pointer type. You can also pass a EUPHORIA atom as a C double or float, and get a C double or float returned to you as a EUPHORIA atom.

Parameter types which use 4 bytes or less are all passed the same way, so it is not necessary to be exact when choosing a 4-byte parameter type. However the distinction between signed and unsigned may be important when you specify the return type of a function.

Currently, there is no way to pass a C structure by value or get a C structure as a return result. You can only pass a pointer to a structure and get a pointer to a structure as a result. However, you can pass a 64 bit integer as two `C_LONG` instead. On calling the routine, pass low doubleword first, then high doubleword.

If you are not interested in using the value returned by the C function, you should instead define it with `define_c_proc()` and call it with `c_proc()`.

If you use `euiw` to call a cdecl C routine that returns a floating-point value, it might not work. This is because the Watcom C compiler (used to build `euiw`) has a non-standard way of handling cdecl floating-point return values.

Passing floating-point values to a machine code routine will be faster if you use `c_func()` rather than `call()` to call the routine, since you won't have to use `atom_to_float64()` and `poke()` to get the floating-point values into memory.

13.1.4.4.5 Example 1:

```
atom user32
integer LoadIcon

-- open user32.dll - it contains the LoadIconA C function
user32 = open_dll("user32.dll")

-- It takes a C pointer and a C int as parameters.
-- It returns a C int as a result.
LoadIcon = define_c_func(user32, "LoadIconA",
                        {C_POINTER, C_INT}, C_INT)
-- We use "LoadIconA" here because we know that LoadIconA
-- needs the stdcall convention, as do
-- all standard .dll routines in the WIN32 API.
-- To specify the cdecl convention, we would have used "+LoadIconA".

if LoadIcon = -1 then
    puts(1, "LoadIconA could not be found!\n")
end if
```

13.1.4.4.6 See Also:

`demo\callmach.ex`, `c_func`, `define_c_proc`, `c_proc`, `open_dll`

13.1.4.5 c_func

```
<built-in> function c_func(integer rid, sequence args={})
```

Call a C function, or machine code function, or translated/compiled EUPHORIA function by routine id.

13.1.4.5.1 Parameters:

1. `rid`: an integer, the `routine_id` of the external function being called.
2. `args`: a sequence, the list of parameters to pass to the function

13.1.4.5.2 Returns:

An **object**, whose type and meaning was defined on calling [define_c_func\(\)](#).

13.1.4.5.3 Errors:

If `rid` is not a valid routine id, or the arguments do not match the prototype of the routine being called, an error occurs.

13.1.4.5.4 Comments:

`rid` must have been returned by [define_c_func\(\)](#), **not** by [routine_id\(\)](#). The type checks are different, and you would get a machine level exception in the best case.

If the function does not take any arguments then `args` should be `{ }`.

If you pass an argument value which contains a fractional part, where the C function expects a C integer type, the argument will be rounded towards 0. e.g. 5.9 will be passed as 5, -5.9 will be passed as -5.

The function could be part of a .dll or .so created by the EUPHORIA To C Translator. In this case, a EUPHORIA atom or sequence could be returned. C and machine code functions can only return integers, or more generally, atoms (IEEE floating-point numbers).

13.1.4.5.5 Example 1:

```
atom user32, hwnd, ps, hdc
integer BeginPaint

-- open user32.dll - it contains the BeginPaint C function
user32 = open_dll("user32.dll")

-- the C function BeginPaint takes a C int argument and
-- a C pointer, and returns a C int as a result:
BeginPaint = define_c_func(user32, "BeginPaint",
                           {C_INT, C_POINTER}, C_INT)

-- call BeginPaint, passing hwnd and ps as the arguments,
-- hdc is assigned the result:
hdc = c_func(BeginPaint, {hwnd, ps})
```

13.1.4.5.6 See Also:

[c_proc](#), [define_c_proc](#), [open_dll](#), [Platform-Specific Issues](#)

13.1.4.6 c_proc

```
<built-in> procedure c_proc(integer rid, sequence args={})
```

Call a C void function, or machine code function, or translated/compiled EUPHORIA procedure by routine id.

13.1.4.6.1 Parameters:

1. `rid`: an integer, the `routine_id` of the external function being called.
2. `args`: a sequence, the list of parameters to pass to the function

13.1.4.6.2 Errors:

If `rid` is not a valid routine id, or the arguments do not match the prototype of the routine being called, an error occurs.

13.1.4.6.3 Comments:

`rid` must have been returned by `define_c_proc()`, **not** by `routine_id()`. The type checks are different, and you would get a machine level exception in the best case.

If the procedure does not take any arguments then `args` should be `{}`.

If you pass an argument value which contains a fractional part, where the C void function expects a C integer type, the argument will be rounded towards 0. e.g. 5.9 will be passed as 5, -5.9 will be passed as -5.

13.1.4.6.4 Example 1:

```
atom user32, hwnd, rect
integer GetClientRect

-- open user32.dll - it contains the GetClientRect C function
user32 = open_dll("user32.dll")

-- GetClientRect is a VOID C function that takes a C int
-- and a C pointer as its arguments:
GetClientRect = define_c_proc(user32, "GetClientRect",
                             {C_INT, C_POINTER})

-- pass hwnd and rect as the arguments
c_proc(GetClientRect, {hwnd, rect})
```

13.1.4.6.5 See Also:

[c_func](#), [define_c_func](#), [open_dll](#), [Platform-Specific Issues](#)

13.1.4.7 call_back

```
include std/dll.e
public function call_back(object id)
```

Get a machine address for an EUPHORIA procedure.

13.1.4.7.1 Parameters:

1. `id`: an object, either the id returned by [routine_id](#) for the function/procedure, or a pair {'+', id}.

13.1.4.7.2 Returns:

An **atom**, the address of the machine code of the routine. It can be used by Windows, or an external C routine in a Windows .dll or Unix-like shared library (.so), as a 32-bit "call-back" address for calling your EUPHORIA routine.

13.1.4.7.3 Errors:

The length of `name` should not exceed 1,024 characters.

13.1.4.7.4 Comments:

By default, your routine will work with the `stdcall` convention. On Windows, you can specify its id as {'+', id}, in which case it will work with the `cdecl` calling convention instead. On non-Microsoft platforms, you should only use simple IDs, as there is just one standard calling convention, i.e. `cdecl`.

You can set up as many call-back functions as you like, but they must all be EUPHORIA functions (or types) with 0 to 9 arguments. If your routine has nothing to return (it should really be a procedure), just return 0 (say), and the calling C routine can ignore the result.

When your routine is called, the argument values will all be 32-bit unsigned (positive) values. You should declare each parameter of your routine as `atom`, unless you want to impose tighter checking. Your routine must return a 32-bit integer value.

You can also use a call-back address to specify a EUPHORIA routine as an exception handler in the Linux/FreeBSD `signal()` function. For example, you might want to catch the `SIGTERM` signal, and do a graceful shutdown. Some Web hosts send a `SIGTERM` to a CGI process that has used too much CPU time.

A call-back routine that uses the `cdecl` convention and returns a floating-point result, might not work with `euiw`. This is because the Watcom C compiler (used to build `euiw`) has a non-standard way of handling `cdecl` floating-point return values.

13.1.4.7.5 Example 1:

See: `demo\win32\window.exw`, `demo\linux\qsort.ex`

13.1.4.7.6 See Also:

[routine_id](#)

13.2 Errors and Warnings

Routines

[crash](#)
[crash_message](#)
[crash_file](#)
[abort](#)
[warning_file](#)
[warning](#)
[crash_routine](#)

13.2.1 Routines

13.2.1.1 crash

```
include std/error.e
public procedure crash(sequence fmt, object data = {})
```

Crash running program, displaying a formatted error message the way `printf()` does.

13.2.1.1.1 Parameters:

1. `fmt` : a sequence representing the message text. It may have format specifiers in it
2. `data` : an object, defaulted to `{}`.

13.2.1.1.2 Comments:

The actual message being shown, both on standard error and in `ex.err` (or whatever file last passed to [crash_file\(\)](#)), is `sprintf(fmt, data)`. The program terminates as for any runtime error.

13.2.1.1.3 Example 1:

```
if PI = 3 then
    crash("The whole structure of universe just changed - please reload solar_system.ex")
end if
```

13.2.1.1.4 Example 2:

```
if token = end_of_file then
    crash("Test file #%d is bad, text read so far is %s\n", {file_number, read_so_far})
end if
```

13.2.1.1.5 See Also:

[crash_file](#), [crash_message](#), [printf](#)

13.2.1.2 crash_message

```
include std/error.e
public procedure crash_message(sequence msg)
```

Specify a final message to display for your user, in the event that EUPHORIA has to shut down your program due to an error.

13.2.1.2.1 Parameters:

1. msg : a sequence to display. It must only contain printable characters.

13.2.1.2.2 Comments:

There can be as many calls to `crash_message()` as needed in a program. Whatever was defined last will be used in case of a runtime error.

13.2.1.2.3 Example 1:

```
crash_message("The password you entered must have at least 8 characters.")
pwd_key = input_text[1..8]
-- if ##input_text## is too short, user will get a more meaningful message than
-- "index out of bounds".
```

13.2.1.2.4 See Also:

[crash](#), [crash_file](#)

13.2.1.3 crash_file

```
include std/error.e
public procedure crash_file(sequence file_path)
```

Specify a file path name in place of "ex.err" where you want any diagnostic information to be written.

13.2.1.3.1 Parameters:

1. `file_path` : a sequence, the new error and traceback file path.

13.2.1.3.2 Comments:

There can be as many calls to `crash_file()` as needed. Whatever was defined last will be used in case of an error at runtime, whether it was triggered by `crash()` or not.

13.2.1.3.3 See Also:

[crash](#), [crash_message](#)

13.2.1.4 abort

```
<built-in> procedure abort(atom error)
```

Abort execution of the program.

13.2.1.4.1 Parameters:

1. `error` : an integer, the exit code to return.

13.2.1.4.2 Comments:

`error` is expected to lie in the 0..255 range. 0 is usually interpreted as the sign of a successful completion.

Other values can indicate various kinds of errors. Windows batch (.bat) programs can read this value using the `errorlevel` feature. Non integer values are rounded down. A EUPHORIA program can read this value using [system_exec\(\)](#).

`abort()` is useful when a program is many levels deep in subroutine calls, and execution must end immediately, perhaps due to a severe error that has been detected.

If you don't use `abort()`, the interpreter will normally return an exit status code of 0. If your program fails with a EUPHORIA-detected compile-time or run-time error then a code of 1 is returned.

13.2.1.4.3 Example 1:

```
if x = 0 then
    puts(ERR, "can't divide by 0 !!!\n")
    abort(1)
else
    z = y / x
end if
```

13.2.1.4.4 See Also:

[crash_message](#), [system_exec](#)

13.2.1.5 warning_file

```
include std/error.e
public procedure warning_file(object file_path)
```

Specify a file path where to output warnings.

13.2.1.5.1 Parameters:

1. `file_path` : an object indicating where to dump any warning that were produced.

13.2.1.5.2 Comments:

By default, warnings are displayed on the standard error, and require pressing the Enter key to keep going. Redirecting to a file enables skipping the latter step and having a console window open, while retaining ability to inspect the warnings in case any was issued.

Any atom ≥ 0 causes standard error to be used, thus reverting to default behaviour.

Any atom < 0 suppresses both warning generation and output. Use this latter in extreme cases only.

On an error, some output to the console is performed anyway, so that whatever warning file was specified is ignored then.

13.2.1.5.3 Example 1:

```
warning_file("warnings.lst")
-- some code
warning_file(0)
-- changed opinion: warnings will go to standard error as usual
```

13.2.1.5.4 See Also:

[without warning](#), [warning](#)

13.2.1.6 warning

```
<built-in> procedure warning(sequence message)
```

Causes the specified warning message to be displayed as a regular warning.

13.2.1.6.1 Parameters:

1. `message` : a double quoted literal string, the text to display.

13.2.1.6.2 Comments:

Writing a library has specific requirements, since the code you write will be mainly used inside code you didn't write. It may be desirable then to influence, from inside the library, that code you didn't write.

This is what `warning()`, in a limited way, does. It enables to generate custom warnings in code that will include yours. Of course, you can also generate warnings in your own code, for instance as a kind of memo. The [without warning](#) top level statement disables such warnings.

The warning is issued with the `custom_warning` level. This level is enabled by default, but can be turned off any time.

Using any kind of expression in `message` will result in a blank warning text.

13.2.1.6.3 Example 1:

```
-- mylib.e
procedure foo(integer n)
    warning("The foo() procedure is obsolete, use bar() instead.")
    ? n
end procedure

-- some_app.exw
include mylib.e
foo(123)
```

will result, when `some_app.exw` is run with `warning`, in the following text being displayed in the console window

```
123
Warning: ( custom_warning ):
The foo() procedure is obsolete, use bar() instead.
```

Press Enter...

13.2.1.6.4 See Also:

[warning_file](#)

13.2.1.7 crash_routine

```
include std/error.e
public procedure crash_routine(integer func)
```

Specify a function to be called when an error takes place at run time.

13.2.1.7.1 Parameters:

1. `func` : an integer, the `routine_id` of the function to link in.

13.2.1.7.2 Comments:

The supplied function must have only one parameter, which should be integer or more general. Defaulted parameters in crash routines are not supported yet.

EUPHORIA maintains a linked list of routines to execute upon a crash. `crash_routine()` adds a new function to the list. The routines defined first are executed last. You cannot unlink a routine once it is linked, nor inspect the crash routine chain.

Currently, the crash routines are passed 0. Future versions may attempt to convey more information to them. If a crash routine returns anything else than 0, the remaining routines in the chain are skipped.

crash routines are not full fledged exception handlers, and they cannot resume execution at current or next statement. However, they can read the generated crash file, and might perform any action, including restarting the program.

13.2.1.7.3 Example 1:

```
function report_error(integer dummy)
  mylib:email("maintainer@remote_site.org", "ex.err")
  return 0 and dummy
end function
crash_routine(routine_id("report_error"))
```

13.2.1.7.4 See Also:

[crash_file](#), [routine_id](#), [Debugging and profiling](#)

13.3 Pseudo Memory

One use is to emulate PBR, such as EUPHORIA's map and stack types.

```
ram_space
malloc
free
valid
allocate_pointer_array
free_pointer_array
allocate_string_pointer_array
```

13.3.1 ram_space

```
include std/eumem.e
export sequence ram_space
```

The (pseudo) RAM heap space. Use [malloc](#) to gain ownership to a heap location and [free](#) to release it back to the system.

13.3.1.1 malloc

```
include std/eumem.e
export function malloc(object mem_struct_p = 1, integer cleanup_p = 1)
```

Allocate a block of (pseudo) memory

13.3.1.1.1 Parameters:

1. `mem_struct_p` : The initial structure (sequence) to occupy the allocated block. If this is an integer, a sequence of zero this long is used. The default is the number 1, meaning that the default initial structure is {0}
2. `cleanup_p` : Identifies whether the memory should be released automatically when the reference count for the handle for the allocated block drops to zero, or when passed to `delete()`. If 0, then the block must be freed using the [free](#) procedure.

13.3.1.1.2 Returns:

A **handle**, to the acquired block. Once you acquire this, you can use it as you need to. Note that if `cleanup_p` is 1, then the variable holding the handle must be capable of storing an atom as a double floating point value (i.e., not an integer).

13.3.1.1.3 Example 1:

```
my_spot = malloc()  
ram_space[my_spot] = my_data
```

13.3.1.2 free

```
include std/eumem.e  
export procedure free(atom mem_p)
```

Deallocate a block of (pseudo) memory

13.3.1.2.1 Parameters:

1. `mem_p` : The handle to a previously acquired [ram_space](#) location.

13.3.1.2.2 Comments:

This allows the location to be used by other parts of your application. You should no longer access this location again because it could be acquired by some other process in your application. This routine should only be called if you passed 0 as `cleanup_p` to [malloc](#).

13.3.1.2.3 Example 1:

```
my_spot = malloc(1,0)  
ram_space[my_spot] = my_data  
-- . . . do some processing . . .  
free(my_spot)
```

13.3.1.3 valid

```
include std/eumem.e  
export function valid(object mem_p, object mem_struct_p = 1)
```

Validates a block of (pseudo) memory

13.3.1.3.1 Parameters:

1. `mem_p` : The handle to a previously acquired `ram_space` location.
2. `mem_struct_p` : If an integer, this is the length of the sequence that should be occupying the `ram_space` location pointed to by `mem_p`.

13.3.1.3.2 Returns:

An integer,

0 if either the `mem_p` is invalid or if the sequence at that location is the wrong length.

1 if the handle and contents is okay.

13.3.1.3.3 Comments:

This can only check the length of the contents at the location. Nothing else is checked at that location.

13.3.1.3.4 Example 1:

```
my_spot = malloc()
ram_space[my_spot] = my_data
. . . do some processing . .
if valid(my_spot, length(my_data)) then
    free(my_spot)
end if
```

13.3.1.4 allocate_pointer_array

```
include std/machine.e
public function allocate_pointer_array(sequence pointers, boolean cleanup = 0)
```

Allocate a NULL terminated pointer array.

13.3.1.4.1 Parameters:

1. `pointers` : a sequence of pointers to add to the pointer array.
2. `cleanup` : an integer, if non-zero, then the returned pointer will be automatically freed when its reference count drops to zero, or when passed as a parameter to `delete`

13.3.1.4.2 Comments:

This function adds the NULL terminator.

13.3.1.4.3 Example 1:

```
atom pa = allocate_pointer_array({ allocate_string("1"), allocate_string("2") })
```

13.3.1.4.4 See Also:

[allocate_string_pointer_array](#), [free_pointer_array](#)

13.3.1.5 free_pointer_array

```
include std/machine.e
public procedure free_pointer_array(atom pointers_array)
```

Free a NULL terminated pointers array.

13.3.1.5.1 Parameters:

1. `pointers_array` : memory address of where the NULL terminated array exists at.

13.3.1.5.2 Comments:

This is for NULL terminated lists, such as allocated by [allocate_pointer_array](#). Do not call `free_pointer_array()` for a pointer that was allocated to be cleaned up automatically. Instead, use [delete](#).

13.3.1.5.3 See Also:

[allocate_pointer_array](#), [allocate_string_pointer_array](#)

13.3.1.6 allocate_string_pointer_array

```
include std/machine.e
public function allocate_string_pointer_array(object string_list, boolean cleanup = 0)
```

Allocate a C-style null-terminated array of strings in memory

13.3.1.6.1 Parameters:

1. `string_list` : sequence of strings to store in RAM.
2. `cleanup` : an integer, if non-zero, then the returned pointer will be automatically freed when its reference count drops to zero, or when passed as a parameter to [delete](#)

13.3.1.6.2 Returns:

An **atom**, the address of the memory block where the string pointer array was stored.

13.3.1.6.3 Example 1:

```
atom p = allocate_string_pointer_array({ "One", "Two", "Three" })  
-- Same as C: char *p = { "One", "Two", "Three", NULL };
```

13.3.1.6.4 See Also:

[free_pointer_array](#)

13.4 Indirect Routine Calling

- [SAFE mode](#)
- [Data Execute mode](#)
- [Accessing EUPHORIA coded routines](#)
 - [routine_id](#)
 - [call_func](#)
 - [call_proc](#)
- [Accessing EUPHORIA internals](#)
 - [machine_func](#)
 - [machine_proc](#)
 - [PAGE_SIZE](#)

13.4.1 SAFE mode

During the development of your application, you can define the word `SAFE` to cause `machine.e` to use alternative memory functions. These functions are slower but help in the debugging stages. In general, `SAFE` mode should not be enabled during production phases but only for development phases.

To define the word `SAFE` run your application with the `-D SAFE` command line option, or add to the top of your main file with `define SAFE`.

13.4.2 Data Execute mode

`Data Execute mode` makes data that will be returned from `allocate()` executable. On some systems `allocate()` will return memory that is not executable unless this mode has been enabled. When writing software you should use `allocate_code()` or `allocate_protect()` to get memory for execution. This is more efficient and more secure than using `Data Execute mode`. However, since on many systems executing memory returned from `allocate()` will work much software will be written 4.0 and yet use `allocate()` for executable memory instead of the afore mentioned routines. Therefore, you may use this switch when you find that your are getting `Data Execute Exceptions` running some software. `SAFE` mode will help you discover what

memory should be changed to what access level. `Data Execute` mode will only work when the EUPHORIA program uses `std/machine.e` not `machine.e`.

13.4.3 Accessing EUPHORIA coded routines

13.4.3.1 routine_id

<built-in> `function routine_id(sequence routine_name)`

Return an integer id number for a user-defined EUPHORIA procedure or function.

13.4.3.1.1 Parameters:

1. `routine_name` : a string, the name of the procedure or function.

13.4.3.1.2 Returns:

An **integer**, known as a routine id, -1 if the named routine can't be found, else zero or more.

13.4.3.1.3 Errors:

`routine_name` should not exceed 1,024 characters.

13.4.3.1.4 Comments:

The id number can be passed to `call_proc()` or `call_func()`, to indirectly call the routine named by `routine_name`. This id depends on the internal process of parsing your code, not on `routine_name`.

The routine named `routine_name` must be visible, i.e. callable, at the place where `routine_id()` is used to get the id number. If it is not, -1 is returned.

Indirect calls to the routine can appear earlier in the program than the definition of the routine, but the id number can only be obtained in code that comes after the definition of the routine - see example 2 below.

Once obtained, a valid routine id can be used at any place in the program to call a routine indirectly via `call_proc()/call_func()`, including at places where the routine is no longer in scope.

Some typical uses of `routine_id()` are:

1. Creating a subroutine that takes another routine as a parameter. (See Example 2 below)

2. Using a sequence of routine id's to make a case (switch) statement. Using the [switch statement](#) is more efficient.
3. Setting up an Object-Oriented system.
4. Getting a routine id so you can pass it to [call_back\(\)](#). (See [Platform-Specific Issues](#))
5. Getting a routine id so you can pass it to [task_create\(\)](#). (See [Multitasking in EUPHORIA](#))
6. Calling a routine that is defined later in a program. This is no longer needed from v4.0 onward.

Note that C routines, callable by EUPHORIA, also have ids, but they cannot be used where routine ids are, because of the different type checking and other technical issues. See [define_c_proc\(\)](#) and [define_c_func\(\)](#).

13.4.3.1.5 Example 1:

```
procedure foo()  
  puts(1, "Hello World\n")  
end procedure  
  
integer foo_num  
foo_num = routine_id("foo")  
  
call_proc(foo_num, {}) -- same as calling foo()
```

13.4.3.1.6 Example 2:

```
function apply_to_all(sequence s, integer f)  
  -- apply a function to all elements of a sequence  
  sequence result  
  result = {}  
  for i = 1 to length(s) do  
    -- we can call add1() here although it comes later in the program  
    result = append(result, call_func(f, {s[i]}))  
  end for  
  return result  
end function  
  
function add1(atom x)  
  return x + 1  
end function  
  
-- add1() is visible here, so we can ask for its routine id  
? apply_to_all({1, 2, 3}, routine_id("add1"))  
-- displays {2,3,4}
```

13.4.3.1.7 See Also:

[call_proc](#), [call_func](#), [call_back](#), [define_c_func](#), [define_c_proc](#), [task_create](#), [Platform-Specific Issues](#), [Indirect routine calling](#)

13.4.3.2 call_func

```
<built-in> function call_func(integer id, sequence args={})
```

Call the user-defined EUPHORIA function by routine id.

13.4.3.2.1 Parameters:

1. `id`: an integer, the routine id of the function to call
2. `args`: a sequence, the parameters to pass to the function.

13.4.3.2.2 Returns:

The **value**, the called function returns.

13.4.3.2.3 Errors:

If `id` is negative or otherwise unknown, an error occurs.

If the length of `args` is not the number of parameters the function takes, an error occurs.

13.4.3.2.4 Comments:

`id` must be a valid routine id returned by `routine_id()`.

`args` must be a sequence of argument values of length `n`, where `n` is the number of arguments required by the called function. Defaulted parameters currently cannot be synthesized while making an indirect call.

If the function with id `id` does not take any arguments then `args` should be `{}`.

13.4.3.2.5 Example 1:

Take a look at the sample program called `demo/csort.ex`

13.4.3.2.6 See Also:

[call_proc](#), [routine_id](#), [c_func](#)

13.4.3.3 call_proc

```
<built-in> procedure call_proc(integer id, sequence args={})
```

Call a user-defined EUPHORIA procedure by routine id.

13.4.3.3.1 Parameters:

1. `id` : an integer, the routine id of the procedure to call
2. `args` : a sequence, the parameters to pass to the function.

13.4.3.3.2 Errors:

If `id` is negative or otherwise unknown, an error occurs.

If the length of `args` is not the number of parameters the function takes, an error occurs.

13.4.3.3.3 Comments:

`id` must be a valid routine id returned by `routine_id()`.

`args` must be a sequence of argument values of length `n`, where `n` is the number of arguments required by the called procedure. Defaulted parameters currently cannot be synthesized while making an indirect call.

If the procedure with id `id` does not take any arguments then `args` should be `{ }`.

13.4.3.3.4 Example 1:

```
public integer foo_id

procedure x()
    call_proc(foo_id, {1, "Hello World\n"})
end procedure

procedure foo(integer a, sequence s)
    puts(a, s)
end procedure

foo_id = routine_id("foo")

x()
```

13.4.3.3.5 See Also:

`call_func`, `routine_id`, `c_proc`

13.4.4 Accessing EUPHORIA internals

13.4.4.1 machine_func

```
<built-in> function machine_func(integer machine_id, object args={})
```

Perform a machine-specific operation that returns a value.

13.4.4.1.1 Returns:

Depends on the called internal facility.

13.4.4.1.2 Comments:

This function is mainly used by the standard library files to implement machine dependent operations such as graphics and sound effects. This routine should normally be called indirectly via one of the library routines in a EUPHORIA include file. User programs normally do not need to call `machine_func`.

A direct call might cause a machine exception if done incorrectly.

13.4.4.1.3 See Also:

[machine_proc](#)

13.4.4.2 machine_proc

```
<built-in> procedure machine_proc(integer machine_id, object args={})
```

Perform a machine-specific operation that does not return a value.

13.4.4.2.1 Comments:

This procedure is mainly used by the standard library files to implement machine dependent operations such as graphics and sound effects. This routine should normally be called indirectly via one of the library routines in a EUPHORIA include file. User programs normally do not need to call `machine_proc`.

A direct call might cause a machine exception if done incorrectly.

13.4.4.2.2 See Also:

[machine_func](#)

13.4.4.3 PAGE_SIZE

```
include std/machine.e
public constant PAGE_SIZE
```

13.5 Types supporting Memory

13.5.1 valid_memory_protection_constant

```
include std/machine.e
public type valid_memory_protection_constant(integer x)
```

protection constants type

13.5.1.1 page_aligned_address

```
include std/machine.e
public type page_aligned_address(atom a)
```

page aligned address type

13.5.1.2 is_DEP_supported

```
include std/machine.e
public function is_DEP_supported()
```

13.5.1.3 is_using_DEP

```
include std/machine.e
public function is_using_DEP()
```

13.5.1.4 DEP_on

```
include std/machine.e
public procedure DEP_on(integer value)
```


13.5.2 Allocating and Writing to memory:

13.5.2.1 `allocate_code`

```
include std/machine.e
public function allocate_code(object data, valid_wordsize wordsize = 1)
```

Allocates and copies data into executable memory.

13.5.2.1.1 Parameters:

1. `a_sequence_of_machine_code` : is the machine code to be put into memory to be later called with `call()`
2. the word `length` : of the said code. You can specify your code as 1-byte, 2-byte or 4-byte chunks if you wish. If your machine code is byte code specify 1. The default is 1.

13.5.2.1.2 Return Value:

An **address**, The function returns the address in memory of the code, that can be safely executed whether DEP is enabled or not or 0 if it fails. On the other hand, if you try to execute a code address returned by `allocate()` with DEP enabled the program will receive a machine exception.

13.5.2.1.3 Comments:

Use this for the machine code you want to run in memory. The copying is done for you and when the routine returns the memory may not be readable or writeable but it is guaranteed to be executable. If you want to also write to this memory **after the machine code has been copied** you should use `allocate_protect()` instead and you should read about having memory executable and writeable at the same time is a bad idea. You mustn't use `free()` on memory returned from this function. You may instead use `free_code()` but since you will probably need the code throughout the life of your program's process this normally is not necessary. If you want to put only data in the memory to be read and written use `allocate`.

13.5.2.1.4 See Also:

`allocate`, `free_code`, `allocate_protect`

13.5.2.2 `std_library_address`

```
include std/machine.e
public type std_library_address(atom addr)
```

Type for memory addresses

an address returned from `allocate()` or `allocate_protect()` or `allocate_code()` or the value 0.

13.5.2.2.1 Return Value:

An **integer**, The type will return 1 if the parameter was returned from one of these functions (and has not yet been freed)

13.5.2.2.2 Comments:

This type is equivalent to `atom` unless `SAFE` is defined. Only values that satisfy this type may be passed into `free` or `free_code`.

13.5.2.3 `allocate_protect`

```
include std/machine.e
public function allocate_protect(object data, valid_wordsize wordsize = 1, valid_memory_protect
```

Allocates and copies data into memory and gives it protection using [Microsoft's Memory Protection Constants](#). The user may only pass in one of these constants. If you only wish to execute a sequence as machine code use `allocate_code()`. If you only want to read and write data into memory use `allocate()`.

See [MSDN: Microsoft's Memory Protection Constants](#)

13.5.2.3.1 Parameters:

1. `data` : is the machine code to be put into memory.
2. `wordsize` : is the size each element of data will take in memory. Are they 1-byte, 2-bytes or 4-bytes long? Specify here. The default is 1.
3. `protection` : is the particular Windows protection.

13.5.2.3.2 Returns:

An **address**, The function returns the address to the required memory or 0 if it fails. This function is guaranteed to return memory on the 4 byte boundary. It also guarantees that the memory returned with at least the protection given (but you may get more).

If you want to call `allocate_protect(data, PAGE_READWRITE)`, you can use [allocate](#) instead. It is more efficient and simpler.

If you want to call `allocate_protect(data, PAGE_EXECUTE)`, you can use [allocate_code\(\)](#) instead. It is simpler.

You mustn't use [free\(\)](#) on memory returned from this function, instead use [free_code\(\)](#).

13.5.2.4 poke_string

```
include std/machine.e
public function poke_string(atom buffaddr, integer buffsize, sequence s)
```

Stores a C-style null-terminated ANSI string in memory

13.5.2.4.1 Parameters:

1. buffaddr: an atom, the RAM address to to the string at.
2. buffsize: an integer, the number of bytes available, starting from buffaddr.
3. s : a sequence, the string to store at address buffaddr.

13.5.2.4.2 Comments:

- This does not allocate an RAM. You must supply the preallocated area.
- This can only be used on ANSI strings. It cannot be used for double-byte strings.
- If s is not a string, nothing is stored and a zero is returned.

13.5.2.4.3 Returns:

An atom. If this is zero, then nothing was stored, otherwise it is the address of the first byte after the stored string.

13.5.2.4.4 Example 1:

```
atom title

title = allocate(1000)
if poke_string(title, 1000, "The Wizard of Oz") then
    -- successful
else
    -- failed
end if
```

13.5.2.4.5 See Also:

[allocate](#), [allocate_string](#)

13.5.2.5 poke_wstring

```
include std/machine.e
public function poke_wstring(atom buffaddr, integer buffsize, sequence s)
```

Stores a C-style null-terminated Double-Byte string in memory

13.5.2.5.1 Parameters:

1. `buffaddr`: an atom, the RAM address to to the string at.
2. `buffsize`: an integer, the number of bytes available, starting from `buffaddr`.
3. `s`: a sequence, the string to store at address `buffaddr`.

13.5.2.5.2 Comments:

- This does not allocate an RAM. You must supply the preallocated area.
- This uses two bytes per string character. **Note** that `buffsize` is the number of *bytes* available in the buffer and not the number of *characters* available.
- If `s` is not a double-byte string, nothing is stored and a zero is returned.

13.5.2.5.3 Returns:

An atom. If this is zero, then nothing was stored, otherwise it is the address of the first byte after the stored string.

13.5.2.5.4 Example 1:

```
atom title

title = allocate(1000)
if poke_wstring(title, 1000, "The Wizard of Oz") then
    -- successful
else
    -- failed
end if
```

13.5.2.5.5 See Also:

[allocate](#), [allocate_wstring](#)

13.5.2.6 `allocate_string`

```
include std/machine.e
public function allocate_string(sequence s, boolean cleanup = 0)
```

Allocate a C-style null-terminated string in memory

13.5.2.6.1 Parameters:

1. `s`: a sequence, the string to store in RAM.
2. `cleanup`: an integer, if non-zero, then the returned pointer will be automatically freed when its

reference count drops to zero, or when passed as a parameter to [delete](#).

13.5.2.6.2 Returns:

An **atom**, the address of the memory block where the string was stored, or 0 on failure.

13.5.2.6.3 Comments:

Only the 8 lowest bits of each atom in `s` is stored. Use `allocate_wstring()` for storing double byte encoded strings.

There is no `allocate_string_low()` function. However, you could easily craft one by adapting the code for `allocate_string`.

Since `allocate_string()` allocates memory, you are responsible to [free\(\)](#) the block when done with it if `cleanup` is zero. If `cleanup` is non-zero, then the memory can be freed by calling [delete](#), or when the pointer's reference count drops to zero.

13.5.2.6.4 Example 1:

```
atom title
title = allocate_string("The Wizard of Oz")
```

13.5.2.6.5 See Also:

[allocate](#), [allocate_wstring](#)

13.5.2.7 `allocate_wstring`

```
include std/machine.e
public function allocate_wstring(sequence s, boolean cleanup = 0)
```

Create a C-style null-terminated `wchar_t` string in memory

13.5.2.7.1 Parameters:

1. `s` : a unicode (utf16) string

13.5.2.7.2 Returns:

An **atom**, the address of the allocated string, or 0 on failure.

13.5.2.7.3 See Also:

[allocate_string](#)

13.5.2.8 peek_wstring

```
include std/machine.e
public function peek_wstring(atom addr)
```

Return a unicode (utf16) string that are stored at machine address a.

13.5.2.8.1 Parameters:

1. `addr` : an atom, the address of the string in memory

13.5.2.8.2 Returns:

The **string**, at the memory position. The terminator is the null word (two bytes equal to 0).

13.5.2.8.3 See Also:

[peek_string](#)

13.5.3 Memory disposal

13.5.3.1 free

```
include std/machine.e
public procedure free(object addr)
```

Free up a previously allocated block of memory. machine:free

13.5.3.1.1 Parameters:

1. `addr`, either a single atom or a sequence of atoms; these are addresses of a blocks to free.

13.5.3.1.2 Comments:

- Use `free()` to return blocks of memory the during execution. This will reduce the chance of running out of memory or getting into excessive virtual memory swapping to disk.
- Do not reference a block of memory that has been freed.
- When your program terminates, all allocated memory will be returned to the system.
- `addr` must have been allocated previously using `allocate()`. You cannot use it to relinquish part of a block. Instead, you have to allocate a block of the new size, copy useful contents from old block there and then `free()` the old block.
- If the memory was allocated and automatic cleanup was specified, then do not call `free()` directly. Instead, use `delete`.
- An `addr` of zero is simply ignored.

13.5.3.1.3 Example 1:

`demo/callmach.ex`

13.5.3.1.4 See Also:

[allocate](#), [free_code](#)

13.5.3.2 free_code

```
include std/machine.e
public procedure free_code( atom addr, integer size, valid_wordsize wordsize = 1 )
```

Frees up allocated code memory

13.5.3.2.1 Parameters:

1. `addr` : must be an address returned by [allocate_code\(\)](#) or [allocate_protect\(\)](#). Do **not** pass memory returned from [allocate\(\)](#) here!
2. `size` : is the length of the sequence passed to `allocate_code()` or the size you specified when you called `allocate_protect()`.
3. `wordsize`: `valid_wordsize` default = 1

13.5.3.2.2 Comments:

Chances are you will not need to call this function because code allocations are typically public scope operations that you want to have available until your process exits.

See Also: [allocate_code](#), [free](#)

13.6 Memory Management - Low-Level

Usage Notes

- edges_only
- positive_int
- machine_addr

Memory allocation

- allocate
- allocate_data
- deallocate

Reading from, Writing to, and Calling into Memory

- peek
- peek2s
- peek2u
- peek4s
- peek4u
- peek_string
- poke
- poke2
- poke4
- mem_copy
- mem_set
- call
- check_calls

Safe memory access

- register_block
- unregister_block
- safe_address
- check_all_blocks
- prepare_block
- BORDER_SPACE
- trailer
- bordered_address
- delete_routine
- delete
- dep_works
- VirtualFree_rid
- free_code

safe.e

- check_calls
- edges_only
- safe_address_list
- machine_addr
- BORDER_SPACE
- trailer
- bordered_address
- safe_address

peek
peek2u
peek2s
peek4s
peek4u
peek_string
poke
poke2
poke4
mem_copy
mem_set
show_block
check_all_blocks
call
c_proc
c_func
register_block
unregister_block
prepare_block
allocate_data
allocate
deallocate
dep_works
VirtualFree_rid
free_code

Microsoft Windows Memory Protection Constants

PAGE_EXECUTE
PAGE_EXECUTE_READ
PAGE_EXECUTE_READWRITE
PAGE_EXECUTE_WRITECOPY
PAGE_WRITECOPY
PAGE_READWRITE
PAGE_READONLY
PAGE_NOACCESS

Standard Library Memory Protection Constants

PAGE_NONE
PAGE_READ_EXECUTE
PAGE_READ_WRITE
PAGE_READ
PAGE_READ_WRITE_EXECUTE
PAGE_WRITE_EXECUTE_COPY
PAGE_WRITE_COPY
valid_memory_protection_constant
test_read
test_write
test_exec
valid_wordsize
DEP_really_works



MEM_COMMIT
MEM_RESERVE
MEM_RESET
MEM_RELEASE
FREE_RID
A_WRITE
A_EXECUTE
M_ALLOC
M_FREE
kernel_dll
memDLL_id
VirtualAlloc_rid
VirtualLock_rid
VirtualUnlock_rid
VirtualProtect_rid
GetLastError_rid
GetSystemInfo_rid
PROT_EXEC
PROT_READ
PROT_WRITE
PROT_NONE
MAP_ANONYMOUS
MAP_PRIVATE
MAP_SHARED
MAP_TYPE
MAP_FIXED
MAP_FILE
get_page_size
mmap
munmap
mlock
munlock
mprotect
is_valid_memory_protection_constant
is_page_aligned_address
Error Code Constants
BMP_SUCCESS
BMP_OPEN_FAILED
BMP_UNEXPECTED_EOF
BMP_UNSUPPORTED_FORMAT
BMP_INVALID_MODE
video_config sequence accessors
VC_COLOR
VC_MODE
VC_LINES
VC_COLUMNS
VC_XPIXELS
VC_YPIXELS
VC_NCOLORS

VC_PAGES
VC_SCRNLINES
VC_SCRNCOLS
Colors
BLACK
BLUE
GREEN
CYAN
RED
MAGENTA
BROWN
WHITE
GRAY
BRIGHT_BLUE
BRIGHT_GREEN
BRIGHT_CYAN
BRIGHT_RED
BRIGHT_MAGENTA
YELLOW
BRIGHT_WHITE
true_color
BLINKING
BYTES_PER_CHAR
color
Routines
mixture
video_config

13.6.1 Usage Notes

safe.e This file is not normally included directly. The normal approach is to include `std/machine.e`, which will automatically include either this file or `std/safe.e` if the `SAFE` symbol has been defined.

Warning: Some of these routines require a knowledge of machine-level programming. You could crash your system!

These routines, along with [peek\(\)](#), [poke\(\)](#) and [call\(\)](#), let you access all of the features of your computer. You can read and write to any memory location, and you can create and execute machine code subroutines.

If you are manipulating 32-bit addresses or values, remember to use variables declared as `atom`. The integer type only goes up to 31 bits.

Writing characters to screen memory with `poke()` is much faster than using `puts()`. Address of start of text screen memory:

- mono: #B0000
- color: #B8000

If you choose to call `machine_proc()` or `machine_func()` directly (to save a bit of overhead) you *must* pass valid arguments or EUPHORIA could crash.

13.6.1.1 Some example programs to look at:

- `demo/callmach.ex` -- calling a machine language routine

See also `include/safe.e`. It's a safe, debugging version of this file.

13.6.1.2 edges_only

```
include std/memory.e
public integer edges_only
```

13.6.1.3 positive_int

```
include std/memory.e
type positive_int(integer x)
```

Positive integer type

13.6.1.4 machine_addr

```
include std/memory.e
public type machine_addr(object a)
```

Machine address type

13.6.2 Memory allocation

13.6.2.1 allocate

```
include std/memory.e
public function allocate(positive_int n, integer cleanup = 0)
```

Allocate a contiguous block of data memory.

13.6.2.1.1 Parameters:

1. `n` : a positive integer, the size of the requested block.
2. `cleanup` : an integer, if non-zero, then the returned pointer will be automatically freed when its reference count drops to zero, or when passed as a parameter to [delete](#).

13.6.2.1.2 Return:

An **atom**, the address of the allocated memory or 0 if the memory can't be allocated.

13.6.2.1.3 Comments:

Since `allocate_string()` allocates memory, you are responsible to [free\(\)](#) the block when done with it if `cleanup` is zero. If `cleanup` is non-zero, then the memory can be freed by calling [delete](#), or when the pointer's reference count drops to zero. When you are finished using the block, you should pass the address of the block to [free\(\)](#) if `cleanup` is zero. If `cleanup` is non-zero, then the memory can be freed by calling [delete](#), or when the pointer's reference count drops to zero. This will free the block and make the memory available for other purposes. When your program terminates, the operating system will reclaim all memory for use with other programs. An address returned by this function shouldn't be passed to [call\(\)](#). For that purpose you may use [allocate_code\(\)](#) instead.

The address returned will be at least 4-byte aligned.

13.6.2.1.4 Example 1:

```
buffer = allocate(100)
for i = 0 to 99 do
    poke(buffer+i, 0)
end for
```

13.6.2.1.5 See Also:

[free](#), [peek](#), [poke](#), [mem_set](#), [allocate_code](#)

13.6.2.2 `allocate_data`

```
include std/memory.e
public function allocate_data(positive_int n, integer cleanup = 0)
```

Allocate `n` bytes of memory and return the address. Free the memory using `free()` below.

13.6.2.3 deallocate

```
include std/memory.e
export procedure deallocate(atom addr)
```

13.6.3 Reading from, Writing to, and Calling into Memory

13.6.3.1 peek

```
<built-in> function peek(object addr_n_length)
```

Fetches a byte, or some bytes, from an address in memory.

13.6.3.1.1 Parameters:

1. `addr_n_length` : an object, either of
 - ◆ an atom `addr` -- to fetch one byte at `addr`, or
 - ◆ a pair `{addr, len}` -- to fetch `len` bytes at `addr`

13.6.3.1.2 Returns:

An **object**, either an integer if the input was a single address, or a sequence of integers if a sequence was passed. In both cases, integers returned are bytes, in the range 0..255.

13.6.3.1.3 Errors:

[Peeking](#) in memory you don't own may be blocked by the OS, and cause a machine exception. If you use the `define safe` these routines will catch these problems with a EUPHORIA error.

When supplying a `{address, count}` sequence, the count must not be negative.

13.6.3.1.4 Comments:

Since addresses are 32-bit numbers, they can be larger than the largest value of type integer (31-bits). Variables that hold an address should therefore be declared as atoms.

It is faster to read several bytes at once using the second form of `peek()` than it is to read one byte at a time in a loop. The returned sequence has the length you asked for on input.

Remember that `peek()` takes just one argument, which in the second form is actually a 2-element sequence.

13.6.3.1.5 Example 1:

```
-- The following are equivalent:
-- method 1
s = {peek(100), peek(101), peek(102), peek(103)}

-- method 2
s = peek({100, 4})
```

13.6.3.1.6 See Also:

[poke](#), [peeks](#), [peek4u](#), [allocate](#), [free](#), [peek2u](#)

13.6.3.2 peeks

<built-in> `function peeks(object addr_n_length)`

Fetches a byte, or some bytes, from an address in memory.

13.6.3.2.1 Parameters:

1. `addr_n_length` : an object, either of
 - ◆ an atom `addr` : to fetch one byte at `addr`, or
 - ◆ a pair `{addr, len}` : to fetch `len` bytes at `addr`

13.6.3.2.2 Returns:

An **object**, either an integer if the input was a single address, or a sequence of integers if a sequence was passed. In both cases, integers returned are bytes, in the range -128..127.

13.6.3.2.3 Errors:

[Peeking](#) in memory you don't own may be blocked by the OS, and cause a machine exception. If you use the [define safe](#) these routines will catch these problems with a EUPHORIA error.

When supplying a `{address, count}` sequence, the count must not be negative.

13.6.3.2.4 Comments:

Since addresses are 32-bit numbers, they can be larger than the largest value of type integer (31-bits). Variables that hold an address should therefore be declared as atoms.

It is faster to read several bytes at once using the second form of `peek()` than it is to read one byte at a time in a loop. The returned sequence has the length you asked for on input.

Remember that `peeks()` takes just one argument, which in the second form is actually a 2-element sequence.

13.6.3.2.5 Example 1:

```
-- The following are equivalent:
-- method 1
s = {peek(100), peek(101), peek(102), peek(103)}

-- method 2
s = peeks({100, 4})
```

13.6.3.2.6 See Also:

[poke](#), [peek4s](#), [allocate](#), [free](#), [peek2s](#), [peek](#)

13.6.3.3 peek2s

```
<built-in> function peek2s(object addr_n_length)
```

Fetches a *signed* word, or some *signed* words, from an address in memory.

13.6.3.3.1 Parameters:

1. `addr_n_length`: an object, either of
 - ◆ an atom `addr` -- to fetch one word at `addr`, or
 - ◆ a pair { `addr`, `len` }, to fetch `len` words at `addr`

13.6.3.3.2 Returns:

An **object**, either an integer if the input was a single address, or a sequence of integers if a sequence was passed. In both cases, integers returned are double words, in the range -32768..32767.

13.6.3.3.3 Errors:

Peeking in memory you don't own may be blocked by the OS, and cause a machine exception. If you use the define safe these routines will catch these problems with a EUPHORIA error.

When supplying a {address, count} sequence, the count must not be negative.

13.6.3.3.4 Comments:

Since addresses are 32-bit numbers, they can be larger than the largest value of type integer (31-bits). Variables that hold an address should therefore be declared as atoms.

It is faster to read several words at once using the second form of `peek()` than it is to read one word at a time in a loop. The returned sequence has the length you asked for on input.

Remember that `peek2s()` takes just one argument, which in the second form is actually a 2-element sequence.

The only difference between `peek2s()` and `peek2u()` is how words with the highest bit set are returned. `peek2s()` assumes them to be negative, while `peek2u()` just assumes them to be large and positive.

13.6.3.3.5 Example 1:

```
-- The following are equivalent:
-- method 1
s = {peek2s(100), peek2s(102), peek2s(104), peek2s(106)}

-- method 2
s = peek2s({100, 4})
```

13.6.3.3.6 See Also:

[poke2](#), [peeks](#), [peek4s](#), [allocate](#), [free peek2u](#)

13.6.3.4 peek2u

```
<built-in> function peek2u(object addr_n_length)
```

Fetches an *unsigned* word, or some *unsigned* words, from an address in memory.

13.6.3.4.1 Parameters:

1. `addr_n_length`: an object, either of
 - ◆ an atom `addr` -- to fetch one double word at `addr`, or
 - ◆ a pair `{addr, len}` -- to fetch `len` double words at `addr`

13.6.3.4.2 Returns:

An **object**, either an integer if the input was a single address, or a sequence of integers if a sequence was passed. In both cases, integers returned are words, in the range 0..65535.

13.6.3.4.3 Errors:

`Peek()` in memory you don't own may be blocked by the OS, and cause a machine exception. If you use the `define safe` these routines will catch these problems with a EUPHORIA error.

When supplying a {address, count} sequence, the count must not be negative.

13.6.3.4.4 Comments:

Since addresses are 32-bit numbers, they can be larger than the largest value of type integer (31-bits). Variables that hold an address should therefore be declared as atoms.

It is faster to read several words at once using the second form of peek() than it is to read one word at a time in a loop. The returned sequence has the length you asked for on input.

Remember that peek2u() takes just one argument, which in the second form is actually a 2-element sequence.

The only difference between peek2s() and peek2u() is how words with the highest bit set are returned. peek2s() assumes them to be negative, while peek2u() just assumes them to be large and positive.

13.6.3.4.5 Example 1:

```
-- The following are equivalent:
-- method 1
Get 4 2-byte numbers starting address 100.
s = {peek2u(100), peek2u(102), peek2u(104), peek2u(106)}

-- method 2
Get 4 2-byte numbers starting address 100.
s = peek2u({100, 4})
```

13.6.3.4.6 See Also:

[poke2](#), [peek](#), [peek2s](#), [allocate](#), [free peek4u](#)

13.6.3.5 peek4s

```
<built-in> function peek4s(object addr_n_length)
```

Fetches a *signed* double words, or some *signed* double words, from an address in memory.

13.6.3.5.1 Parameters:

1. `addr_n_length`: an object, either of
 - ◆ an atom `addr` -- to fetch one double word at `addr`, or
 - ◆ a pair { `addr`, `len` } -- to fetch `len` double words at `addr`

13.6.3.5.2 Returns:

An **object**, either an atom if the input was a single address, or a sequence of atoms if a sequence was passed. In both cases, atoms returned are double words, in the range $-\text{power}(2,31)..\text{power}(2,31)-1$.

13.6.3.5.3 Errors:

Peeking in memory you don't own may be blocked by the OS, and cause a machine exception. If you use the define safe these routines will catch these problems with a EUPHORIA error.

When supplying a {address, count} sequence, the count must not be negative.

13.6.3.5.4 Comments:

Since addresses are 32-bit numbers, they can be larger than the largest value of type integer (31-bits). Variables that hold an address should therefore be declared as atoms.

It is faster to read several double words at once using the second form of `peek()` than it is to read one double word at a time in a loop. The returned sequence has the length you asked for on input.

Remember that `peek4s()` takes just one argument, which in the second form is actually a 2-element sequence.

The only difference between `peek4s()` and `peek4u()` is how double words with the highest bit set are returned. `peek4s()` assumes them to be negative, while `peek4u()` just assumes them to be large and positive.

13.6.3.5.5 Example 1:

```
-- The following are equivalent:
-- method 1
s = {peek4s(100), peek4s(104), peek4s(108), peek4s(112)}

-- method 2
s = peek4s({100, 4})
```

13.6.3.5.6 See Also:

[poke4](#), [peeks](#), [peek4u](#), [allocate](#), [free](#), [peek2s](#)

13.6.3.6 peek4u

```
<built-in> function peek4u(object addr_n_length)
```

Fetches an *unsigned* double word, or some *unsigned* double words, from an address in memory.

13.6.3.6.1 Parameters:

1. `addr_n_length` : an object, either of
 - ◆ an atom `addr` -- to fetch one double word at `addr`, or
 - ◆ a pair `{addr, len}` -- to fetch `len` double words at `addr`

13.6.3.6.2 Returns:

An **object**, either an atom if the input was a single address, or a sequence of atoms if a sequence was passed. In both cases, atoms returned are double words, in the range $0..power(2,32)-1$.

13.6.3.6.3 Errors:

`Peek()` in memory you don't own may be blocked by the OS, and cause a machine exception. If you use the define safe these routines will catch these problems with a EUPHORIA error.

When supplying a `{address, count}` sequence, the count must not be negative.

13.6.3.6.4 Comments:

Since addresses are 32-bit numbers, they can be larger than the largest value of type integer (31-bits). Variables that hold an address should therefore be declared as atoms.

It is faster to read several double words at once using the second form of `peek()` than it is to read one double word at a time in a loop. The returned sequence has the length you asked for on input.

Remember that `peek4u()` takes just one argument, which in the second form is actually a 2-element sequence.

The only difference between `peek4s()` and `peek4u()` is how double words with the highest bit set are returned. `peek4s()` assumes them to be negative, while `peek4u()` just assumes them to be large and positive.

13.6.3.6.5 Example 1:

```
-- The following are equivalent:
-- method 1
s = {peek4u(100), peek4u(104), peek4u(108), peek4u(112)}

-- method 2
s = peek4u({100, 4})
```

13.6.3.6.6 See Also:

[poke4](#), [peek](#), [peek4s](#), [allocate](#), [free](#), [peek2u](#)

13.6.3.7 peek_string

<built-in> `procedure peek_string(atom addr)`

Read an ASCII string in RAM, starting from a supplied address.

13.6.3.7.1 Parameters:

1. `addr` : an atom, the address at which to start reading.

13.6.3.7.2 Returns:

A **sequence**, of bytes, the string that could be read.

13.6.3.7.3 Errors:

Further, `peek()` memory that doesn't belong to your process is something the operating system could prevent, and you'd crash with a machine level exception.

13.6.3.7.4 Comments:

An ASCII string is any sequence of bytes and ends with a 0 byte. If you `peek_string()` at some place where there is no string, you will get a sequence of garbage.

13.6.3.7.5 See Also:

[peek](#), [peek_wstring](#), [allocate_string](#)

13.6.3.8 poke

<built-in> `procedure poke(atom addr, object x)`

Stores one or more bytes, starting at a memory location.

13.6.3.8.1 Parameters:

1. `addr` : an atom, the address at which to store
2. `x` : an object, either a byte or a non empty sequence of bytes.

13.6.3.8.2 Errors:

Poke() in memory you don't own may be blocked by the OS, and cause a machine exception. The -D SAFE option will make `poke()` catch this sort of issues.

13.6.3.8.3 Comments:

The lower 8-bits of each byte value, i.e. `remainder(x, 256)`, is actually stored in memory.

It is faster to write several bytes at once by poking a sequence of values, than it is to write one byte at a time in a loop.

Writing to the screen memory with `poke()` can be much faster than using `puts()` or `printf()`, but the programming is more difficult. In most cases the speed is not needed. For example, the EUPHORIA editor, `ed`, never uses `poke()`.

13.6.3.8.4 Example 1:

```
a = allocate(100)    -- allocate 100 bytes in memory

-- poke one byte at a time:
poke(a, 97)
poke(a+1, 98)
poke(a+2, 99)

-- poke 3 bytes at once:
poke(a, {97, 98, 99})
```

13.6.3.8.5 Example 2:

`demo/callmach.ex`

13.6.3.8.6 See Also:

[peek](#), [peeks](#), [poke4](#), [allocate](#), [free](#), [poke2](#), [call](#), [mem_copy](#), [mem_set](#)

13.6.3.9 poke2

<built-in> `procedure poke2(atom addr, object x)`

Stores one or more words, starting at a memory location.

13.6.3.9.1 Parameters:

1. `addr` : an atom, the address at which to store
2. `x` : an object, either a word or a non empty sequence of words.

13.6.3.9.2 Errors:

`Poke()` in memory you don't own may be blocked by the OS, and cause a machine exception. If you use the `define safe` these routines will catch these problems with a EUPHORIA error.

13.6.3.9.3 Comments:

There is no point in having `poke2s()` or `poke2u()`. For example, both 32768 and -32768 are stored as `#F000` when stored as words. It's up to whoever reads the value to figure it out.

It is faster to write several words at once by poking a sequence of values, than it is to write one words at a time in a loop.

Writing to the screen memory with `poke2()` can be much faster than using `puts()` or `printf()`, but the programming is more difficult. In most cases the speed is not needed. For example, the EUPHORIA editor, `ed`, never uses `poke2()`.

The 2-byte values to be stored can be negative or positive. You can read them back with either `peek2s()` or `peek2u()`. Actually, only `remainder(x,65536)` is being stored.

13.6.3.9.4 Example 1:

```
a = allocate(100)    -- allocate 100 bytes in memory

-- poke one 2-byte value at a time:
poke2(a, 12345)
poke2(a+2, #FF00)
poke2(a+4, -12345)

-- poke 3 2-byte values at once:
poke4(a, {12345, #FF00, -12345})
```

13.6.3.9.5 See Also:

[peek2s](#), [peek2u](#), [poke](#), [poke4](#), [allocate](#), [free](#), [call](#)

13.6.3.10 `poke4`

```
<built-in> procedure poke4(atom addr, object x)
```

Stores one or more double words, starting at a memory location.

13.6.3.10.1 Parameters:

1. `addr` : an atom, the address at which to store
2. `x` : an object, either a double word or a non empty sequence of double words.

13.6.3.10.2 Errors:

`Poke()` in memory you don't own may be blocked by the OS, and cause a machine exception. If you use the `define safe` these routines will catch these problems with a EUPHORIA error.

13.6.3.10.3 Comments:

There is no point in having `poke4s()` or `poke4u()`. For example, both `+power(2,31)` and `-power(2,31)` are stored as `#F0000000`. It's up to whoever reads the value to figure it out.

It is faster to write several double words at once by poking a sequence of values, than it is to write one double words at a time in a loop.

Writing to the screen memory with `poke4()` can be much faster than using `puts()` or `printf()`, but the programming is more difficult. In most cases the speed is not needed. For example, the EUPHORIA editor, `ed`, never uses `poke4()`.

The 4-byte values to be stored can be negative or positive. You can read them back with either `peek4s()` or `peek4u()`. However, the results are unpredictable if you want to store values with a fractional part or a magnitude greater than `power(2,32)`, even though EUPHORIA represents them all as atoms.

13.6.3.10.4 Example 1:

```
a = allocate(100)    -- allocate 100 bytes in memory

-- poke one 4-byte value at a time:
poke4(a, 9712345)
poke4(a+4, #FF00FF00)
poke4(a+8, -12345)

-- poke 3 4-byte values at once:
poke4(a, {9712345, #FF00FF00, -12345})
```

13.6.3.10.5 See Also:

[peek4s](#), [peek4u](#), [poke](#), [poke2](#), [allocate](#), [free](#), [call](#)

13.6.3.11 mem_copy

<built-in> `procedure mem_copy(atom destination, atom origin, integer len)`

Copy a block of memory from an address to another.

13.6.3.11.1 Parameters:

1. `destination` : an atom, the address at which data is to be copied
2. `origin` : an atom, the address from which data is to be copied
3. `len` : an integer, how many bytes are to be copied.

13.6.3.11.2 Comments:

The bytes of memory will be copied correctly even if the block of memory at `destination` overlaps with the block of memory at `origin`.

`mem_copy(destination, origin, len)` is equivalent to: `poke(destination, peek({origin, len}))` but is much faster.

13.6.3.11.3 Example 1:

```
dest = allocate(50)
src = allocate(100)
poke(src, {1,2,3,4,5,6,7,8,9})
mem_copy(dest, src, 9)
```

13.6.3.11.4 See Also:

[mem_set](#), [peek](#), [poke](#), [allocate](#), [free](#)

13.6.3.12 mem_set

<built-in> `procedure mem_set(atom destination, integer byte_value, integer how_many)`

Sets a contiguous range of memory locations to a single value.

13.6.3.12.1 Parameters:

1. `destination` : an atom, the address starting the range to set.
2. `byte_value` : an integer, the value to copy at all addresses in the range.
3. `how_many` : an integer, how many bytes are to be set.

13.6.3.12.2 Comments:

The low order 8 bits of `byte_value` are actually stored in each byte. `mem_set(destination, byte_value, how_many)` is equivalent to: `poke(destination, repeat(byte_value, how_many))` but is much faster.

13.6.3.12.3 Example 1:

```
destination = allocate(1000)
mem_set(destination, ' ', 1000)
-- 1000 consecutive bytes in memory will be set to 32
-- (the ASCII code for ' ')
```

13.6.3.12.4 See Also:

[peek](#), [poke](#), [allocate](#), [free](#), [mem_copy](#)

13.6.3.13 call

<built-in> `procedure call(atom addr)`

Call a machine language routine which was stored in memory prior.

13.6.3.13.1 Parameters:

1. `addr` : an atom, the address at which to transfer execution control.

13.6.3.13.2 Comments:

The machine code routine must execute a RET instruction #C3 to return control to EUPHORIA. The routine should save and restore any registers that it uses.

You can allocate a block of memory for the routine and then poke in the bytes of machine code using `allocate_code()`. You might allocate other blocks of memory for data and parameters that the machine code can operate on using `allocate()`. The addresses of these blocks could be part of the machine code.

If your machine code uses the stack, use `c_proc()` instead of `call()`.

13.6.3.13.3 Example 1:

`demo/callmach.ex`

13.6.3.13.4 See Also:

[allocate_code](#), [free_code](#), [c_proc](#), [define_c_proc](#)

13.6.3.14 check_calls

```
include std/memory.e
public integer check_calls
```

13.6.4 Safe memory access

13.6.4.1 register_block

```
include std/memory.e
public procedure register_block(atom block_addr, atom block_len, integer protection)
```

Description: Add a block of memory to the list of safe blocks maintained by safe.e (the debug version of memory.e). The block starts at address a. The length of the block is i bytes.

13.6.4.1.1 Parameters:

1. `block_addr` : an atom, the start address of the block
2. `block_len` : an integer, the size of the block.
3. `protection` : a constant integer, of the memory protection constants found in machine.e, that describes what access we have to the memory.

13.6.4.1.2 Comments:

In memory.e, this procedure does nothing. It is there to simplify switching between the normal and debug version of the library.

This routine is only meant to be used for debugging purposes. safe.e tracks the blocks of memory that your program is allowed to [peek\(\)](#), [poke\(\)](#), [mem_copy\(\)](#) etc. These are normally just the blocks that you have allocated using EUPHORIA's [allocate\(\)](#) routine, and which you have not yet freed using EUPHORIA's [free\(\)](#). In some cases, you may acquire additional, external, blocks of memory, perhaps as a result of calling a C routine.

If you are debugging your program using safe.e, you must register these external blocks of memory or safe.e will prevent you from accessing them. When you are finished using an external block you can unregister it using [unregister_block\(\)](#).

13.6.4.1.3 Example 1:

```
atom addr

addr = c_func(x, {})
register_block(addr, 5)
poke(addr, "ABCDE")
unregister_block(addr)
```

13.6.4.1.4 See Also:

[unregister_block](#), [safe.e](#)

13.6.4.2 unregister_block

```
include std/memory.e
public procedure unregister_block(atom block_addr)
```

Remove a block of memory from the list of safe blocks maintained by [safe.e](#) (the debug version of [memory.e](#)).

13.6.4.2.1 Parameters:

1. `block_addr` : an atom, the start address of the block

13.6.4.2.2 Comments:

In [memory.e](#), this procedure does nothing. It is there to simplify switching between the normal and debug version of the library.

This routine is only meant to be used for debugging purposes. Use it to unregister blocks of memory that you have previously registered using [register_block\(\)](#). By unregistering a block, you remove it from the list of safe blocks maintained by [safe.e](#). This prevents your program from performing any further reads or writes of memory within the block.

See [register_block\(\)](#) for further comments and an example.

13.6.4.2.3 See Also:

[register_block](#), [safe.e](#)

13.6.4.3 safe_address

```
include std/memory.e
public function safe_address(atom start, integer len, positive_int action)
```

Scans the list of registered blocks for any corruption.

13.6.4.3.1 Comments:

safe.e maintains a list of acquired memory blocks. Those gained through `allocate()` are automatically included. Any other block, for debugging purposes, must be registered by [register_block\(\)](#) and unregistered by [unregister_block\(\)](#).

The list is scanned and, if any block shows signs of corruption, it is displayed on the screen and the program terminates. Otherwise, nothing happens.

In `memory.e`, this routine does nothing. It is there to make switching between debugged and normal version of your program easier.

13.6.4.3.2 See Also:

[register_block](#), [unregister_block](#)

13.6.4.4 check_all_blocks

```
include std/memory.e
public procedure check_all_blocks()
```

13.6.4.5 prepare_block

```
include std/memory.e
export function prepare_block(atom addr, integer a, integer protection)
```

13.6.4.6 BORDER_SPACE

```
include std/memory.e
export constant BORDER_SPACE
```

13.6.4.7 trailer

```
include std/memory.e
export constant trailer
```

13.6.4.8 bordered_address

```
include std/memory.e
export type bordered_address(atom addr)
```

13.6.4.9 delete_routine

```
<built-in>function delete_routine( object x, integer rid )
```

Associates a routine for cleaning up after a euphoria object.

13.6.4.9.1 Comments:

`delete_routine()` associates a euphoria object with a routine id meant to clean up any allocated resources. It always returns an atom (double) or a sequence, depending on what was passed (integers are promoted to atoms).

The routine specified by `delete_routine()` should be a procedure that takes a single parameter, being the object to be cleaned up after. Objects are cleaned up under one of two circumstances. The first is if it's called as a parameter to `delete()`. After the call, the association with the delete routine is removed.

The second way for the delete routine to be called is when its reference count is reduced to 0. Before its memory is freed, the delete routine is called.

`delete_routine()` may be called multiple times for the same object. In this case, the routines are called in reverse order compared to how they were associated.

13.6.4.10 delete

```
<built-in>procedure delete( object x )
```

Calls the cleanup routines associated with the object, and removes the association with those routines.

13.6.4.10.1 Comments:

The cleanup routines associated with the object are called in reverse order than they were added. If the object is an integer, or if no cleanup routines are associated with the object, then nothing happens.

After the cleanup routines are called, the value of the object is unchanged, though the cleanup routine will no longer be associated with the object.

13.6.4.11 dep_works

```
include std/memory.e
export function dep_works()
```

13.6.4.12 VirtualFree_rid

```
include std/memory.e
export atom VirtualFree_rid
```

13.6.4.13 free_code

```
include std/memory.e
public procedure free_code(atom addr, integer size, valid_wordsize wordsize = 1)
```

13.6.5 safe.e

This is a slower DEBUGGING VERSION of machine.e

13.6.5.1 How To Use This File:

1. If your program doesn't already include machine.e add: include std/machine.e to your main .ex[w][u] file at the top.

2. To turn debug version on, issue

```
with define SAFE
```

in your main program, before the statement including machine.e.

3. If necessary, call register_block(address, length, memory_protection) to add additional "external" blocks of memory to the safe_address_list. These are blocks of memory that are safe to use but which you did not acquire through EUPHORIA's allocate(), allocate_data(), allocate_code() or memory_protect(). Call unregister_block(address) when you want to prevent further access to an external block.

4. Run your program. It might be 10x slower than normal but it's worth it to catch a nasty bug.

5. If a bug is caught, you will hear some "beep" sounds. Press Enter to clear the screen and see the error message. There will be a "divide by zero" traceback in ex.err so you can find the statement that is making the

illegal memory access.

6. To switch between normal and debug versions, simply comment in or out the "with define SAFE" directive. In means debugging and out means normal. Alternatively, you can use -D SAFE as a switch on the command line (debug) or not (normal).

7. The older method of switching files and renaming them ***no longer works***. machine.e conditionally includes safe.e.

This file is equivalent to machine.e, but it overrides the built-in

13.6.5.1.1 routines:

poke, peek, poke4, peek4s, peek4u, call, mem_copy, and mem_set

13.6.5.1.2 and it provides alternate versions of:

allocate, free

Your program will only be allowed to read/write areas of memory that it allocated (and hasn't freed), as well as areas in low memory that you list below, or add dynamically via register_block().

13.6.5.2 check_calls

```
include std/safe.e
public integer check_calls
```

Define block checking policy.

13.6.5.2.1 Comments:

If this integer is 1, (the default), check all blocks for edge corruption after each `call()`, `c_proc()` or `c_func()`. To save time, your program can turn off this checking by setting check_calls to 0.

13.6.5.3 edges_only

```
include std/safe.e
public integer edges_only
```

Determine whether to flag accesses to remote memory areas.

13.6.5.3.1 Comments:

If this integer is 1 (the default under *WIN32*), only check for references to the leader or trailer areas just outside each registered block, and don't complain about addresses that are far out of bounds (it's probably a legitimate block from another source)

For a stronger check, set this to 0 if your program will never read/write an unregistered block of memory.

On *WIN32* people often use unregistered blocks.

13.6.5.4 safe_address_list

```
include std/safe.e
public sequence safe_address_list
```

13.6.5.5 machine_addr

```
include std/safe.e
public type machine_addr(atom a)
```

13.6.5.6 BORDER_SPACE

```
include std/safe.e
export constant BORDER_SPACE
```

13.6.5.7 trailer

```
include std/safe.e
export constant trailer
```

13.6.5.8 bordered_address

```
include std/safe.e
export type bordered_address(atom addr)
```

13.6.5.9 safe_address

```
include std/safe.e
public function safe_address(atom start, integer len, positive_int action)
```

13.6.5.10 peek

```
include std/safe.e
override function peek(object
```

13.6.5.11 peeks

```
include std/safe.e
override function peeks(object
```

13.6.5.12 peek2u

```
include std/safe.e
override function peek2u(object
```

13.6.5.13 peek2s

```
include std/safe.e
override function peek2s(object
```

13.6.5.14 peek4s

```
include std/safe.e
override function peek4s(object
```

13.6.5.15 peek4u

```
include std/safe.e
override function peek4u(object
```

13.6.5.16 peek_string

```
include std/safe.e
override function peek_string(object
```

13.6.5.17 poke

```
include std/safe.e
override procedure poke(atom
```

13.6.5.18 poke2

```
include std/safe.e
override procedure poke2(atom
```

13.6.5.19 poke4

```
include std/safe.e
override procedure poke4(atom
```

13.6.5.20 mem_copy

```
include std/safe.e
override procedure mem_copy(machine_addr
```

13.6.5.21 mem_set

```
include std/safe.e
override procedure mem_set(machine_addr
```

13.6.5.22 show_block

```
include std/safe.e
public procedure show_block(sequence block_info)
```

13.6.5.23 check_all_blocks

```
include std/safe.e
public procedure check_all_blocks()
```

13.6.5.24 call

```
include std/safe.e
override procedure call(bordered_address
```

13.6.5.25 c_proc

```
include std/safe.e
override procedure c_proc(integer
```

13.6.5.26 c_func

```
include std/safe.e
override function c_func(integer
```

13.6.5.27 register_block

```
include std/safe.e
public procedure register_block(machine_addr block_addr, positive_int block_len, valid_memory_p
```

13.6.5.28 unregister_block

```
include std/safe.e
public procedure unregister_block(machine_addr block_addr)
```

13.6.5.29 prepare_block

```
include std/safe.e
export function prepare_block(atom a, integer n, natural protection)
```

13.6.5.30 allocate_data

```
include std/safe.e
public function allocate_data(positive_int n, integer cleanup = 0)
```

13.6.5.31 allocate

```
include std/safe.e
public function allocate(positive_int n, integer cleanup = 0)
```

13.6.5.32 deallocate

```
include std/safe.e
export procedure deallocate(atom a)
```

13.6.5.33 dep_works

```
include std/safe.e
export function dep_works()
```

13.6.5.34 VirtualFree_rid

```
include std/safe.e
export atom VirtualFree_rid
```

13.6.5.35 free_code

```
include std/safe.e
public procedure free_code(atom addr, integer size, valid_wordsize wordsize = 1)
```

13.6.6 Microsoft Windows Memory Protection Constants

These constant names are taken right from Microsoft's Memory Protection constants.

13.6.6.1 PAGE_EXECUTE

```
include std/memconst.e
public constant PAGE_EXECUTE
```

You may run the data in this page

13.6.6.2 PAGE_EXECUTE_READ

```
include std/memconst.e
public constant PAGE_EXECUTE_READ
```

You may read or run the data

13.6.6.3 PAGE_EXECUTE_READWRITE

```
include std/memconst.e
public constant PAGE_EXECUTE_READWRITE
```

You may run, read or write in this page

13.6.6.4 PAGE_EXECUTE_WRITECOPY

```
include std/memconst.e
public constant PAGE_EXECUTE_WRITECOPY
```

You may run or write in this page

13.6.6.5 PAGE_WRITECOPY

```
include std/memconst.e
public constant PAGE_WRITECOPY
```

You may write to this page.

13.6.6.6 PAGE_READWRITE

```
include std/memconst.e
public constant PAGE_READWRITE
```

You may read or write in this page.

13.6.6.7 PAGE_READONLY

```
include std/memconst.e
public constant PAGE_READONLY
```

You may only read data in this page

13.6.6.8 PAGE_NOACCESS

```
include std/memconst.e
public constant PAGE_NOACCESS
```

You have no access to this page

13.6.7 Standard Library Memory Protection Constants

Memory Protection Constants are the same constants names and meaning across all platforms yet possibly of different numeric value. They are only necessary for [allocate_protect](#)

The constant names are created like this: You have four aspects of protection READ, WRITE, EXECUTE and COPY. You take the word PAGE and you concatenate an underscore and the aspect in the order above. For example: PAGE_WRITE_EXECUTE The sole exception to this nomenclature is when you will have no access to the page the constant is called PAGE_NOACCESS.

13.6.7.1 PAGE_NONE

```
include std/memconst.e
public constant PAGE_NONE
```

You have no access to this page An alias to PAGE_NOACCESS

13.6.7.2 PAGE_READ_EXECUTE

```
include std/memconst.e
public constant PAGE_READ_EXECUTE
```

You may read or run the data An alias to PAGE_EXECUTE_READ

13.6.7.3 PAGE_READ_WRITE

```
include std/memconst.e
public constant PAGE_READ_WRITE
```

You may read or write to this page An alias to PAGE_READWRITE

13.6.7.4 PAGE_READ

```
include std/memconst.e
public constant PAGE_READ
```

You may only read to this page An alias to PAGE_READONLY

13.6.7.5 PAGE_READ_WRITE_EXECUTE

```
include std/memconst.e
public constant PAGE_READ_WRITE_EXECUTE
```

You may run, read or write in this page An alias to PAGE_EXECUTE_READWRITE

13.6.7.6 PAGE_WRITE_EXECUTE_COPY

```
include std/memconst.e
public constant PAGE_WRITE_EXECUTE_COPY
```

You may run or write to this page. Data will copied for use with other processes when you first write to it.

13.6.7.7 PAGE_WRITE_COPY

```
include std/memconst.e
public constant PAGE_WRITE_COPY
```

You may write to this page. Data will copied for use with other processes when you first write to it.

13.6.7.8 valid_memory_protection_constant

```
include std/memconst.e
export type valid_memory_protection_constant(integer x)
```

13.6.7.9 test_read

```
include std/memconst.e
export function test_read(valid_memory_protection_constant protection)
```


13.6.7.10 test_write

```
include std/memconst.e
export function test_write(valid_memory_protection_constant protection)
```

13.6.7.11 test_exec

```
include std/memconst.e
export function test_exec(valid_memory_protection_constant protection)
```

13.6.7.12 valid_wordsize

```
include std/memconst.e
export type valid_wordsize(integer i)
```

13.6.7.13 DEP_really_works

```
include std/memconst.e
export integer DEP_really_works
```

13.6.7.14 MEM_COMMIT

```
include std/memconst.e
export constant MEM_COMMIT
```

13.6.7.15 MEM_RESERVE

```
include std/memconst.e
export constant MEM_RESERVE
```

13.6.7.16 MEM_RESET

```
include std/memconst.e
export constant MEM_RESET
```

13.6.7.17 MEM_RELEASE

```
include std/memconst.e
export constant MEM_RELEASE
```

13.6.7.18 FREE_RID

```
include std/memconst.e
export integer FREE_RID public enum A_READ
```

13.6.7.19 A_WRITE

```
include std/memconst.e
export integer A_WRITE
```

13.6.7.20 A_EXECUTE

```
include std/memconst.e
export integer A_EXECUTE
```

13.6.7.21 M_ALLOC

```
include std/memconst.e
export constant M_ALLOC
```

13.6.7.22 M_FREE

```
include std/memconst.e
export constant M_FREE
```

13.6.7.23 kernel_dll

```
include std/memconst.e
export atom kernel_dll
```

13.6.7.24 memDLL_id

```
include std/memconst.e
export atom memDLL_id
```

13.6.7.25 VirtualAlloc_rid

```
include std/memconst.e
export atom VirtualAlloc_rid
```

13.6.7.26 VirtualLock_rid

```
include std/memconst.e
export atom VirtualLock_rid
```

13.6.7.27 VirtualUnlock_rid

```
include std/memconst.e
export atom VirtualUnlock_rid
```

13.6.7.28 VirtualProtect_rid

```
include std/memconst.e
export atom VirtualProtect_rid
```

13.6.7.29 GetLastError_rid

```
include std/memconst.e
export atom GetLastError_rid
```

13.6.7.30 GetSystemInfo_rid

```
include std/memconst.e
export atom GetSystemInfo_rid
```

13.6.7.31 PROT_EXEC

```
include std/unix/mmap.e
public constant PROT_EXEC
```

13.6.7.32 PROT_READ

```
include std/unix/mmap.e
public constant PROT_READ
```

13.6.7.33 PROT_WRITE

```
include std/unix/mmap.e
public constant PROT_WRITE
```

13.6.7.34 PROT_NONE

```
include std/unix/mmap.e
public constant PROT_NONE
```

13.6.7.35 MAP_ANONYMOUS

```
include std/unix/mmap.e
public constant MAP_ANONYMOUS
```

13.6.7.36 MAP_PRIVATE

```
include std/unix/mmap.e
public constant MAP_PRIVATE
```

13.6.7.37 MAP_SHARED

```
include std/unix/mmap.e
public constant MAP_SHARED
```

13.6.7.38 MAP_TYPE

```
include std/unix/mmap.e
public constant MAP_TYPE
```

13.6.7.39 MAP_FIXED

```
include std/unix/mmap.e
public constant MAP_FIXED
```

13.6.7.40 MAP_FILE

```
include std/unix/mmap.e
public constant MAP_FILE
```

13.6.7.41 get_page_size

```
include std/unix/mmap.e
public function get_page_size()
```

13.6.7.42 mmap

```
include std/unix/mmap.e
public function mmap(object start, integer length, valid_memory_protection_constant protection,
```

13.6.7.43 munmap

```
include std/unix/mmap.e
public function munmap(atom addr, integer length)
```

13.6.7.44 mlock

```
include std/unix/mmap.e
public function mlock(atom addr, integer length)
```

13.6.7.45 munlock

```
include std/unix/mmap.e
public function munlock(atom addr, integer length)
```

13.6.7.46 mprotect

```
include std/unix/mmap.e
public function mprotect(atom addr, integer length, valid_memory_protection_constant protection)
```

13.6.7.47 is_valid_memory_protection_constant

```
include std/unix/mmap.e
public function is_valid_memory_protection_constant(integer x)
```

13.6.7.48 is_page_aligned_address

```
include std/unix/mmap.e
public function is_page_aligned_address(atom a)
```

14 Graphics

- Error Code Constants
- video_config sequence accessors
- Routines
- Graphics - Cross Platform
 - Routines
 - Graphics Modes
- Graphics - Image Routines
 - Bitmap handling

14.1 Error Code Constants

14.1.1 BMP_SUCCESS

```
include std/graphcst.e
public enum BMP_SUCCESS
```

14.1.1.1 BMP_OPEN_FAILED

```
include std/graphcst.e
public enum BMP_OPEN_FAILED
```

14.1.1.2 BMP_UNEXPECTED_EOF

```
include std/graphcst.e
public enum BMP_UNEXPECTED_EOF
```

14.1.1.3 BMP_UNSUPPORTED_FORMAT

```
include std/graphcst.e
public enum BMP_UNSUPPORTED_FORMAT
```

14.1.1.4 BMP_INVALID_MODE

```
include std/graphcst.e
public enum BMP_INVALID_MODE
```

14.1.2 video_config sequence accessors

14.1.2.1 VC_COLOR

```
include std/graphcst.e  
public enum VC_COLOR
```

14.1.2.2 VC_MODE

```
include std/graphcst.e  
public enum VC_MODE
```

14.1.2.3 VC_LINES

```
include std/graphcst.e  
public enum VC_LINES
```

14.1.2.4 VC_COLUMNS

```
include std/graphcst.e  
public enum VC_COLUMNS
```

14.1.2.5 VC_XPIXELS

```
include std/graphcst.e  
public enum VC_XPIXELS
```

14.1.2.6 VC_YPIXELS

```
include std/graphcst.e  
public enum VC_YPIXELS
```

14.1.2.7 VC_NCOLORS

```
include std/graphcst.e  
public enum VC_NCOLORS
```


14.1.2.8 VC_PAGES

```
include std/graphcst.e
public enum VC_PAGES
```

14.1.2.9 VC_SCRNLINES

```
include std/graphcst.e
public enum VC_SCRNLINES
```

14.1.2.10 VC_SCRNCOLS

```
include std/graphcst.e
public enum VC_SCRNCOLS
```

14.1.2.11 Colors

14.1.2.12 BLACK

```
include std/graphcst.e
public constant BLACK
```

14.1.2.13 BLUE

```
include std/graphcst.e
public constant BLUE
```

14.1.2.14 GREEN

```
include std/graphcst.e
public constant GREEN
```

14.1.2.15 CYAN

```
include std/graphcst.e
public constant CYAN
```

14.1.2.16 RED

```
include std/graphcst.e
public constant RED
```

14.1.2.17 MAGENTA

```
include std/graphcst.e
public constant MAGENTA
```

14.1.2.18 BROWN

```
include std/graphcst.e
public constant BROWN
```

14.1.2.19 WHITE

```
include std/graphcst.e
public constant WHITE
```

14.1.2.20 GRAY

```
include std/graphcst.e
public constant GRAY
```

14.1.2.21 BRIGHT_BLUE

```
include std/graphcst.e
public constant BRIGHT_BLUE
```

14.1.2.22 BRIGHT_GREEN

```
include std/graphcst.e
public constant BRIGHT_GREEN
```

14.1.2.23 BRIGHT_CYAN

```
include std/graphcst.e
public constant BRIGHT_CYAN
```

14.1.2.24 BRIGHT_RED

```
include std/graphcst.e
public constant BRIGHT_RED
```

14.1.2.25 BRIGHT_MAGENTA

```
include std/graphcst.e
public constant BRIGHT_MAGENTA
```

14.1.2.26 YELLOW

```
include std/graphcst.e
public constant YELLOW
```

14.1.2.27 BRIGHT_WHITE

```
include std/graphcst.e
public constant BRIGHT_WHITE
```

14.1.2.28 true_color

```
include std/graphcst.e
export constant true_color
```

14.1.2.29 BLINKING

```
include std/graphcst.e
public constant BLINKING
```

Add to color to get blinking text

14.1.2.30 BYTES_PER_CHAR

```
include std/graphcst.e
public constant BYTES_PER_CHAR
```

14.1.2.31 color

```
include std/graphcst.e
public type color(integer x)
```

14.1.3 Routines

14.1.3.1 mixture

```
include std/graphcst.e
public type mixture(sequence s)
```

Mixture Type

14.1.3.1.1 Comments:

A mixture is a {red, green, blue} triple of intensities, which enables you to define custom colors. Intensities must be from 0 (weakest) to 63 (strongest). Thus, the brightest white is {63, 63, 63}.

14.1.3.2 video_config

```
include std/graphcst.e
public function video_config()
```

14.1.3.2.1 Return a description of the current video configuration:

14.1.3.2.2 Returns:

A **sequence**, of 10 non-negative integers, laid out as follows:

1. color monitor? -- 1 0 if monochrome, 1 otherwise
2. current video mode
3. number of text rows in console buffer
4. number of text columns in console buffer
5. screen width in pixels

6. screen height in pixels
7. number of colors
8. number of display pages
9. number of text rows for current screen size
10. number of text columns for current screen size

14.1.3.2.3 Comments:

14.1.3.2.4 A public enum is available for convenient access to the returned configuration data:

- VC_COLOR
- VC_MODE
- VC_LINES
- VC_COLUMNS
- VC_XPIXELS
- VC_YPIXELS
- VC_NCOLORS
- VC_PAGES
- VC_LINES
- VC_COLUMNS
- VC_SCRNLINES
- VC_SCRNCOLS

This routine makes it easy for you to parameterize a program so it will work in many different graphics modes.

14.1.3.2.5 Example:

```
vc = video_config()  
-- vc could be {1, 3, 300, 132, 0, 0, 32, 8, 37, 90}
```

14.1.3.2.6 See Also:

[graphics_mode](#)

14.2 Graphics - Cross Platform

Routines

[position](#)
[get_position](#)
[text_color](#)
[bk_color](#)
[wrap](#)
[scroll](#)

Graphics Modes
graphics_mode

14.2.1 Routines

14.2.1.1 position

```
<built-in> procedure position(integer row, integer column)
```

14.2.1.1.1 Parameters:

1. `row` : an integer, the index of the row to position the cursor on.
2. `column` : an integer, the index of the column to position the cursor on.

Set the cursor to line `row`, column `column`, where the top left corner of the screen is line 1, column 1. The next character displayed on the screen will be printed at this location. `position()` will report an error if the location is off the screen. The *Windows* console does not check for rows, as the physical height of the console may be vastly less than its logical height.

14.2.1.1.2 Example 1:

```
position(2,1)
-- the cursor moves to the beginning of the second line from the top
```

14.2.1.1.3 See Also:

[get_position](#)

14.2.1.2 get_position

```
include std/graphics.e
public function get_position()
```

Return the current line and column position of the cursor

14.2.1.2.1 Returns:

A **sequence**, `{line, column}`, the current position of the text mode cursor.

14.2.1.2.2 Comments:

The coordinate system for displaying text is different from the one for displaying pixels. Pixels are displayed such that the top-left is (x=0,y=0) and the first coordinate controls the horizontal, left-right location. In pixel-graphics modes you can display both text and pixels. `get_position()` returns the current line and column for the text that you are displaying, not the pixels that you may be plotting. There is no corresponding routine for getting the current pixel position, because there is not such a thing.

14.2.1.2.3 See Also:

[position](#)

14.2.1.3 text_color

```
include std/graphics.e
public procedure text_color(color c)
```

Set the foreground text color.

14.2.1.3.1 Parameters:

1. `c` : the new text color. Add `BLINKING` to get blinking text in some modes.

14.2.1.3.2 Comments:

Text that you print after calling `text_color()` will have the desired color.

When your program terminates, the last color that you selected and actually printed on the screen will remain in effect. Thus you may have to print something, maybe just `'\n'`, in `WHITE` to restore white text, especially if you are at the bottom line of the screen, ready to scroll up.

14.2.1.3.3 Example:

```
text_color(BRIGHT_BLUE)
```

14.2.1.3.4 See Also:

[bk_color](#) , [clear_screen](#)

14.2.1.4 bk_color

```
include std/graphics.e
public procedure bk_color(color c)
```

Set the background color to one of the 16 standard colors.

14.2.1.4.1 Parameters:

1. `c` : the new text color. Add `BLINKING` to get blinking text in some modes.

14.2.1.4.2 Comments:

To restore the original background color when your program finishes, e.g. `0` - `BLACK`, you must call `bk_color(0)`. If the cursor is at the bottom line of the screen, you may have to actually print something before terminating your program. Printing `'\n'` may be enough.

14.2.1.4.3 Example:

```
bk_color(BLACK)
```

14.2.1.4.4 See Also:

[text_color](#)

14.2.1.5 wrap

```
include std/graphics.e
public procedure wrap(boolean on)
```

Determine whether text will wrap when hitting the rightmost column.

14.2.1.5.1 Parameters:

1. `on` : a boolean, `0` to truncate text, nonzero to wrap.

14.2.1.5.2 Comments:

By default text will wrap.

Use `wrap()` in text modes or pixel-graphics modes when you are displaying long lines of text.

14.2.1.5.3 Example:

```
puts(1, repeat('x', 100) & "\n\n")
-- now have a line of 80 'x' followed a line of 20 more 'x'
wrap(0)
puts(1, repeat('x', 100) & "\n\n")
-- creates just one line of 80 'x'
```

14.2.1.5.4 See Also:

[puts](#), [position](#)

14.2.1.6 scroll

```
include std/graphics.e
public procedure scroll(integer amount, positive_int top_line, positive_int bottom_line)
```

Scroll a region of text on the screen.

14.2.1.6.1 Parameters:

1. `amount` : an integer, the number of lines by which to scroll. This is >0 to scroll up and <0 to scroll down.
2. `top_line` : the 1-based number of the topmost line to scroll.
3. `bottom_line` : the 1-based number of the bottom-most line to scroll.

14.2.1.6.2 Comments:

inclusive. New blank lines will appear at the top or bottom.

You could perform the scrolling operation using a series of calls to `[:puts] ()`, but `scroll()` is much faster.

The position of the cursor after scrolling is not defined.

14.2.1.6.3 Example 1:

```
bin/ed.ex
```

14.2.1.6.4 See Also:

[clear_screen](#), [text_rows](#)

14.2.2 Graphics Modes

14.2.2.1 graphics_mode

```
include std/graphics.e
public function graphics_mode(mode m = - 1)
```

Attempt to set up a new graphics mode.

14.2.2.1.1 Parameters:

1. *m* : an integer, ignored.

14.2.2.1.2 Returns:

An **integer**, always returns zero.

14.2.2.1.3 Comments:

- This has no effect on Unix platforms.
- On Windows, it causes a console to be shown if one has not already been created.

14.2.2.1.4 See Also:

[video_config](#)

14.3 Graphics - Image Routines

[graphics_point](#)
Bitmap handling
[read_bitmap](#)
[save_bitmap](#)

14.3.1 graphics_point

```
include std/image.e
public type graphics_point(sequence p)
```

14.3.2 Bitmap handling

14.3.2.1 read_bitmap

```
include std/image.e
public function read_bitmap(sequence file_name)
```

Read a bitmap (.BMP) file into a 2-d sequence of sequences (image)

14.3.2.1.1 Parameters:

1. `file_name` : a sequence, the path to a .bmp file to read from. The extension is not assumed if missing.

14.3.2.1.2 Returns:

An **object**, on success, a sequence of the form {palette, image}. On failure, an error code is returned.

14.3.2.1.3 Comments:

In the returned value, the first element is a list of mixtures, each of which defines a color, and the second, a list of point rows. Each pixel in a row is represented by its color index.

The file should be in the bitmap format. The most common variations of the format are supported.

Bitmaps of 2, 4, 16 or 256 colors are supported. If the file is not in a good format, an error code (atom) is returned instead

```
public constant
    BMP_OPEN_FAILED = 1,
    BMP_UNEXPECTED_EOF = 2,
    BMP_UNSUPPORTED_FORMAT = 3
```

You can create your own bitmap picture files using Windows Paintbrush and many other graphics programs. You can then incorporate these pictures into your EUPHORIA programs.

14.3.2.1.4 Example 1:

```
x = read_bitmap("c:\\windows\\arcade.bmp")
```

14.3.2.1.5 Note:

double backslash needed to get single backslash in a string

14.3.2.1.6 See Also:

[save_bitmap](#)

14.3.2.2 save_bitmap

```
include std/image.e
public function save_bitmap(two_seq palette_n_image, sequence file_name)
```

Create a .BMP bitmap file, given a palette and a 2-d sequence of sequences of colors.

14.3.2.2.1 Parameters:

1. `palette_n_image` : a {palette, image} pair, like [read_bitmap\(\)](#) returns
2. `file_name` : a sequence, the name of the file to save to.

14.3.2.2.2 Returns:

An **integer**, 0 on success.

14.3.2.2.3 Comments:

This routine does the opposite of [read_bitmap\(\)](#). The first element of `palette_n_image` is a sequence of [mixtures](#) defining each color in the bitmap. The second element is a sequence of sequences of colors. The inner sequences must have the same length.

14.3.2.2.4 The result will be one of the following codes:

```
public constant
    BMP_SUCCESS = 0,
    BMP_OPEN_FAILED = 1,
    BMP_INVALID_MODE = 4 -- invalid graphics mode
                        -- or invalid argument
```

`save_bitmap()` produces bitmaps of 2, 4, 16, or 256 colors and these can all be read with `read_bitmap()`. Windows Paintbrush and some other tools do not support 4-color bitmaps.

14.3.2.2.5 Example 1:

```
code = save_bitmap({paletteData, imageData},  
                  "c:\\example\\a1.bmp")
```

14.3.2.2.6 See Also:

[read_bitmap](#)

15 EUPHORIA Routines

EUPHORIA Information

- Numeric Version Information
- Compiled Platform Information
- String Version Information
- Copyright Information

Keyword Data

Syntax Coloring

Source Tokenizer

Unit Testing Framework

- Background
- Constants
- Setup Routines
- Reporting
- Tests

15.1 EUPHORIA Information

- Numeric Version Information
- Compiled Platform Information
 - platform_name
 - version
 - version_major
 - version_minor
 - version_patch
 - version_revision
- String Version Information
 - version_type
 - version_string
 - version_string_short
 - version_string_long
- Copyright Information
 - euphoria_copyright
 - pcre_copyright
 - all_copyrights

15.1.1 Numeric Version Information

15.1.2 Compiled Platform Information

15.1.2.1 platform_name

```
include info.e
public function platform_name()
```

Get the platform name

15.1.2.1.1 Returns:

A **sequence**, containing the platform name, i.e. Windows, Linux, DOS, FreeBSD or OS X.

15.1.2.2 version

```
include info.e
public function version()
```

Get the version, as an integer, of the host EUPHORIA

15.1.2.2.1 Returns:

An **integer**, representing Major, Minor and Patch versions. Version 4.0.0 will return 40000, 4.0.1 will return 40001, 5.6.2 will return 50602, 5.12.24 will return 512624, etc...

15.1.2.3 version_major

```
include info.e
public function version_major()
```

Get the major version of the host EUPHORIA

15.1.2.3.1 Returns:

An **integer**, representing the Major version number. Version 4.0.0 will return 4, version 5.6.2 will return 5, etc...

15.1.2.4 version_minor

```
include info.e
public function version_minor()
```

Get the minor version of the hosting EUPHORIA

15.1.2.4.1 Returns:

An **integer**, representing the Minor version number. Version 4.0.0 will return 0, 4.1.0 will return 1, 5.6.2 will return 6, etc...

15.1.2.5 version_patch

```
include info.e
public function version_patch()
```

Get the patch version of the hosting EUPHORIA

15.1.2.5.1 Returns:

An **integer**, representing the Path version number. Version 4.0.0 will return 0, 4.0.1 will return 1, 5.6.2 will return 2, etc...

15.1.2.6 version_revision

```
include info.e
public function version_revision()
```

Get the source code revision of the hosting EUPHORIA

15.1.2.6.1 Returns:

A text **sequence**, containing the source code management system's revision number that the executing EUPHORIA was built from.

15.1.3 String Version Information

15.1.3.1 version_type

```
include info.e
public function version_type()
```

Get the type version of the hosting EUPHORIA

15.1.3.1.1 Returns:

A **sequence**, representing the Type version string. Version 4.0.0 alpha 1 will return `alpha 1`. 4.0.0 beta 2 will return `beta 2`. 4.0.0 final, or release, will return `release`.

15.1.3.2 version_string

```
include info.e
public function version_string()
```

Get a normal version string

15.1.3.2.1 Returns:

A **#sequence**, representing the Major, Minor, Patch, Type and Revision all in one string.

15.1.3.2.2 Example return values:

- "4.0.0 alpha 3 (r1234)"
- "4.0.0 release (r271)"
- "4.0.2 beta 1 (r2783)"

15.1.3.3 version_string_short

```
include info.e
public function version_string_short()
```

Get a short version string

15.1.3.3.1 Returns:

A **sequence**, representing the Major, Minor and Patch all in one string.

15.1.3.3.2 Example return values:

- "4.0.0"
- "4.0.2"
- "5.6.2"

15.1.3.4 version_string_long

```
include info.e
public function version_string_long()
```

Get a long version string

15.1.3.4.1 Returns:

Same **value**, as [version_string](#) with the addition of the platform name.

15.1.3.4.2 Example return values:

- "4.0.0 alpha 3 for Windows"
- "4.0.0 release for Linux"
- "5.6.2 release for OS X"

15.1.4 Copyright Information

15.1.4.1 euphoria_copyright

```
include info.e
public function euphoria_copyright()
```

Get the copyright statement for EUPHORIA

15.1.4.1.1 Returns:

A **sequence**, containing 2 sequences: product name and copyright message

15.1.4.1.2 Example 1:

```
sequence info = euphoria_copyright()
-- info = {
--     "EUPHORIA v4.0.0 alpha 3",
--     "Copyright (c) XYZ, ABC\n" &
--     "Copyright (c) ABC, DEF"
-- }
```

15.1.4.2 pcre_copyright

```
include info.e
public function pcre_copyright()
```

Get the copyright statement for PCRE.

15.1.4.2.1 Returns:

A **sequence**, containing 2 sequences: product name and copyright message.

15.1.4.2.2 See Also:

[euphoria_copyright\(\)](#)

15.1.4.3 all_copyrights

```
include info.e
public function all_copyrights()
```

Get all copyrights associated with this version of EUPHORIA.

15.1.4.3.1 Returns:

A **sequence**, of product names and copyright messages.

```
{
  { ProductName, CopyrightMessage },
  { ProductName, CopyrightMessage },
  ...
}
```

15.2 Keyword Data

15.2.1 keywords

```
include keywords.e
public constant keywords
```

Sequence of EUPHORIA keywords

15.2.1.1 builtins

```
include keywords.e
public constant builtins
```

Sequence of EUPHORIA's built-in function names

15.3 Syntax Coloring

Syntax Color Break EUPHORIA statements into words with multiple colors. The editor and pretty printer (eprint.ex) both use this file.

15.3.1 set_colors

```
include syncolor.e
public procedure set_colors(sequence pColorList)
```

15.3.1.1 init_class

```
include syncolor.e
public procedure init_class()
```

15.3.1.2 SyntaxColor

```
include syncolor.e
public function SyntaxColor(sequence pline)
```

15.4 Source Tokenizer

15.4.1 T_EOF

```
include tokenize.e
public constant T_EOF
```

15.4.1.1 T_NULL

```
include tokenize.e
public constant T_NULL
```

15.4.1.2 TF_HEX

```
include tokenize.e
public constant TF_HEX
```

15.4.1.3 TF_INT

```
include tokenize.e
public constant TF_INT
```

15.4.1.4 TF_ATOM

```
include tokenize.e
public constant TF_ATOM
```

15.4.1.5 TTYPE

```
include tokenize.e
public enum TTYPE
```

15.4.1.6 TDATA

```
include tokenize.e
public enum TDATA
```

15.4.1.7 TLNUM

```
include tokenize.e
public enum TLNUM
```

15.4.1.8 TLPOS

```
include tokenize.e
public enum TLPOS
```

15.4.1.9 TFORM

```
include tokenize.e
public enum TFORM
```

15.4.1.10 et_error_string

```
include tokenize.e
public function et_error_string(integer err)
```

15.4.1.11 et_keep_blanks

```
include tokenize.e
public procedure et_keep_blanks(integer toggle)
```

return blank lines as tokens default is FALSE

15.4.1.12 et_keep_comments

```
include tokenize.e
public procedure et_keep_comments(integer toggle)
```

return comments as tokens default is FALSE

15.4.1.13 et_string_numbers

```
include tokenize.e
public procedure et_string_numbers(integer toggle)
```

return TDATA for all T_NUMBER tokens in "string" format

15.4.1.13.1 by default:

T_NUMBER tokens return atoms T_CHAR tokens return single integer chars T_EOF tokens return undefined data all other tokens return strings

15.4.1.14 ET_TOKENS

```
include tokenize.e
public constant ET_TOKENS
```

15.4.1.15 ET_ERROR

```
include tokenize.e
public constant ET_ERROR
```

15.4.1.16 ET_ERR_LINE

```
include tokenize.e
public constant ET_ERR_LINE
```

15.4.1.17 ET_ERR_COLUMN

```
include tokenize.e
public constant ET_ERR_COLUMN
```

15.4.1.18 et_tokenize_string

```
include tokenize.e
public function et_tokenize_string(sequence code)
```

15.4.1.19 et_tokenize_file

```
include tokenize.e
public function et_tokenize_file(sequence fname)
```

15.5 Unit Testing Framework

Background

Constants

TEST_QUIET

TEST_SHOW_FAILED_ONLY

TEST_SHOW_ALL

Setup Routines

set_test_verbosity

```
    set_wait_on_summary
    set_accumulate_summary
    set_test_abort
Reporting
    test_report
Tests
    test_equal
    test_not_equal
    test_true
    assert
    test_false
    test_fail
    test_pass
```

15.5.1 Background

Unit testing is the process of assuring that the smallest programming units are actually delivering functionality that complies with their specification. The units in question are usually individual routines rather than whole programs or applications.

The theory is that if the components of a system are working correctly, then there is a high probability that a system using those components can be made to work correctly.

In EUPHORIA terms, this framework provides the tools to make testing and reporting on functions and procedures easy and standardized. It gives us a simple way to write a test case and to report on the findings. Example:

```
include std/unittest.e

test_equal( "Power function test #1", 4, power(2, 2))
test_equal( "Power function test #2", 4, power(16, 0.5))

test_report()
```

Name your test file in the special manner, `t_NAME.e` and then simply run `eutest` in that directory.

```
C:\EUPHORIA> eutest
t_math.e:
failed: Bad math, expected: 100 but got: 8
2 tests run, 1 passed, 1 failed, 50.0% success

==== Test failure summary:
FAIL: t_math.e

2 file(s) run 1 file(s) failed, 50.0% success--
```

In this example, we use the `test_equal` function to record the result of a test. The first parameter is the name of the test, which can be anything and is displayed if the test fails. The second parameter is the expected result -- what we expect the function being tested to return. The third parameter is the actual result returned by the function being tested. This is usually written as a call to the function itself.

It is typical to provide as many test cases as would be required to give us confidence that the function is being truly exercised. This includes calling it with typical values and edge-case or exceptional values. It is also useful to test the function's error handling by calling it with bad parameters.

When a test fails, the framework displays a message, showing the test's name, the expected result and the actual result. You can configure the framework to display each test run, regardless of whether it fails or not.

After running a series of tests, you can get a summary displayed by calling the `test_report()` procedure. To get a better feel for unit testing, have a look at the provided test cases for the standard library in the *tests* directory.

When included in your program, `unittest.e` sets a crash handler to log a crash as a failure.

15.5.2 Constants

15.5.2.1 TEST_QUIET

```
include std/unittest.e
public enum TEST_QUIET
```

15.5.2.2 TEST_SHOW_FAILED_ONLY

```
include std/unittest.e
public enum TEST_SHOW_FAILED_ONLY
```

15.5.2.3 TEST_SHOW_ALL

```
include std/unittest.e
public enum TEST_SHOW_ALL
```

15.5.3 Setup Routines

15.5.3.1 set_test_verbosity

```
include std/unittest.e
public procedure set_test_verbosity(atom verbosity)
```

Set the amount of information that is returned about passed and failed tests.

15.5.3.1.1 Parameters:

1. `verbosity` : an atom which takes predefined values for verbosity levels.

15.5.3.1.2 Comments:

The following values are allowable for `verbosity`:

- `TEST_QUIET -- 0`,
- `TEST_SHOW_FAILED_ONLY -- 1`
- `TEST_SHOW_ALL -- 2`

However, anything less than `TEST_SHOW_FAILED_ONLY` is treated as `TEST_QUIET`, and everything above `TEST_SHOW_ALL` is treated as `TEST_SHOW_ALL`.

- At the lowest verbosity level, only the score is shown, ie the ratio passed tests/total tests.
- At the medium level, in addition, failed tests display their name, the expected outcome and the outcome they got. This is the initial setting.
- At the highest level of verbosity, each test is reported as passed or failed.

If a file crashes when it should not, this event is reported no matter the verbosity level.

The command line switch `""-failed"` causes verbosity to be set to medium at startup. The command line switch `""-all"` causes verbosity to be set to high at startup.

15.5.3.1.3 See Also:

[test_report](#)

15.5.3.2 `set_wait_on_summary`

```
include std/unittest.e
public procedure set_wait_on_summary(integer to_wait)
```

Request the test report to pause before exiting.

15.5.3.2.1 Parameters:

1. `to_wait` : an integer, zero not to wait, nonzero to wait.

15.5.3.2.2 Comments:

Depending on the environment, the test results may be invisible if `set_wait_on_summary(1)` was not called prior, as this is not the default. The command line switch `""-wait"` performs this call.

15.5.3.2.3 See Also:

[test_report](#)

15.5.3.3 set_accumulate_summary

```
include std/unittest.e
public procedure set_accumulate_summary(integer accumulate)
```

Request the test report to save run stats in "unittest.dat" before exiting.

15.5.3.3.1 Parameters:

1. `accumulate` : an integer, zero not to accumulate, nonzero to accumulate.

15.5.3.3.2 Comments:

The file "unittest.dat" is appended to with {t,f}

where

t is total number of tests run

f is the total number of tests that failed

15.5.3.4 set_test_abort

```
include std/unittest.e
public function set_test_abort(integer abort_test)
```

Set behavior on test failure, and return previous value.

15.5.3.4.1 Parameters:

1. `abort_test` : an integer, the new value for this setting.

15.5.3.4.2 Returns:

An **integer**, the previous value for the setting.

15.5.3.4.3 Comments:

By default, the tests go on even if a file crashed.

15.5.4 Reporting

15.5.4.1 test_report

```
include std/unittest.e
public procedure test_report()
```

Output test report

15.5.4.1.1 Comments:

The report components are described in the comments section for [set_test_verbosity](#). Everything prints on the standard error device.

15.5.4.1.2 See Also:

[set_test_verbosity](#)

15.5.5 Tests

15.5.5.1 test_equal

```
include std/unittest.e
public procedure test_equal(sequence name, object expected, object outcome)
```

Records whether a test passes by comparing two values.

15.5.5.1.1 Parameters:

1. name : a string, the name of the test
2. expected : an object, the expected outcome of some action
3. outcome : an object, some actual value that should equal the reference expected.

15.5.5.1.2 Comments:

- For floating point numbers, a fuzz of 1e-9 is used to assess equality.

A test is recorded as passed if equality holds between `expected` and `outcome`. The latter is typically a function call, or a variable that was set by some prior action.

While `expected` and `outcome` are processed symmetrically, they are not recorded symmetrically, so be careful to pass `expected` before `outcome` for better test failure reports.

15.5.5.1.3 See Also:

[test_not_equal](#), [test_true](#), [test_false](#), [test_pass](#), [test_fail](#)

15.5.5.2 test_not_equal

```
include std/unittest.e
public procedure test_not_equal(sequence name, object a, object b)
```

Records whether a test passes by comparing two values.

15.5.5.2.1 Parameters:

1. `name` : a string, the name of the test
2. `expected` : an object, the expected outcome of some action
3. `outcome` : an object, some actual value that should equal the reference `expected`.

15.5.5.2.2 Comments:

- For atoms, a fuzz of 1e-9 is used to assess equality.
- For sequences, no such fuzz is implemented.

A test is recorded as passed if equality does not hold between `expected` and `outcome`. The latter is typically a function call, or a variable that was set by some prior action.

15.5.5.2.3 See Also:

[test_equal](#), [test_true](#), [test_false](#), [test_pass](#), [test_fail](#)

15.5.5.3 test_true

```
include std/unittest.e
public procedure test_true(sequence name, object outcome)
```

Records whether a test passes.

15.5.5.3.1 Parameters:

1. name : a string, the name of the test
2. outcome : an object, some actual value that should not be zero.

15.5.5.3.2 Comments:

This assumes an expected value different from 0. No fuzz is applied when checking whether an atom is zero or not. Use [test_equal\(\)](#) instead in this case.

15.5.5.3.3 See Also:

[test_equal](#), [test_not_equal](#), [test_false](#), [test_pass](#), [test_fail](#)

15.5.5.4 assert

```
include std/unittest.e
public procedure assert(object name, object outcome)
```

Records whether a test passes. If it fails, the program also fails.

15.5.5.4.1 Parameters:

1. name : a string, the name of the test
2. outcome : an object, some actual value that should not be zero.

15.5.5.4.2 Comments:

This is identical to `test_true()` except that if the test fails, the program will also be forced to fail at this point.

15.5.5.4.3 See Also:

[test_equal](#), [test_not_equal](#), [test_false](#), [test_pass](#), [test_fail](#)

15.5.5.5 test_false

```
include std/unittest.e
public procedure test_false(sequence name, object outcome)
```

Records whether a test passes by comparing two values.

15.5.5.5.1 Parameters:

1. name : a string, the name of the test
2. outcome : an object, some actual value that should be zero

15.5.5.5.2 Comments:

This assumes an expected value of 0. No fuzz is applied when checking whether an atom is zero or not. Use [test_equal\(\)](#) instead in this case.

15.5.5.5.3 See Also:

[test_equal](#), [test_not_equal](#), [test_true](#), [test_pass](#), [test_fail](#)

15.5.5.6 test_fail

```
include std/unittest.e
public procedure test_fail(sequence name)
```

Records that a test failed.

15.5.5.6.1 Parameters:

1. name : a string, the name of the test

15.5.5.6.2 See Also:

[test_equal](#), [test_not_equal](#), [test_true](#), [test_false](#), [test_pass](#)

15.5.5.7 test_pass

```
include std/unittest.e
public procedure test_pass(sequence name)
```

Records that a test passed.

15.5.5.7.1 Parameters:

1. `name` : a string, the name of the test

15.5.5.7.2 See Also:

[test_equal](#), [test_not_equal](#), [test_true](#), [test_false](#), [test_fail](#)

16 Windows Routines

Windows Message Box
 Style Constants
 Return Value Constants
 Routines
Windows Sound

16.1 Windows Message Box

Style Constants

MB_ABORTRETRYIGNORE
MB_APPLMODAL
MB_DEFAULT_DESKTOP_ONLY
MB_DEFBUTTON1
MB_DEFBUTTON2
MB_DEFBUTTON3
MB_DEFBUTTON4
MB_HELP
MB_ICONASTERISK
MB_ICONERROR
MB_ICONEXCLAMATION
MB_ICONHAND
MB_ICONINFORMATION
MB_ICONQUESTION
MB_ICONSTOP
MB_ICONWARNING
MB_OK
MB_OKCANCEL
MB_RETRYCANCEL
MB_RIGHT
MB_RTLREADING
MB_SERVICE_NOTIFICATION
MB_SETFOREGROUND
MB_SYSTEMMODAL
MB_TASKMODAL
MB_YESNO
MB_YESNOCANCEL

Return Value Constants

IDABORT
IDCANCEL
IDIGNORE
IDNO
IDOK
IDRETRY

IDYES
 [get_active_id](#)
Routines
 [message_box](#)

16.1.1 Style Constants

Possible style values for `message_box()` style sequence

16.1.1.1 MB_ABORTRETRYIGNORE

```
include std/win32/msgbox.e  
public constant MB_ABORTRETRYIGNORE
```

Abort, Retry, Ignore

16.1.1.2 MB_APPLMODAL

```
include std/win32/msgbox.e  
public constant MB_APPLMODAL
```

User must respond before doing something else

16.1.1.3 MB_DEFAULT_DESKTOP_ONLY

```
include std/win32/msgbox.e  
public constant MB_DEFAULT_DESKTOP_ONLY
```

16.1.1.4 MB_DEFBUTTON1

```
include std/win32/msgbox.e  
public constant MB_DEFBUTTON1
```

First button is default button

16.1.1.5 MB_DEFBUTTON2

```
include std/win32/msgbox.e  
public constant MB_DEFBUTTON2
```

Second button is default button

16.1.1.6 MB_DEFBUTTON3

```
include std/win32/msgbox.e
public constant MB_DEFBUTTON3
```

Third button is default button

16.1.1.7 MB_DEFBUTTON4

```
include std/win32/msgbox.e
public constant MB_DEFBUTTON4
```

Fourth button is default button

16.1.1.8 MB_HELP

```
include std/win32/msgbox.e
public constant MB_HELP
```

Windows 95: Help button generates help event

16.1.1.9 MB_ICONASTERISK

```
include std/win32/msgbox.e
public constant MB_ICONASTERISK
```

16.1.1.10 MB_ICONERROR

```
include std/win32/msgbox.e
public constant MB_ICONERROR
```

16.1.1.11 MB_ICONEXCLAMATION

```
include std/win32/msgbox.e
public constant MB_ICONEXCLAMATION
```

Exclamation-point appears in the box

16.1.1.12 MB_ICONHAND

```
include std/win32/msgbox.e  
public constant MB_ICONHAND
```

A hand appears

16.1.1.13 MB_ICONINFORMATION

```
include std/win32/msgbox.e  
public constant MB_ICONINFORMATION
```

Lowercase letter i in a circle appears

16.1.1.14 MB_ICONQUESTION

```
include std/win32/msgbox.e  
public constant MB_ICONQUESTION
```

A question-mark icon appears

16.1.1.15 MB_ICONSTOP

```
include std/win32/msgbox.e  
public constant MB_ICONSTOP
```

16.1.1.16 MB_ICONWARNING

```
include std/win32/msgbox.e  
public constant MB_ICONWARNING
```

16.1.1.17 MB_OK

```
include std/win32/msgbox.e  
public constant MB_OK
```

Message box contains one push button: OK

16.1.1.18 MB_OKCANCEL

```
include std/win32/msgbox.e  
public constant MB_OKCANCEL
```

Message box contains OK and Cancel

16.1.1.19 MB_RETRYCANCEL

```
include std/win32/msgbox.e  
public constant MB_RETRYCANCEL
```

Message box contains Retry and Cancel

16.1.1.20 MB_RIGHT

```
include std/win32/msgbox.e  
public constant MB_RIGHT
```

Windows 95: The text is right-justified

16.1.1.21 MB_RTLREADING

```
include std/win32/msgbox.e  
public constant MB_RTLREADING
```

Windows 95: For Hebrew and Arabic systems

16.1.1.22 MB_SERVICE_NOTIFICATION

```
include std/win32/msgbox.e  
public constant MB_SERVICE_NOTIFICATION
```

Windows NT: The caller is a service

16.1.1.23 MB_SETFOREGROUND

```
include std/win32/msgbox.e  
public constant MB_SETFOREGROUND
```

Message box becomes the foreground window

16.1.1.24 MB_SYSTEMMODAL

```
include std/win32/msgbox.e  
public constant MB_SYSTEMMODAL
```

All applications suspended until user responds

16.1.1.25 MB_TASKMODAL

```
include std/win32/msgbox.e  
public constant MB_TASKMODAL
```

Similar to MB_APPLMODAL

16.1.1.26 MB_YESNO

```
include std/win32/msgbox.e  
public constant MB_YESNO
```

Message box contains Yes and No

16.1.1.27 MB_YESNOCANCEL

```
include std/win32/msgbox.e  
public constant MB_YESNOCANCEL
```

Message box contains Yes, No, and Cancel

16.1.2 Return Value Constants

possible values returned by MessageBox(). 0 means failure

16.1.2.1 IDABORT

```
include std/win32/msgbox.e  
public constant IDABORT
```

Abort button was selected.

16.1.2.2 IDCANCEL

```
include std/win32/msgbox.e  
public constant IDCANCEL
```

Cancel button was selected.

16.1.2.3 IDIGNORE

```
include std/win32/msgbox.e  
public constant IDIGNORE
```

Ignore button was selected.

16.1.2.4 IDNO

```
include std/win32/msgbox.e  
public constant IDNO
```

No button was selected.

16.1.2.5 IDOK

```
include std/win32/msgbox.e  
public constant IDOK
```

OK button was selected.

16.1.2.6 IDRETRY

```
include std/win32/msgbox.e  
public constant IDRETRY
```

Retry button was selected.

16.1.2.7 IDYES

```
include std/win32/msgbox.e  
public constant IDYES
```

Yes button was selected.

16.1.2.8 get_active_id

```
include std/win32/msgbox.e
public constant get_active_id
```

16.1.3 Routines

16.1.3.1 message_box

```
include std/win32/msgbox.e
public function message_box(sequence text, sequence title, object style)
```

Displays a window with a title, message, buttons and an icon, usually known as a message box.

16.1.3.1.1 Parameters:

1. `text`: a sequence, the message to be displayed
2. `title`: a sequence, the title the box should have
3. `style`: an object which defines which icon should be displayed, if any, and which buttons will be presented.

16.1.3.1.2 Returns:

An **integer**, the button which was clicked to close the message box, or 0 on failure.

16.1.3.1.3 Comments:

See [Style Constants](#) above for a complete list of possible values for `style` and [Return Value Constants](#) for the returned value. If `style` is a sequence, its elements will be or'ed together.

16.2 Windows Sound

```
SND_DEFAULT
SND_STOP
SND_QUESTION
SND_EXCLAMATION
SND_ASTERISK
sound
```


16.2.1 SND_DEFAULT

```
include std/win32/sounds.e
public constant SND_DEFAULT
```

16.2.1.1 SND_STOP

```
include std/win32/sounds.e
public constant SND_STOP
```

16.2.1.2 SND_QUESTION

```
include std/win32/sounds.e
public constant SND_QUESTION
```

16.2.1.3 SND_EXCLAMATION

```
include std/win32/sounds.e
public constant SND_EXCLAMATION
```

16.2.1.4 SND_ASTERISK

```
include std/win32/sounds.e
public constant SND_ASTERISK
```

16.2.1.5 sound

```
include std/win32/sounds.e
public procedure sound(atom sound_type = SND_DEFAULT)
```

Makes a sound.

16.2.1.5.1 Parameters:

1. sound_type: An atom. The type of sound to make. The default is SND_DEFAULT.

16.2.1.5.2 Comments:

The `sound_type` value can be one of ...

- `SND_ASTERISK`
- `SND_EXCLAMATION`
- `SND_STOP`
- `SND_QUESTION`
- `SND_DEFAULT`

These are sounds associated with the same Windows events via the Control Panel.

16.2.1.5.3 Example:

```
sound( SND_EXCLAMATION )
```

17 Release Notes

Version 4.0

- Bug Fixes
 - Changes
 - New Features
 - New Routines/Constants
-

17.1 Version 4.0

17.1.1 Bug Fixes

- 1855414. `open()` max path length is now determined by the underlying operating system and not a generic default. `open()` also now returns -1 when the filename is too long instead of causing a fatal error.
- 1608870. `dir()` now handles *.abc correctly, not showing a file ending with .abcd. `dir()` also now supports wildcard characters (* and ?) on all platforms.

17.1.2 Changes

- DOS support has been withdrawn. OpenEUPHORIA from version 4 onwards will not be specifically supporting DOS editions of the language.
- Comments may now be embedded in data passed to **value()** in **get.e**.
- **Moved** function **reverse()**, **sprint()** from **misc.e**, **lower()** and **upper()** from **wildcard.e** to **sequence.e**.
- **Moved** functions/constant **PI**, **arccos**, **arcsin** moved from **misc.e** to **math.e**.
- **Moved** **pretty_print()** from **misc.e** to **pretty.e**
- **Moved** **bin/syncolor.e** and **bin/keywords.e** to **include/euphoria**.
- **Moved** **instance()** and **sleep()** from **misc.e** to **os.e**
- **Moved** **sound()** and **tick_rate()** from **machine.e** to **os.e**
- **Moved** **free_console()** from **dll.e** to **console.e**
- **Moved** text mode graphic commands to **console.e**
- Documentation moved to a new format.
- **file.e** deprecated. **filesys.e**, **io.e** and **console.e** created in it's place. **file.e** still exists and includes the 3 replacements so old apps will still function.
- Switching between safe and normal memory routines no longer requires renaming files.

17.1.3 New Features

- New standard include files are in `include/std` to resolve many conflicts.
- Include file names with accent characters now supported.
- Enhanced symbol resolution to take into account information regarding which files were included by which files.
- Namespaces for a source file now can be used for identifiers in the specified file and for global identifiers in all files included by the specified file.
- Command line arguments for the translator allow for creating binaries with debugging symbols, and to specify a different runtime library.
- In trace mode, '?' will show the last defined variable of the requested name.
- Include directories can now be specified based on command line arguments and config files in addition to environment variables.
- Improved accuracy in scanning numbers in scientific notation. Scanned numbers are accurate to the full precision of the IEEE 754 floating point standard.
- **Translator:** Added **-lccopt-off** option for translator to disable using the optimization flag when compiling with LCC, as it has problems sometimes.
- New **loop do ... until** *condition* end loop construct, which differs from a while loop in that it performs its test at the end of the block, rather than at the start.
- New keywords to give greater control over the instruction flow:
 - ◆ **continue**: start next iteration of a loop;
 - ◆ **retry**: restarts the current iteration of a loop
 - ◆ **entry**: marks the entry point into a loop, skipping initial test
 - ◆ **break**: exit an if block or switch block
 - ◆ **goto**: jump to a label that is in the same scope
- The **exit**, **break**, **continue** and **retry** keywords now can take an optional parameter, which enables to exit several blocks at a time, or (re)starting an iteration of a loop which is not the innermost one.
- Block headers now may mention a label. This label can be used as the optional parameter of flow control keywords.
- Variables can now be initialised right on the spot at which they are declared, just like constants.
- Any routine parameter can be defaulted, i.e. given a default value that is plugged in if omitted on a call. Any expression can be used, and parameters of the same call can even be used.
- New **switch ... end switch** construct, which more efficiently implements a series of **elsif**, using the compact **case** statement.
- **Unit testing added to EUPHORIA.** Over 1,000 tests. To run, `cd tests` and type `eutest`
- Condition compiling keywords (**ifdef**, **elsifdef**, **end ifdef**) and **with define=xyz** or command line **-D XYZ** to insert/omit code in interpreter IL code and in translated C code.
- New enum keyword that allows for *parse time* sequential constant creation.
- The namespace **eu** is predefined, and can be used to fully qualify built-in routines.
- `with warning` has been enhanced in order to individually turn warnings on or off.

- New scope: **export**. Identifiers with the export scope can only be seen from files that:
 1. directly include the file where the identifiers are defined
- New scope: **public**. Identifiers with the public scope can only be seen from files that:
 1. directly include the file where the identifiers are defined
 2. directly include a file that uses the "public include file.e" construct to pass public identifiers
- Routine resolution changes
 1. Routines the same name as an internal no longer override the internal by default. You must use the keyword **override**.
 2. An unqualified call to routine that exists as an internal calls the internal unless overridden with the override keyword. global, public and export functions are not called. A namespace must be used.
- -STRICT option added that will display **all** warnings regardless of the file's with/without warning setting.
- -BATCH option designed to run in an automated environment. Causes any "Press Enter" type prompt due to error to be suppressed. Exit code will be 1 on success, 0 on failure as normal.
- -TEST option allows for editing/IDE environments to perform a syntax check on the euphoria code in question. Causes euphoria interpreter to do all parsing, syntax checking, etc... but does not execute the code. Exit code will be 1 on success, 0 on failure as normal. Editors/IDE's may need both -test and -batch.
- dis.ex (in the source directory) will parse a euphoria program and output the symbol table and the IL code in a readable format.
- Variables may be in any part of a routine, or in **for**, **while**, **if**, **loop** and **switch** blocks, in which case the scope of the variable ends when its block ends.

17.1.4 New Routines/Constants

- **value()** and **get()** in **get.e** may return additional information about the parsing activity.
- New builtin routines: **peeks()**, **peek2s()**, **peek2u()**, **peek_string()**, **poke2()**
- Began **library expansion**. A large percentage of these functions have been found in The Archive. Many thanks must go to the original developers who have contributed these functions.
- New routines in **error.e**: **crash()**.
- New routines in **math.e**: **ceil()**, **round_to()**, **round()**, **sign()**, **abs()**, **sum()**, **average()**, **min()**, **max()**, **deg2rad()**, **rad2deg()**, **log10()**, **atan2()**, **rand_range()**, **mod()**, **sinh**, **cosh**, **tanh**, **arcsinh**, **arccosh**, **arctanh**. Global Constants **E**, **HALFPI**, **TWOPI**, **QUARTPI**, **LN2**, **INVLOG2**, **EULER_GAMMA**, **EULER_NORMAL****.
- New routines in **sequence.e** and **search.e**: **insert()**, **splice()**, **remove()**, **replace()**, **head()**, **mid()**, **slice()**, **tail()**, **split()**, **split_adv()**, **join()**, **find_any()**, **find_any_from()**, **trim()**, **trim_head()**, **trim_tail()**, **truncate()**, **pad_head()**, **pad_tail()**, **chunk()**, **flatten()**, **find_all()**, **match_all()**, **find_replace()**, **vslice()**, **rfind()**, **rfind_from()**, **rmatch()**, **rmatch_from()**, **filter()**, **apply()**, **can_add**, **linear**, **repeat_pattern**, **project**, **extract**, **valid_index**, **find_nested****.
- New **sets.e** library to handle sets, set maps and set operations.
- New **stats.e** library for statistical and engineering math routines.
- New routines in **sort.e**: **sort_columns()**. All routines can use ascending or descending order.

- New routines in **fileSYS.e**, **io.e** and **os.e**: **read_lines()**, **read_file()**, **write_lines()**, **append_lines**, **write_file()**, **pathinfo()**, **dirname()**, **filename()**, **create_directory()**, **remove_directory()**, **file_exists()**, **copy_file()**, **rename_file()**, **delete_file()**, **move_file()**, **file_type()**, **file_length()**, **driveid()**, **fileext()**, **show_help()**, **cmd_parse()**, **uname()**, **is_win_nt()**, **setenv()**, **unsetenv()** and **Global Constants SLASH** and **CRLF****.
- New routines in **datetime.e**: **new()**, **from_date()**, **now()**, **weeks_day()**, **years_day()**, **to_unix()**, **from_unix()**, **add()**, **subtract()**, **diff()**. Global constants: **DT_YEAR**, **DT_MONTH**, **DT_DAY**, **DT_HOUR**, **DT_MINUTE**, **DT_SECOND**, **SECONDS**, **MINUTES**, **HOURS**, **DAYS**, **WEEKS**, **MONTHS**, **YEARS**. Global sequences: **month_names**, **month_abbrs**, **day_names**, **day_abbrs**, **ampm**. Note: These were made global for localization.
- New map routines in **map.e**: **new()**, **has()**, **get()**, **put()**, **remove()**, **size()**, **keys()**, **values()**, **hash()** (built-in), and more.
- New unit testing library: **unittest.e**: see tests/all.ex for examples on how to use with your own libraries.
- New stack routines in **stack.e**: **new()**, **is_empty()**, **push()**, **top()**, **pop()**, **clear()**, **swap()**, **dup()**. Global constants: **FIFO**, **FILO**.
- New locale routines in **locale.e**: **set()**, **get()**, **number()**, **money()**, **datetime()**, **set_po_path()**, **get_po_path()**, **po_load()**, **w()**.
- New os routines in **os.e**: **cmd_parse()**.
- New procedure **warning()** enables to specify custom warnings to be displayed at compile time, for instance to mark a routine as obsolete. The text of the warning must be a literal constant.
- Warnings also may be redirected or suppressed using **warning_file()**.
- New routines in **eds.e** (formerly **database.e**): **db_clear()**, **db_clear_table()**, **db_get_recid()**, **db_replace_recid()**, **db_current_table()**, **db_record_recid()**, **db_set_caching()**
- New **primes.e** library devoted to prime numbers
- New **regex.e** library on regular expressions, using the standard PCRE library
- New **socket.e** library for network/internet data interchange
- New character sets types and routines in **types.e**
- New **unicode.e** for encoding/decoding/storing strings in UTF16 format
- New **task.e** extending the multitasking capabilities of EUPHORIA
- New forward referencing: routines may be used before they are declared, and variables in other files may be referenced before they are declared (such as when two files include each other)
- New: Parser is smarter about identifying namespaces vs regular symbols
- New: Default namespace may be declared by a file.
- New: Some routines may be inlined for optimization. `with/without inline` give some control over which routines will be inlined to the coder.
- New: `with/without indirect_includes` help with using pre-eu4 code that uses global symbols with eu4 code, especially with respect to the respective standard libraries.
- New: `delete_routine()` and `delete()` for automatic garbage collection using user defined routines. Also, regexes automatically clean up after themselves.

18 Index

-p - define a pre-processor	-pf - force pre-processing
24 reasons why you are going to write your next program in euphoria!	:udt
?	ADD
ADDR_ADDRESS	ADDR_FAMILY
ADDR_FLAGS	ADDR_PROTOCOL
ADDR_TYPE	ADD_APPEND
ADD_PREPEND	ADD_SORT_DOWN
ADD_SORT_UP	ADLER32
AF_APPLETALK	AF_BTH
AF_INET	AF_INET6
AF_UNIX	AF_UNSPEC
ANCHORED	ANSI
ANY_UP	APPEND
ASCENDING	AT_EXPANSION
AUTO_CALLOUT	A_EXECUTE
A_WRITE	BAD_RECNO
BAD_SEEK	BINARY_MODE
BK_LEN	BK_PIECES
BLACK	BLINKING
BLOCK_CURSOR	BLUE
BMP_INVALID_MODE	BMP_OPEN_FAILED
BMP_SUCCESS	BMP_UNEXPECTED_EOF
BMP_UNSUPPORTED_FORMAT	BOM_8
BORDER_SPACE	BRIGHT_BLUE
BRIGHT_CYAN	BRIGHT_GREEN
BRIGHT_MAGENTA	BRIGHT_RED
BRIGHT_WHITE	BROWN
BSR_ANYCRLF	BSR_UNICODE
BYTES_PER_CHAR	BYTES_PER_SECTOR
CASELESS	CHILD
CMD_SWITCHES	COMBINE_SORTED
COMBINE_UNSORTED	CONCAT
COUNT_DIRS	COUNT_FILES
COUNT_SIZE	COUNT_TYPES
CS_FIRST	CYAN
C_BOOL	C_BYTE
C_CHAR	C_DOUBLE

C_DWORD	C_DWORDLONG
C_FLOAT	C_HANDLE
C_HRESULT	C_HWND
C_INT	C_LONG
C_LPARAM	C_POINTER
C_SHORT	C_SIZE_T
C_UBYTE	C_UCHAR
C_UINT	C_ULONG
C_USHORT	C_WORD
C_WPARAM	DBL_PTR
DB_EXISTS_ALREADY	DB_FATAL_FAIL
DB_LOCK_EXCLUSIVE	DB_LOCK_FAIL
DB_LOCK_NO	DB_LOCK_SHARED
DB_OK	DB_OPEN_FAIL
DEFAULT	DEGREES_TO_RADIANS
DESCENDING	DFA_RESTART
DFA_SHORTEST	DISPLAY_ASCII
DIVIDE	DNS
DNS_QUERY_ACCEPT_TRUNCATED_RESPONSE	DNS_QUERY_BYPASS_CACHE
DNS_QUERY_DONT_RESET_TTL_VALUES	DNS_QUERY_NO_HOSTS_FILE
DNS_QUERY_NO_LOCAL_NAME	DNS_QUERY_NO_NETBT
DNS_QUERY_NO_RECURSION	DNS_QUERY_NO_WIRE_QUERY
DNS_QUERY_RESERVED	DNS_QUERY_RETURN_MESSAGE
DNS_QUERY_STANDARD	DNS_QUERY_TREAT_AS_FQDN
DNS_QUERY_USE_TCP_ONLY	DNS_QUERY_WIRE_ONLY
DOLLAR_ENDONLY	DOS_TEXT
DOTALL	DUPNAMES
DUP_TABLE	D_ATTRIBUTES
D_DAY	D_HOUR
D_MINUTE	D_MONTH
D_NAME	D_SECOND
D_SIZE	D_YEAR
E	EOF
EOL	EOLSEP
ERROR_BADCOUNT	ERROR_BADMAGIC
ERROR_BADNEWLINE	ERROR_BADOPTION
ERROR_BADPARTIAL	ERROR_BADUTF8
ERROR_BADUTF8_OFFSET	ERROR_CALLOUT
ERROR_DFA_RECURSE	ERROR_DFA_UCOND
ERROR_DFA_UITEM	ERROR_DFA_UMLIMIT
ERROR_DFA_WSSIZE	ERROR_INTERNAL

ERROR_MATCHLIMIT	ERROR_NOMATCH
ERROR_NOMEMORY	ERROR_NOSUBSTRING
ERROR_NULL	ERROR_NULLWSLIMIT
ERROR_PARTIAL	ERROR_RECURSIONLIMIT
ERROR_UNKNOWN_NODE	ERROR_UNKNOWN_OPCODE
ERR_ACCESS	ERR_ADDRINUSE
ERR_ADDRNOTAVAIL	ERR_AFNOSUPPORT
ERR_AGAIN	ERR_ALREADY
ERR_CONNABORTED	ERR_CONNREFUSED
ERR_CONNRESET	ERR_DESTADDRREQ
ERR_FAULT	ERR_HOSTUNREACH
ERR_INPROGRESS	ERR_INTR
ERR_INVALID	ERR_IO
ERR_ISCONN	ERR_ISDIR
ERR_LOOP	ERR_MFILE
ERR_MSGSIZE	ERR_NAMETOOLONG
ERR_NETDOWN	ERR_NETRESET
ERR_NETUNREACH	ERR_NFILE
ERR_NOBUFS	ERR_NOENT
ERR_NOTCONN	ERR_NOTDIR
ERR_NOTINITIALISED	ERR_NOTSOCK
ERR_OPNOTSUPP	ERR_PROTONOSUPPORT
ERR_PROTOTYPE	ERR_ROFS
ERR_SHUTDOWN	ERR_SOCKTNOSUPPORT
ERR_TIMEOUT	ERR_WOULDBLOCK
ET_ERROR	ET_ERR_COLUMN
ET_ERR_LINE	ET_TOKENS
EULER_GAMMA	EXTENDED
EXTRA	EXT_COUNT
EXT_NAME	EXT_SIZE
E_ATOM	E_INTEGER
E_OBJECT	E_SEQUENCE
FALSE	FIFO
FILETYPE_DIRECTORY	FILETYPE_FILE
FILETYPE_NOT_FOUND	FILETYPE_UNDEFINED
FIRSTLINE	FLETCHER32
FP_FORMAT	FREEBSD
FREE_BYTES	FREE_RID
GET_EOF	GET_FAIL
GET_LONG_ANSWER	GET_NOTHING
GET_SHORT_ANSWER	GET_SUCCESS

GRAY	GREEN
HALFPI	HALFSQRT2
HALF_BLOCK_CURSOR	HAS_CASE
HAS_PARAMETER	HELP
HELP_RID	HOST_ALIASES
HOST_IPS	HOST_OFFICIAL_NAME
HOST_TYPE	HSIEH32
HTTP	HTTP_HEADER_ACCEPT
HTTP_HEADER_ACCEPTCHARSET	HTTP_HEADER_ACCEPTENCODING
HTTP_HEADER_ACCEPTLANGUAGE	HTTP_HEADER_ACCEPTRANGES
HTTP_HEADER_AUTHORIZATION	HTTP_HEADER_CACHECONTROL
HTTP_HEADER_CONNECTION	HTTP_HEADER_CONTENTLENGTH
HTTP_HEADER_CONTENTTYPE	HTTP_HEADER_DATE
HTTP_HEADER_FROM	HTTP_HEADER_GET
HTTP_HEADER_HOST	HTTP_HEADER_HTTPVERSION
HTTP_HEADER_IFMODIFIEDSINCE	HTTP_HEADER_KEEPALIVE
HTTP_HEADER_POST	HTTP_HEADER_POSTDATA
HTTP_HEADER_REFERER	HTTP_HEADER_USERAGENT
I/O	IDABORT
IDCANCEL	IDIGNORE
IDNO	IDOK
IDRETRY	IDYES
INDENT	INSERT_FAILED
INT_FORMAT	INVLN10
INVLN2	INVSQ2PI
IS_ATOM	IS_ATOM_DBL
IS_ATOM_INT	IS_DBL_OR_SEQUENCE
IS_SEQUENCE	LAST_ERROR_CODE
LEAVE	LEFT_DOWN
LEFT_UP	LINE_BREAKS
LINUX	LN10
LN2	LOCK_EXCLUSIVE
LOCK_SHARED	MAGENTA
MAKE_INT	MAKE_SEQ
MANDATORY	MAP_ANONYMOUS
MAP_FILE	MAP_FIXED
MAP_PRIVATE	MAP_SHARED
MAP_TYPE	MAX_ASCII
MAX_LINES	MB_ABORTRETRYIGNORE
MB_APPLMODAL	MB_DEFAULT_DESKTOP_ONLY
MB_DEFBUTTON1	MB_DEFBUTTON2

MB_DEFBUTTON3	MB_DEFBUTTON4
MB_HELP	MB_ICONASTERISK
MB_ICONERROR	MB_ICONEXCLAMATION
MB_ICONHAND	MB_ICONINFORMATION
MB_ICONQUESTION	MB_ICONSTOP
MB_ICONWARNING	MB_OK
MB_OKCANCEL	MB_RETRYCANCEL
MB_RIGHT	MB_RTLREADING
MB_SERVICE_NOTIFICATION	MB_SETFOREGROUND
MB_SYSTEMMODAL	MB_TASKMODAL
MB_YESNO	MB_YESNOCANCEL
MD5	MEM_COMMIT
MEM_RELEASE	MEM_RESERVE
MEM_RESET	MIDDLE_DOWN
MIDDLE_UP	MINF
MIN_ASCII	MISSING_END
MOVE	MSG_CONFIRM
MSG_CTRUNC	MSG_DONTROUTE
MSG_DONTWAIT	MSG_EOR
MSG_ERRQUEUE	MSG_FIN
MSG_MORE	MSG_NOSIGNAL
MSG_OOB	MSG_PEEK
MSG_PROXY	MSG_RST
MSG_SYN	MSG_TRUNC
MSG_TRYHARD	MSG_WAITALL
MULTILINE	MULTIPLE
MULTIPLY	M_ALLOC
M_FREE	NESTED_ALL
NESTED_ANY	NESTED_BACKWARD
NESTED_INDEX	NETBSD
NEWLINE_ANY	NEWLINE_ANYCRLF
NEWLINE_CR	NEWLINE_CRLF
NEWLINE_LF	NORMAL_ORDER
NOTBOL	NOTEMPTY
NOTEOL	NO_AT_EXPANSION
NO_AUTO_CAPTURE	NO_CASE
NO_CURSOR	NO_DATABASE
NO_HELP	NO_PARAMETER
NO_TABLE	NO_UTF8_CHECK
NO_VALIDATION	NO_VALIDATION_AFTER_FIRST_EXTRA
NS_C_ANY	NS_C_IN

NS_KT_DH	NS_KT_DSA
NS_KT_PRIVATE	NS_KT_RSA
NS_T_A	NS_T_A6
NS_T_AAAA	NS_T_ANY
NS_T_MX	NS_T_NS
NS_T_PTR	NULL
NULLDEVICE	NUMBER_OF_FREE_CLUSTERS
NUM_ENTRIES	OBJ_ATOM
OBJ_INTEGER	OBJ_SEQUENCE
OBJ_UNASSIGNED	OK
ONCE	OPENBSD
OPTIONAL	OPT_CNT
OPT_IDX	OPT_REV
OPT_VAL	OSX
PAGE_EXECUTE	PAGE_EXECUTE_READ
PAGE_EXECUTE_READWRITE	PAGE_EXECUTE_WRITECOPY
PAGE_NOACCESS	PAGE_NONE
PAGE_READ	PAGE_READONLY
PAGE_READWRITE	PAGE_READ_EXECUTE
PAGE_READ_WRITE	PAGE_READ_WRITE_EXECUTE
PAGE_SIZE	PAGE_WRITECOPY
PAGE_WRITE_COPY	PAGE_WRITE_EXECUTE_COPY
PARENT	PARTIAL
PATHSEP	PATH_BASENAME
PATH_DIR	PATH_DRIVEID
PATH_FILEEXT	PATH_FILENAME
PAUSE_MSG	PHI
PI	PID
PINF	PISQR
PRETTY_DEFAULT	PROT_EXEC
PROT_NONE	PROT_READ
PROT_WRITE	PUT
QUARTPI	RADIANS_TO_DEGREES
RD_INPLACE	RED
REVERSE_ORDER	RIGHT_DOWN
RIGHT_UP	ROTATE_LEFT
ROTATE_RIGHT	SCREEN
SD_BOTH	SD_RECEIVE
SD_SEND	SECTORS_PER_CLUSTER
SELECT_IS_ERROR	SELECT_IS_READABLE
SELECT_IS_WRITABLE	SELECT_SOCKET

SEQ_NOALT	SHA256
SHARED_LIB_EXT	SHOW_ONLY_OPTIONS
SIDE_NONE	SLASH
SLASHES	SMALLMAP
SM_TEXT	SND_ASTERISK
SND_DEFAULT	SND_EXCLAMATION
SND_QUESTION	SND_STOP
SOCKET_SOCKADDR_IN	SOCKET_SOCKET
SOCK_DGRAM	SOCK_RAW
SOCK_RDM	SOCK_SEQPACKET
SOCK_STREAM	SOL_SOCKET
SO_ACCEPTCONN	SO_BROADCAST
SO_CONNDATA	SO_CONNDATALEN
SO_CONNOPT	SO_CONNOPTLEN
SO_DEBUG	SO_DISCDATA
SO_DISCDATALEN	SO_DISCOPT
SO_DISCOPTLEN	SO_DONTLINGER
SO_DONTROUTE	SO_ERROR
SO_KEEPAIVE	SO_LINGER
SO_MAXDG	SO_MAXPATHDG
SO_OOBINLINE	SO_OPENTYPE
SO_RCVBUF	SO_RCVLOWAT
SO_RCVTIMEO	SO_REUSEADDR
SO_REUSEPORT	SO_SNDBUF
SO_SNDLOWAT	SO_SNDTIMEO
SO_SYNCHRONOUS_ALERT	SO_SYNCHRONOUS_NONALERT
SO_TYPE	SO_USELOOPBACK
SQRT2	SQRT3
SQRTE	START_COLUMN
STDERR	STDFLTR_ALPHA
STDIN	STDOUT
STRING_OFFSETS	ST_ALLNUM
ST_FULLPOP	ST_IGNSTR
ST_SAMPLE	ST_ZEROSTR
SUBTRACT	SUNOS
TDATA	TEST_QUIET
TEST_SHOW_ALL	TEST_SHOW_FAILED_ONLY
TEXT_MODE	TFORM
TF_ATOM	TF_HEX
TF_INT	THICK_UNDERLINE_CURSOR
TLNUM	TLPOS

TOTAL_BYTES	TOTAL_NUMBER_OF_CLUSTERS
TRUE	TTYPE
TWOPI	T_EOF
T_NULL	UNDERLINE_CURSOR
UNGREEDY	UNIX_TEXT
URL_ENTIRE	URL_HOSTNAME
URL_HTTP_DOMAIN	URL_HTTP_PATH
URL_HTTP_QUERY	URL_MAIL_ADDRESS
URL_MAIL_DOMAIN	URL_MAIL_QUERY
URL_MAIL_USER	URL_PASSWORD
URL_PATH	URL_PORT
URL_PROTOCOL	URL_QUERY_STRING
URL_USER	USED_BYTES
UTF	UTF8
UTF_16	UTF_16BE
UTF_16LE	UTF_32
UTF_32BE	UTF_32LE
UTF_8	VALIDATE_ALL
VC_COLOR	VC_COLUMNS
VC_LINES	VC_MODE
VC_NCOLORS	VC_PAGES
VC_SCRNCOLS	VC_SCRNLINES
VC_XPIXELS	VC_YPIXELS
WHITE	WIN32
WRAP	W_BAD_PATH
YEAR	YEARS
YELLOW	[:special statements
[:type_check	a public enum is available for convenient access to the returned configuration data:
a quick example	a return value of:
a small example	a trivial example
a type is expected here	a_execute
a_write	abort
abs	absolute_path
accept	accessing c structures
accessing c variables	accessing data
accessing euphoria coded routines	accessing euphoria internals
accessor constants	accumulation
add	add_append
add_item	add_prepend
add_sort_down	add_sort_up

add_to	adding to sequences
addr_address	addr_family
addr_flags	addr_protocol
addr_type	adler32
advanced examples	af_appletalk
af_bth	af_inet
af_inet6	af_unix
af_unspec	after user input, left margin problem
all_copyrights	all_left_units
all_matches	all_right_units
allocate	allocate_code
allocate_data	allocate_pointer_array
allocate_protect	allocate_string
allocate_string_pointer_array	allocate_wstring
allocating and writing to memory:	allow_break
amalgamated_sum	ampm
an example of parse options:	an important message for all c/c++ programmers...
anchored	and it provides alternate versions of:
and_bits	ansi
answer types	any_key
any_up	append
append(sequence s1, object o1)	append_lines
application definitions	apply
approx	arccos
arccosh	arcsin
arcsinh	arctan
arctanh	arithmetic operators
as a first programming language:	as an extension to languages you already know:
ascending	ascii_string
assembly file	assert
assignment statement	assignment with operator
assignment with operators	assumption:
at	at_expansion
atan2	atom
atom_to_float32	atom_to_float64
atoms and sequences	attr_to_colors
auto_callout	avedev
average	background
bad_recno	bad_seek
basic routines	basic set-theoretic operations.

basics	batch_command_line
begins	belongs_to
benchmark	binary_mode
binary_search	bind
binding and shrouding	bitmap handling
bits_to_int	bitwise operations
bk_color	bk_len
bk_pieces	black
blinking	block of key pointers
block of table headers	block_cursor
blue	bmp_invalid_mode
bmp_open_failed	bmp_success
bmp_unexpected_eof	bmp_unsupported_format
bom_16be	bom_16le
bom_32be	bom_32le
bom_8	boolean
border_space	bordered_address
branching statements	break
break statement	breakup
bright_blue	bright_cyan
bright_green	bright_magenta
bright_red	bright_white
brown	bsr_anycrlf
bsr_unicode	bug fixes
build_commandline	build_list
building sequences	builtins
but, my favorite language is ...	by default:
byte_range	bytes_per_char
bytes_per_sector	bytes_to_int
c compilers supported	c type constants
c_bool	c_byte
c_char	c_double
c_dword	c_dwordlong
c_float	c_func
c_handle	c_hresult
c_hwnd	c_int
c_long	c_lparam
c_pointer	c_proc
c_short	c_size_t
c_ubyte	c_uchar
c_uint	c_ulong

c_ushort	c_word
c_wparam	calc_hash
calc_primes	call
call-backs to your euphoria routines	call_back
call_func	call_proc
calling a function	calling a procedure
calling c functions	calling euphoria routines by id.
calling euphoria's internals	can_add
canon2win	canonical
canonical_path	cardinal
caseless	ceil
central_moment	cgi program hangs/no output
challenging	chance
change_target	changes
changing the shape of a sequence	char_count
char_test	character strings and individual characters
chars_before	chdir
check_all_blocks	check_break
check_calls	check_free_list
child	clear
clear_directory	clear_screen
client side only	close
cmd_parse	cmd_switches
code_length	color
colors	colors_to_attr
columnize	combine
combine_maps	combine_sorted
combine_unsorted	command line handling
command line options	command line switches
command-line options	command_line
comments	common internet routines
common problems	common problems and solutions
compare	comparison with earlier multitasking schemes
comparison with multithreading	compiled platform information
complex with/without options	compose_map
concat	concatenation of sequences and atoms - the '&' operator
configuration file format	connect
console	console window disappeared
constants	contain three serialized sequences:
continue	continue statement

contributors	conventions used in the manual
converting multiplies to adds in a loop	copy
copy_file	copyright information
core sockets	cos
cosh	count
count_dirs	count_files
count_size	count_types
crash	crash_file
crash_message	crash_routine
create	create/destroy
create_directory	creole markup
cross platform text graphics	cs_first
curdir	current authors
current_dir	cursor
cursor style constants	cursor type constants:
custom_sort	cutting and pasting
cyan	d_attributes
d_day	d_hour
d_minute	d_month
d_name	d_second
d_size	d_year
data execute mode	data structures
data type conversion	database file format
date	date/time
datetime	day_abbrs
day_names	days_in_month
days_in_year	db_cache_clear
db_clear_table	db_close
db_compress	db_create
db_create_table	db_current
db_current_table	db_delete_record
db_delete_table	db_dump
db_exists_already	db_fatal_fail
db_fatal_id	db_fetch_record
db_find_key	db_get_errors
db_get_recid	db_insert
db_lock_exclusive	db_lock_fail
db_lock_no	db_lock_shared
db_ok	db_open
db_open_fail	db_record_data
db_record_key	db_record_recid

db_rename_table	db_replace_data
db_replace_recid	db_select
db_select_table	db_set_caching
db_table_list	db_table_size
deallocate	debugging
debugging and profiling	decanonical
declarations	decode
default	defaulted_value
defaulttext	define_c_func
define_c_proc	define_c_var
define_map	define_operation
defining the scope of an identifier	definition
deg2rad	degrees_to_radians
delete	delete_file
delete_routine	delta
dep_on	dep_really_works
dep_works	dequote
descending	deserialize
dfa_restart	dfa_shortest
diagram_commutes	diff
difference	dir
dir_size	direct_map
directory handling	dirname
disclaimer	disclaimer:
discover euphoria	disk_metrics
disk_size	display
display_ascii	display_text_image
distributes_over	distributing a program
divide	dll/shared library interface
dns_query_accept_truncated_response	dns_query_bypass_cache
dns_query_dont_reset_ttl_values	dns_query_no_hosts_file
dns_query_no_local_name	dns_query_no_netbt
dns_query_no_recursion	dns_query_no_wire_query
dns_query_reserved	dns_query_return_message
dns_query_standard	dns_query_treat_as_fqdn
dns_query_use_tcp_only	dns_query_wire_only
documentation software	documentation tags
dollar_endonly	dos_text
dotall	driveid
dump	dup
dup_table	dupnames

dynamic link libraries (shared libraries)

dyncall

e_atom

e_object

ed - euphoria editor

edit sample files

embed_union

emovavg

encode

ensure_in_list

enum

eof

eolsep

equality

err_addrinuse

err_afnosupport

err_already

err_connrefused

err_destaddrreq

err_hostunreach

err_intr

err_io

err_isdir

err_mfile

err_nametoolong

err_netreset

err_nfile

err_noent

err_notdir

err_notsock

err_protonosupport

err_rofs

err_socktnosupport

err_wouldblock

error constants

error status constants

error_badmagic

error_badoptio

error_badutf8

error_callout

error_dfa_recurse

dynamic linking to external code

e

e_integer

e_sequence

edges_only

editing a program

embedding

emptyseq

ends

ensure_in_range

environment.

eol

equal

err_access

err_addrnotavail

err_again

err_connaborted

err_connreset

err_fault

err_inprogress

err_inval

err_isconn

err_loop

err_msgsize

err_netdown

err_netunreach

err_nobufs

err_notconn

err_notinitialised

err_opnotsupp

err_prototype

err_shutdown

err_timedout

error code constants

error information

error_badcount

error_badnewline

error_badpartial

error_badutf8_offset

error_code

error_dfa_ucond

error_dfa_uitem
error_dfa_wssize
error_matchlimit
error_no
error_nomemory
error_null
error_partial
error_unknown_node
errors and warnings
escape commands
et_err_column
et_error
et_keep_blanks
et_string_numbers
et_tokenize_string
eu.cfg
euler_gamma
euphoria database (eds)
euphoria has qualities that go beyond the elegance of sequences
euphoria internals
euphoria license
euphoria programs
euphoria to c translator
euphoria version definitions
eptest - unit testing
example programs
executable size and compression
exit statement
expected to see...
ext_count
ext_size
external euphoria type constants
extract
fallthru
fetch
fiber_product
file handling
file reading/writing
file types
error_dfa_umlimit
error_internal
error_message
error_nomatch
error_nosubstring
error_nullwslimit
error_recursionlimit
error_unknown_opcode
escape
escaped characters
et_err_line
et_error_string
et_keep_comments
et_tokenize_file
et_tokens
eudoc - source documentation tool
euphoria credits
euphoria database system (eds)
euphoria information
euphoria is unique
euphoria programming language v4.0
euphoria routines
euphoria trouble-shooting guide
euphoria_copyright
example :
exec
exit
exp
expressions
ext_name
extended
extra
extracting, removing, replacing from/into a sequence
false
fiber_over
fifo
file name parsing
file system
file_exists

file_length	file_number
file_position	file_timestamp
file_type	filebase
fileext	filename
filetype_directory	filetype_file
filetype_not_found	filetype_undefined
filter	find
find/match	find_all
find_any	find_from
find_nested	find_replace
find_replace_callback	find_replace_limit
finding	finish_datesub.ex
first, euphoria delivers the "expected" features of a modern language:	firstline
flatten	fletcher32
float32_to_atom	float64_to_atom
floating-point calculations not exact	floor
flow control	flow control statements
flush	for
for instance:	for statement
for_each	formal syntax
format	fp_format
frac	free
free list	free_bytes
free_code	free_console
free_pointer_array	free_rid
freebsd	frequently asked questions
from_date	from_unix
functions	further notes
gcd	general behavior
general notes	general routine reference
general routines	general tips
general use	generic documentation
geomean	get
get_active_id	get_bytes
get_char	get_charsets
get_def_lang	get_dstring
get_encoding_properties	get_eof
get_fail	get_http
get_http_use_cookie	get_integer16
get_integer32	get_key

get_lang_path	get_long_answer
get_mouse	get_nothing
get_option	get_ovector_size
get_page_size	get_pid
get_position	get_rcvheader
get_screen_char	get_sendheader
get_seqlen	get_short_answer
get_success	get_text
get_url	getc
getenv	getlasterror_rid
getnumericvalue	gets
getsysteminfo_rid	getting a routine identifier
goto	goto statement
graphics	graphics - cross platform
graphics - image routines	graphics modes
graphics_mode	graphics_point
gray	green
half_block_cursor	halfpi
halfsqrt2	harder
harmean	has
has_case	has_inverse
has_match	has_parameter
has_unit	hash
hashing algorithms	head
header	header labels
header management	hello, world
help	help_rid
hex_text	high-level win32 programming
host_aliases	host_by_addr
host_by_name	host_ips
host_official_name	host_type
how can i make my program run even faster?	how does storage get recycled?
how much of a speed-up should i expect?	how to access the data
how to run the translator	how to speed-up loops
how to uninstall euphoria	how to use this file:
hsieh32	http_header_accept
http_header_acceptcharset	http_header_acceptencoding
http_header_acceptlanguage	http_header_acceptranges
http_header_authorization	http_header_cachecontrol
http_header_connection	http_header_contentlength
http_header_contenttype	http_header_date

http_header_from	http_header_get
http_header_host	http_header_httpversion
http_header_ifmodifiedsince	http_header_keepalive
http_header_post	http_header_postdata
http_header_referer	http_header_useragent
hyperbolic trigonometry	idabort
idcancel	identifiers
idignore	idno
idok	idretry
idyces	if
if statement	ifdef statement
image	in-lining of routine calls
include	include_paths
included tools	inclusion and belonging.
indent	index
index block	indirect calling a routine coded in euphoria
indirect routine calling	indirect_includes
information	init_class
init_curdir	input routines
insert	insert(sequence in_what, object what, atom position)
insert_failed	insertion_sort
installation	installing euphoria
instance	int_format
int_to_bits	int_to_bytes
intdiv	integer
integer_array	interacting with the os
interfacing with c code (win32, linux, freebsd)	interpreter vs. translator
intersection	introduction
invln10	invln2
invsq2pi	ip address handling
is_associative	is_bijective
is_dep_supported	is_empty
is_even	is_even_obj
is_in_list	is_in_range
is_inetaddr	is_injective
is_leap_year	is_match
is_page_aligned_address	is_subset
is_surjective	is_symmetric
is_unit	is_using_dep
is_valid_memory_protection_constant	is_win_nt

isalpha	isalphanum
isblock	iscontrol
iscurrency	isdigit
isdirltr	isdirneutral
isdirrtl	isdirseparator
isdirstrong	isdirweak
isdirwhitespace	isformat
isgraph	isline
islower	ismark
ismirroring	isnonbreaking
isnonspacing	isnumber
isother	isparagraph
isprint	isprivate
ispunctuation	issegment
isseparator	isspace
issurrogate	issymbol
istitle	isuchar
isupper	join
kernel_dll	keyboard related routines
keys	keyvalues
keyword data	keywords
kill	kurtosis
lang_load	language reference
largest	last
last_error_code	leave
left_down	left_up
legal restrictions	length
length(sequence s)	library definitions
library routines	licensing
line terminator	line_breaks
linear	linux
linux and freebsd	listen
ln10	ln2
load	load_map
locale name translation	locale routines
locale_canonical	localized variables
locate_file	lock type constants
lock_exclusive	lock_file
lock_shared	lock_type
log	log10
logarithms and powers.	logical operators

long lines	lookup
loop	loop statements
loop until statement	low level file/device handling
low level routines	low-level win32 programming
lower	m_alloc
m_free	machine:free
machine_addr	machine_func
machine_proc	magenta
major	malloc
managing databases	managing records
managing tables	mandatory
manually editing environment variables	map
map (hash table)	map_anonymous
map_file	map_fixed
map_private	map_shared
map_type	mapping
maps between sets.	match
match_all	match_from
match_replace	matches
matching	math
math constants	max
max_ascii	max_lines
maximum file size	maybe_any_key
mb_abortretryignore	mb_applmodal
mb_default_desktop_only	mb_defbutton1
mb_defbutton2	mb_defbutton3
mb_defbutton4	mb_help
mb_iconasterisk	mb_iconerror
mb_iconexclamation	mb_iconhand
mb_iconinformation	mb_iconquestion
mb_iconstop	mb_iconwarning
mb_ok	mb_okcancel
mb_retrycancel	mb_right
mb_rtlreading	mb_service_notification
mb_setforeground	mb_systemmodal
mb_taskmodal	mb_yesno
mb_yesnocancel	md5
measuring performance	median
mem_commit	mem_copy
mem_release	mem_reserve
mem_reset	mem_set

memdll_id	memory allocation
memory disposal	memory management - low-level
merge	message translation functions
message_box	microsoft windows memory protection constants
mid	middle_down
middle_up	min
min_ascii	minf
mini-guides	minsize
miscellaneous	missing_end
mixture	mlock
mmap	mod
mode	modify the tutorial programs
money	month_abbrs
month_names	mouse
mouse_events	mouse_pointer
movavg	move
move_file	mprotect
msg_confirm	msg_ctrunc
msg_dontroute	msg_dontwait
msg_eor	msg_errqueue
msg_fin	msg_more
msg_nosignal	msg_oob
msg_peek	msg_proxy
msg_rst	msg_syn
msg_trunc	msg_tryhard
msg_waitall	multi-tasking
multiline	multiple
multiply	multitasking in euphoria
munlock	munmap
my_dir	nested_all
nested_any	nested_backward
nested_get	nested_index
nested_put	netbsd
networking routines	new
new features	new routines/constants
new_extra	new_from_kvpairs
new_from_string	new_time
newline_any	newline_ancrlf
newline_cr	newline_crlf
newline_lf	next_prime
no_at_expansion	no_auto_capture

no_case
no_database
no_parameter
no_utf8_check
no_validation_after_first_extra
normal_order
notbol
note:
noteol
notes:
now_gmt
ns_c_in
ns_kt_dsa
ns_kt_rsa
ns_t_a6
ns_t_any
ns_t_ns
null
num_entries
number to a string?
number_of_free_clusters
obj_atom
obj_sequence
object
ok
once
only a subset of the format specification is currently supported:
open
openbsd
operating system constants
operation
operations on sequences
opt_cnt
opt_rev
optimize
option records have the following structure:
optional
options detail
or_all
osx
no_cursor
no_help
no_table
no_validation
non-text files
not_bits
note to linux, freebsd and osx users:
notempty
notes
now
ns_c_any
ns_kt_dh
ns_kt_private
ns_t_a
ns_t_aaaa
ns_t_mx
ns_t_ptr
nulldevice
number
number_array
numeric version information
obj_integer
obj_unassigned
objects
on/off options
only 10 significant digits during printing
only first character in printf
open_dll
opening/closing
operating system helpers
operation codes for put
operations on sets
opt_idx
opt_val
option constants
option_switches
options
options:
or_bits
other examples of pre-processors include:

other ideas
others
pad_head
page_aligned_address
page_execute_read
page_execute_writecopy
page_none
page_read_execute
page_read_write_execute
page_readwrite
page_write_copy
page_writecopy
parameter:
parse
parse_commandline
parse_querystring
parse_url
partial
patch
path_dir
path_fileext
pathinfo
pause_msg
peek
peek2u
peek4u
peek_string
peek_wstring
performance tips
pi
pinf
pisqr
pixel-graphics tips
platform definitions
platform specific issues
platform_locale
platforms:
poke2
poke_string
pop
positive_int
other operations on sequences
outputs:
pad_tail
page_execute
page_execute_readwrite
page_noaccess
page_read
page_read_write
page_readonly
page_size
page_write_execute_copy
pairs
parent
parse options:
parse_ip_address
parse_recvheader
parsing
past authors
path_basename
path_driveid
path_filename
pathsep
pcre_copyright
peek2s
peek4s
peek_end
peek_top
peeks
phi
pid
pipe input/output
pivot
platform
platform issues
platform:
platform_name
poke
poke4
poke_wstring
position
posix locale names:

posix_names	possible modes are:
possible problems	power
powof2	pre-process details
precedence chart	predefined character sets:
prepare_block	prepend
prepend(sequence s1, object o1)	press enter
pretty printing	pretty_default
pretty_print	pretty_sprint
prime numbers	prime_list
print	printf
procedures	process
process_lines	product
product_map	products
profile_file	profiling
program has no errors, no output	project
prompt_number	prompt_string
proper	prot_exec
prot_none	prot_read
prot_write	pseudo memory
push	put
put_integer16	put_integer32
put_screen_char	puts
q_print	quartpi
quick start	quote
rad2deg	radians_to_degrees
ram_space	rand
rand_range	random numbers
range	raw_frequency
rd_inplace	read
read the manual	read/write ports?
read/write process	read/write routines
read_bitmap	read_file
read_lines	reading from, writing to, and calling into memory
recalling previous strings	receive
receive_from	red
redefine my for-loop variable?	regex
register_block	regular expressions
rehash	relational operators
release notes	remainder
remove	remove_all
remove_directory	remove_dups

remove_from
remove_subseq
repeat
repeat_pattern
replace_all
reporting
restrict
retry
return a description of the current video configuration:

return value:

reverse
reverse_map
rfind
right_up
rnd
roll
rotate_bits
rotate_right
roundings and remainders
routine not declared, my file
routines
run the tutorial programs
running under windows
safe mode
safe_address
sample
save_map
saving results in variables
scope
scroll
sd_receive
searching
sectors_per_cluster
see also:
select
select_is_error
select_is_writable
send
send_to

remove_item
rename_file
repeat(object o1, integer times)
replace
replacement
requirements
retain_all
retry statement
return value constants
returns an object with the same value as x. x must be with in the integer range of a legal euphoria integer type.
reverse mappings
reverse_order
right_down
rmatch
rnd_1
rotate
rotate_left
round
routine not declared
routine_id
run the demo programs
running a program
safe memory access
safe.e
safe_address_list
save_bitmap
save_text_image
scalability
screen
sd_both
sd_send
section
security / multi-user access
seek
select accessor constants
select_is_readable
select_socket
send flags
seq_noalt

sequence
sequence manipulation
sequence_array
sequences_to_map
serialize
server side only
service_by_port
set_accumulate_summary
set_colors
set_def_lang
set_encoding_properties
set_option
set_sendheader
set_sendheader_useragent_msie
set_test_verbosity
setenv
setup routines
shared_lib_ext
short-circuit evaluation
show_help
shuffle
shutdown options
side effects:
sign
signature:
simple
sinh
skewness
slashes
slice
sm_text
smallest
snd_asterisk
snd_exclamation
snd_stop
so_broadcast
so_conndatalen
so_connoptlen
so_discdata
so_discopt
so_dontlinger
sequence centric routines
sequence-formation
sequence_to_set
serialization of euphoria objects
server and client sides
service_by_name
set
set_charsets
set_decimal_mark
set_default_charsets
set_lang_path
set_rand
set_sendheader_default
set_test_abort
set_wait_on_summary
sets
sha256
shift_bits
show_block
show_only_options
shutdown
side effect:
side_none
sign and comparisons
sim_index
sin
size
slash
sleep
slicing of sequences
small
smallmap
snd_default
snd_question
so_acceptconn
so_conndata
so_connopt
so_debug
so_discdatalen
so_discoptlen
so_dontroute

[so_error](#)
[so_linger](#)
[so_maxpathdg](#)
[so_opentype](#)
[so_rcvlowat](#)
[so_reuseaddr](#)
[so_sndbuf](#)
[so_sndtimeo](#)
[so_synchronous_nonalert](#)
[so_uselookback](#)
[sock_raw](#)
[sock_seqpacket](#)
[socket](#)
[socket type constants](#)
[socket_socket](#)
[some example programs to look at:](#)
[some special case optimizations](#)
[sort](#)
[sorting](#)
[source code](#)
[source tokenizer](#)
[special regex characters are:](#)
[special top-level statements](#)
[splice](#)
[split](#)
[split_limit](#)
[sprint](#)
[sqrt](#)
[sqrt3](#)
[st_allnum](#)
[st_ignstr](#)
[st_zerostr](#)
[standard library memory protection constants](#)
[statements](#)
[std_library_address](#)
[stdev](#)
[stdin](#)
[store](#)
[string centric routines](#)
[so_keepalive](#)
[so_maxdg](#)
[so_oobinline](#)
[so_rcvbuf](#)
[so_rcvtimeo](#)
[so_reuseport](#)
[so_sndlowat](#)
[so_synchronous_alert](#)
[so_type](#)
[sock_dgram](#)
[sock_rdm](#)
[sock_stream](#)
[socket options](#)
[socket_sockaddr_in](#)
[sol_socket](#)
[some further notes on time profiling](#)
[some typical devices that you can open on windows are:](#)
[sort_columns](#)
[sound](#)
[source documentation](#)
[special keys](#)
[special replacement operators:](#)
[specifying the type of a variable](#)
[splice\(sequence in_what, object what, atom position\)](#)
[split_any](#)
[splitting](#)
[sprintf](#)
[sqrt2](#)
[sqre](#)
[st_fullpop](#)
[st_sample](#)
[stack](#)
[start_column](#)
[statistics](#)
[stderr](#)
[stdfltr_alpha](#)
[stdout](#)
[string](#)
[string to a number?](#)

string version information
structure of an eds database
subscripting of sequences
subtract
sum_central_moments
sunos
swap
switch statement
synopsis for creating report from the log
synopsis:
syntaxcolor
system_exec
t_alpha
t_boolean
t_cntrl
t_digit
t_eof
t_identifier
t_null
t_punct
t_specword
t_upper
t_xdigit
tail
tan
task_clock_start
task_create
task_list
task_self
task_suspend
tdata
test_equal
test_fail
test_not_equal
test_quiet
test_report
test_show_failed_only
test_write
text manipulation
text_mode
tf_atom
string_offsets
style constants
subsets
sum
summary
support functions
switch
synopsis
synopsis for running the tests
syntax coloring
system
t_alnum
t_ascii
t_bytearray
t_consonant
t_display
t_graph
t_lower
t_print
t_space
t_text
t_vowel
table header
taking advantage of cache memory
tanh
task_clock_stop
task_delay
task_schedule
task_status
task_yield
temp_file
test_exec
test_false
test_pass
test_read
test_show_all
test_true
tests
text_color
text_rows
tf_hex

tf_int	tform
the attributes element is a string sequence containing characters chosen from:	the bind command
the c representation of a euphoria object	the c representations of a euphoria sequence and a euphoria double
the complete set of resolution rules	the current database.
the current table.	the equation for average is:
the equation for standard deviation is:	the error control files
the euphoria data structures	the euphoria representation of a euphoria object
the eutest program	the following flags are available to fine tune the search:
the override qualifier	the paths are ordered in the order they are searched:
the result will be one of the following codes:	the return codes are:
the shroud command	the trace file
the trace screen	the unit test files
the unix platforms	the user defined pre-processor
the visibility of public and export identifiers	the win32 platform
the with/without trace directive	thick_underline_cursor
threshold	time
time/number translation	tlnum
tlpos	to_integer
to_number	to_unix
token types recognized on the command line:	tolower
tomirror	top
total_bytes	total_number_of_clusters
totitle	toupper
toutf	trace
trace a demo	trailer
transform	translate
trigonometry	trim
trim_head	trim_tail
trsprintf	true
true_color	trunc
ttype	twopi
type_of	types
types - extended	types of maps
types of tasks	types supporting memory
udp only	uname
underline_cursor	ungreedy
unicode	union
unit testing framework	unix_text

unknown escape character
unregister_block
upper
url handling
url_entire
url_http_domain
url_http_query
url_mail_domain
url_mail_user
url_path
url_protocol
url_user
use of a configuration file
used_bytes
using ifdef
using namespaces
utf
utf_16
utf_16le
utf_32be
utf_8
valid
valid_index
valid_wordsize
validate_all
values
vc_color
vc_lines
vc_ncolors
vc_scrncols
vc_xpixels
version
version_major
version_patch
version_string
version_string_short
video_config
virtualalloc_rid
virtuallock_rid
virtualunlock_rid
vslice
unlock_file
unsetenv
url encoding and decoding
url parsing
url_hostname
url_http_path
url_mail_address
url_mail_query
url_password
url_port
url_query_string
usage notes
use of tabs
using euphoria
using machine code and c
using the euphoria to c translator
utf8
utf_16be
utf_32
utf_32le
utility routines
valid operations are:
valid_memory_protection_constant
validate
value
variables
vc_columns
vc_mode
vc_pages
vc_scrnlines
vc_ypixels
version 4.0
version_minor
version_revision
version_string_long
version_type
video_config sequence accessors
virtualfree_rid
virtualprotect_rid
vlookup
w32_name_canonical

w32_names	w_bad_path
wait_key	walk_dir
warning	warning:
warning: use the right file mode	warning_file
watch out for the following common mistake:	web interface
weeks_day	what if i want to change the compile or link options in emake.bat?
what to do?	where
while	while statement
white	why call indirectly?
why multitask?	why scopes, and what are they?
wildcard matching	wildcard_file
wildcard_match	win32
windows	windows message box
windows routines	windows sound
with / without	with / without define
with / without inline	with entry statement
with/without warning	with_batch
with_define	with_inline
wrap	write
write your own	write_file
write_lines	writeln
writeln	xor_bits
year	years
years_day	yellow
yet another programming language?	you can generate these files by
you can refer to the elements of an entry with the following constants:	you can use this to run things that might be difficult to quote out: