| See also: AL - Perl Perl @ Wikipedia Perl.org PerlMonks.org O'Reilly Books Perl mailing lists | Perl Intro - a quick introduction to Perl. PerlCheat , Learn Perl in Y minutes, or in 2 hours 30 minutes Online Perl books and tutorials : Beginning Perl , Modern Perl (html) , Perl Maven Tutorial Perl Cookbook or (PLEAC Perl: list of Perl code solutions) Learning Perl LPor, Intermediate Perl or , Mastering Perl or , Effective Perl Programming or Other exist but are not recommended for various reasons. | | | | |
|---|--|--|---|---|--|
| Perl Guidelines and tools | Perl Style Guide, 10 Essential Development Practices, Books: Perl Best Practices o, Modern Perl Best Practices (course) o perlcritic script uses Perl::Critic to scan Perl code. The pel-perl-critic command invokes it to check code in buffer. The perltidy application reformats Perl code. Older perltidy home page. PerlTidy @ Wikipedia, PBP recommended .perltidyrc | | | | |
| • In Emacs: C-c C-h F | peridoc: about peridoc itself peritoc: table of content: names of all pages perisyn: Peri syntax perifunc: Peri built-in functions | Use period to find if a Perl module is installed, a period local::lib prints the document period -Mlocal::lib is useful to get module | ation of local::lib if it is in | stalled. | |
| CPAN (@ Wikipedia) • Search CPAN — meta::cpan | The Zen of Comprehensive Archive Networks PAUSE - Perl Authors Upload Server Installing Local Perl Modules with CPAN | Command line tools interacting with CPAN to instate cpan: (requires config. but has defaults). Use lotate Type cpan to open the cpan shell, then type incomplus, or cpanminus: cpanm: (no config recomplus). | cal::lib; cpan will be able nstall The::Modulet | to install into your ~/perl5 tree. to install packages. | |

Perl scripts

| Writing Perl scripts | Impose strictures in Perl files t | Impose strictures in Perl files to prevent errors by adding one of the following use lines. Also see the strictures package. | | |
|--|--|--|--|--|
| Use the following at the beginning of Perl script files. perldiag @ perldoc | <pre>#!/usr/bin/env perl use strict; use warnings; # for testing only:</pre> | #! /usr/bin/perl -w use v5.12; # loads strict use v5.35; # &loads warnings A use diagnostics produces more info but increases startup time. | Executable Perl script should have a valid shebang-line identifying the appropriate-location of the Perl interpreter. It may have to be modified at installation time (OpenGroup/SUS). It's best to: use warnings; perl-w generates warning for all Perl code in the program including modules used by the program. Also use the goption to check syntax. But most Perl code should also activate the strict Perl rules and warnings to detect warnings. See: Barewords in Perl | |
| | <pre>use diagnostics;</pre> | Alternative: perl -Mdiagnostics . Emacs p | el-perl-critic command can report diagnostic. | |
| use version/features | <u>use</u> v5.36; | This can be used to enable both the strict and warning pramas as well as several <u>named features</u> . • See the <u>table listing the feature bundles per Perl versions</u> . | | |

```
Perl 5 Operators
                              Perl has a large number of operators, listed below with their precedence and associativity.

• <u>C Operators missing from Perl</u>: unary &, unary * and (type)

• <u>Quote and Quote-like operators</u>: in Perl quotes are operators and they provide various kind of interpolating and pattern matching capabilities.
Perl 5 Operators
                    Note:
Associativity: one of:
                              left
                                            terms and list operators (leftward)
                                                                                          ( )
rightleft
                              left
                                            Arrow Operator:
                              NA
                                            Auto-increment and Auto-decrement: ++ --
• NA : not associative:
                              right
                                            Exponentiation:
  cannot use more than
                                            Symbolic Unary Operators:
                                                                                                  -. \ and unary + and -
                              right
                                                                                                                                                Note: The operator \ <u>creates a reference</u>. See <u>example</u>.
  one of these operators
                              left.
                                            Binding operators:
                                                                                          =~ !~
in sequence.

CH: chained
                                                                                         * /
                              left
                                            Multiplicative Operators:
                                                                                                 용
                                                                                                       x
                              left
                                            Additive Operators:
                              left.
                                            Shift Operators:
                                                                                          <<
                                                                                                >>
                              NA
To get this information,
                                            named unary operators
                              NA
                                            Class instance Operator:
perldoc perlop
                              CH
                                            Relational Operators:
                                                                                         as numbers: < >
                                                                                                                                     as strings: 1t
                                                                                                                                                           gt
                                                                                                                                                                   le
                              CH/NA
                                                                                         as numbers: == != <=>
                                            Equality Operators:
                                                                                                                                     as strings: eq
                                                                                                                                                           ne
                                                                                                                                                                   cmp
                              left.
Note: or The
                                            Bitwise And:
                                                                                        & &.
                                            Bitwise Or and Exclusive Or:
Bitwise String Operators
                              1eft
                                                                                             |.
                              left
are:
                                            C-style Logical And:
                                                                                        &&
                                                                                              ^^
                              left
                                                                                                    11
                                            Logical Defined-Or:
                                                                                        Ш
              | · · · · | · · · · |
      & .=
                              NA
                                            Range Operators:
                              right
                                            Conditional Operator:
                                                                                        ?:
                              right
                                            Assignment Operators:
                                                                                                              x=
                                                                                        goto last next redo dump
                              left
                                            Comma, fat-comma Operators:
                              NA
                                            <u>list operators (rightward)</u>
                              right
                                            Logical Not:
                                                                                       not
                                            Logical And:
                              left
                                            Logical or and Exclusive or:
                                                                                       or xor
                                            Converts a string that starts with digits into a number.
                                                                                                                                                             -+- is essentially - + - or - - but a + to allow placing
                                                                                                                print -+- '22les poulets!';
trick operators 🔔
                                                                                                                                                             them together. The 0+ does the same as -+- but
                              0+
                                                                                                                 # prints 22
Do not use in
                                                                                                                                                             the second has higher precedence.
production code!
But understanding how
                                            Called the 'goatse' operator. It causes the right side
                                                                                                                my $str = "A 22 before 33 does not make 9, it is 44!";
                              =()=
these work does help
                                                                                                                my $digit_count =()= $str =~ /\d/g;
print "$digit_count"; # prints
                                            expression to be evaluated in array context. Used to assign
understand Perl.
                                                                                                                                                           # prints '7',the number of digits in $str
                                            the array/list size to a scalar.
These are not real Perl
                                                                                "@{[something]}" is
                                                                                                                print "these people @{[get names()]} get promoted"
                                            Interpolate an array in a string:
operators; they are
                              0{[]}
concatenation of other operators that achieve a
                                                                                                                                                             $ perl -le 'print ~~localtime'
Mon Nov 30 09:06:13 2009
                                            Force scalar context.
                                                                                    In scalar context <u>localtime</u> returns human readable time,
specific effect.
                                                                                    but in list context it returns a 9-tuple with date elements.
Truth and falsehood
                              · False in a boolean

    Negation of a true value by "!" or "not"

                                                                                                                So the following scalar values are
                                                                                                                                                             All other scalar values, including the following are
                                                                    returns a special false value.
                                                                                                                considered false:
                                 context:
                                                                                                                                                             true:
                                                                   When evaluated as a string it is treated as ", but as a number, it is treated as 0.
                                                                                                                                                             1 any non-0 number' ' the string with a space in it
                                  the number 0.
                                                                                                                · undef - the undefined value
 Remember that the
                                    the strings '0' and '',
                                                                                                                • 0 the number 0, even if you write it
strings '0' and " mean
                                                                                                                   as 000 or 0.0
" the empty string.
                                                                                                                                                             • '00' two or more 0 characters in a string
                                   the empty list (),
      The output of
                                                                                                                                                             • "0\n" a 0 followed by a newline
glob() may return a file
                                    "undef'
                                                                                                                • '0', a single 0 in the string.
                                                                                                                                                              'true'

    All other values are true.

named '0'!
                                                                                                                                                             • 'false' . Even the string 'false' evaluates to true.
 🛕 a bareword false has
                               One way to define valid true and false constant symbols that can be used in assignments (but see +):
                                                                                                                                                      use constant { true => 1, false => 0 };
a truth value of true!!!!
                                                                                                                                                       if (-e $fname && -f _ && -r _ ) {
  print("$fname exists, is readable\n"); }
File test operators
See filetest -X
                              File tests can be stacked (-r -w -e $fname) or combined as in the following example or:
                                 Notice the underscore in the example: it's the {\it virtual filehandle}\ \_ accessing the last {\it \underline{stat}} or {\it \underline{lstat}} result :
The operators check if
                                            is readable by effective uid/gid
                                                                                            exists.
                                                                                                                                                             is a block special file.
                                                                                                                                                             is a character special file.
                                            is writable by effective uid/gid
the file...
                               -w
                                                                                     -z
                                                                                            is empty.
                                            is executable by effective uid/gid is owned by effective uid
                                                                                                                                                             handle is opened to a tty. has setuid bit set.
See also:
                                                                                    -s
-f
                                                                                            has nonzero size (returns size in bytes).
                                                                                                                                                       -t
                              -o
-R
-W
                                                                                            is a plain file.

    File Tests or

                                                                                                                                                       -u
                                                                                            is a directory.
                                            is readable by real uid/gid is writable by real uid/gid
                                                                                    -d
                                                                                                                                                             has setgid bit set.
• <u>File test operators</u> @
                                                                                            is a symbolic link
                                                                                                                                                             has sticky bit set.
  perl tutorial
```

is a named pipe (FIFO) or Filehandle is a pipe.

Days between start time and file access time

is an ASCII text file (heuristic guess). is a "binary" file (opposite of -T).

Days between start time and node change time (in

-C

-X -O -M

See also:

localtime

· IO::Interactive

is executable by real uid/gid

Days between start time and file

file is owned by real uid.

modification time

Perl 5 Constants and Variables

Perl Constants • Perl pragma to declare constants. ... But be aware that these are still not read-only, that they inject sub-routines and have several limitations. Read the doc!! <u>CPAN modules for defining constants by Neil Bowers</u>. Of particular interest: <u>Const::Fast</u> and <u>Attribute::Constant</u> for efficient read-only constants. **Perl Variables Names Array Naming Conventions** All: underscore or letter of the first character. Case is significant in all names. ASCII by Module names are MixedCaseNoUnderscoresConstants are UPPERCASE_WITH_UNDERSCORES · Local variables: \$lowercase Similar conventions, except that array names should be plural. Global variables: \$Title Case @locals @Global_Arrays Package wide vars are Mixed_Case_With_Underscores Functions/methods are lowercase_with_underscores default, UTF-8 if the utf8 Constants: \$UPPER_CASE pragma is used. · All variables: words separated by underscores. @CONSTANT_ARRAYS Avoid ALLUPPERCASE: used by Perl special variables. Perl types Sigil **Examples** Meaning **Scalar** \$ \$foo Simple scalar value 29th element of array @days \$days[28] Value associated with the Feb key of hash %days \$days{'Feb'} Same as \$days, but unambiguous before alphanumerics. Useful inside strings for interpolation of variables followed by other letters. \${days} \$Dog::days The \$days variable inside the Dog package. Same as above. However this is an archaic use of the single quote. \$Dog'days \$#days \$days->[28] Last index of array @days. 29th element of array pointed to by reference \$days. \$days[0][2] \$d{99}{'Feb'} \$d{99, 'Feb'} Multi-dimensional array Multi-dimensional hash Multi-dimensional hash emulation list and Array Array containing (\$days[0], \$days[1], ... #days[\$#days]) . • A list is an ordered collection of scalars (of any type). @days Array slice containing (\$days[3], \$days[4], \$days[5]).

Array slice containing (\$days[3], \$days[4], \$days[5]). 0-based indexed (first @days[3,4,5] An array is a variable that contains a list. index is 0). @days[3..5] · Reading beyond the end of array returns undef Last index of array • Negative indices used in read access from the end: -1 is last item. @name is **\$**#nam Use these negative indices to access from the end. Do not compute index with \$#name -3, if the list size is 2, this will give invalid results. Use a slice to select multiple elements from a list, array, or hash. · An Ivalue slice imposes list context on the righthand side • slices Don't use a slice when you know you need exactly one element. What are the advantages of anonymous array? @ StackOverflow Anonymous arrays Anonymous array := a type of array reference. Array reference allows Perl to treat the array as a single item.
 This can be used to build nested data structures. Perlref @ Perldoc, Perl reference tutorial @ Perldoc This can be used to build, nested data structures. Associative array (hash): keys-value pairs. Can be initialized as: Initialize a hash slice with array context: Hash/associative array %days %days = (Jan => 31, Feb => \$leap? 29 : 28, ...)
%days = ("Jan", 31, 'Feb', \$leap? 29 : 28, ... @char_to_num{'A' .. 'Z'} = 1 .. 26; @days{'J',F'} Hash slice containing ($days{'J'}$, $days{'F'}$). & is needed to create reference to subroutine Subroutine &foo Typeglob *foo See: Advanced Perl Programming, 1st Edition Section 3.2 7 kinds of package scalar variables 4. subroutine name 6. file handles array variables
 hash variables variables or variable 5. format names 7. directory handles like elements in Perl: how to format output in Perl?, Perl-Formats See write and select Scalar values Numeric literals examples. Useful related builtin functions Note: leading 0 work only for literals, not for string-to-number conversions my \$x = 12345;· integer: using the system's native format. numeric integer oct - supports binary, octal, my \$x = 12345.67;bigint - transparent big integer support.
bignum - transparent big number support. # floating point
scientific notation hex mу \$x 6.02e23; <u>hex</u> power² exponent: Perl >= v5.22 underline for legibility floating-point: using the system's native format. \$x = 0x1f.0p3;POSIX::ceil my bigrat - transparent big rational number support. 4_294_967_296; my \$x POSIX::floor my \$x = 0x1234 5678;# underline in hex is also OK abs \$x = 0.377; \$x = 0.377; \$x = 0.0377;A variable holding an integer can be converted to тy # octal also floating-point if the operation done to it requires it mv Per1 >= v5.34 my \$x = 003//, my \$x = 0xffff; my \$x = 0b1100_0010; # becal also
hexadecimal
binary (such as dividing 1 by 2). string • double-quoted strings: perform backslash and variable interpolation of expression that begin with \$ (a scalar) or @ (an array). Hashes cannot be interpolated. single-quote strings: only perform \ ' and \ \ substitution (to ' and \ respectively), nothing else. Single quote and double quote strings can spread multiple lines: it embeds the newline character on each new line. But \n is only expanded in double quoted strings! In single quote string it is treated as two characters; no substitution is done (as explained above). Unicode support To use Unicode literally in a program, add the utf8 pragma: See: Perl Unicode Tutorial, Perl Unicode Introduction, Perl Unicode Support @ perldoc use utf8; Interpolates? Generic · Quote constructs Meaning Notes Literal string No • Not all characters can be used as the / separator. { }, () and < > can also be q// qq// Literal string Strings in Perl: Yes used. quoted, interpolated qx// Command execution Yes You can use whitespace between the quote specifier and its initial bracketing character: qw// World list No my \$chuck_of_code () $= q {$ and escaped m// Pattern match Yes if (Scondition) { s/// Pattern substitution s/// print "Salut! Character translation tr/// v/// No Regular expression • It's also possible to write: s<foo>(bar) and tr(a-f)[A-F] as well as separating them on 2 lines: tr (a-f) Array variables are interpolated by joining all elements with the separator specified by the <u>\$" special variable (\$LIST_SEPARATOR)</u>. ESC character Character escapes Alert (bell) Any Unicode code point, by name: (only inside Backspace \033 ESC in octal double quoted \e \f ESC character \o{33} ESC in octal \N{LATIN SMALL LETTER E WITH ACUTE} Form feed strings) DEL in hexadecimal \N{ U+E9 } Newline (usually LF) \x{263a} Character number 0x263A \n Carriage return (Usually CR) \cC Control-C \t Horizontal tab translation \u Force next character to titlecase \U Force all following characters to uppercase. Ends at \E ۱E Ends \U. \L. \F or \Q escapes (inside double quoted Force next character to lowercase \L \F Force all following characters to lowercase. Ends at \E Force all following characters to Unicode fold case. Ends at \E Backslash all following non alphanumeric characters. Ends at **\E** strings) \Q bareword In Perl, a bareword refers to a sequence of characters suitable for an identifier. It's not quoted. By default Perl allows barewords to behave like strings. This is not allowed when any of use strict; or use strict "subs"; or use v5.12; is specified. Here documentsHere docs @ Perl Perl here-documents are a form of line oriented quoting. There are several forms of here documents, where the identifier (like EOF used below, but can be any word) must be placed at the beginning of the terminating line: maven Perl here doc · Default: <<EOF: Supports variable interpolation. <<"EOF"; Supports variable interpolation. Can also be written with whitespace as in << "EOF Double quotes: Does not support interpolation. Can also be written with whitespace as in << 'EOF';
Execute commands in a shell and return text printed on stdout. Can also be written with whitespace as in << 'EOF'; @Wikipedia Single quotes: <<'EOF': backticks: indented: <<-EOF; Allows indenting the here-doc string. Can also use the ~ with the other forms: <<-\EOF, <<-"EOF", Perl Regexp PCRE cheatsheet Regexp Tutorial Debuggex regexp tester info. cheatsheets & **Learn PCRE in X minutes** regex101

RegEx Pal

regexp testers

| Perl Special Variables Perl Variables Deprecated and | To get information about \$< | erl special variable from the command line use: peridoc -v '\$<' | e use the peridoc -v command. | | | |
|--|--|---|---|--|--|--|
| Deprecated and | \$# \$* \$[\${^E | | | | | |
| removed variables: | | NCODING \${^WIN32 SLOPE | PY_STAT} | | | |
| General variables | | | | | | |
| | • \$ARG • \$_ | | subroutine parameters | • @ARG • @_ | | |
| | • \$LIST_SEPARATOR • \$" | | Subscript separator for multidimensional array emulation | • \$SUBSCRIPT_SEPARATOR • \$SUBSEP • \$; | | |
| | • \$PROGRAM_NAME • \$0 | | Name used to execute the current copy of Perl | • \$EXECUTABLE_ • \$^X | NAME | |
| | • \$PROCESS_ID • \$PID • \$\$ | Process real GID | • \$REAL_GROUP_ID • \$GID • \$(| Process effective GID | • \$EFFECTIVE_GROUP_I D • \$EGID • \$) | |
| | • \$REAL_USER_ID • \$UIG • \$< | | Process effective UID | • \$EFFECTIVE_US • \$EUID • \$> | ER_ID\$ | |
| • | • \$a The Perl sort function of the state of | ion uses global variables \$a and \$b. sort @sorted = sort { \$a <=> \$b } @ | | ion that uses the <=> equ | uality operator to force numerical | |
| Current environment | %ENV | | ccessed as an associative array (a hoss shell environment variables the | | rays. | |
| | \$OLD_PERL_VERSION\$] | | Perl interpreter revision, version and subversion | • \$PERL_VERSION • \$^V | 1 | |
| | • \$SYSTEM_FD_MAX • \$^F | | Fields of each line when autosplit mode is on. | @F | | |
| Include Directories | @INC | Included filenames | %INC | Hook localization (?) | \$INC | |
| | • \$INPLACE_EDIT • \$^I | Package's class parent classes | @ISA | Emergency memory pool | \$^M | |
| Maximum block nesting | \${^MAX_NESTED_EVAL_ | BEGIN_BLOCKS} | | Time when program began running | • \$BASETIME • \$^T | |
| | • \$OSNAME • \$^O | Signal handlers | %SIG | Coderefs for various perl keywords | %{^HOOK} | |
| Regexp Variables | | | | | | |
| captured sub-patterns | \$ <digit>(\$1, \$2,)</digit> | | Capture buffer content | @{^CAPTURE} | | |
| | • \$MATCH • \$& | | String matched (compiled regexp) | \${^MATCH} | | |
| | • \$PREMATCH • \$` | | String preceding match (compiled regexp) | \${^PREMATCH} | | |
| | • \$POSTMATCH • \$' | | String following match (compiled regexp) | {^POSTMATCH} | | |
| | • \$LAST_PAREN_MATCH • \$+ | I | Most recently closed capture group | • \$LAST_SUBMAT • \$^N | CH_RESULT | |
| <u>values</u> | %{^CAPTURE}%LAST_PAREN_MATC%+ | Н | Maximum regexp nested group | \${^RE_COMPILE_R | ECURSION_LIMIT} | |
| | @LAST_MATCH_START | | • @LAST_MATCH_END • @+ | Named captured groups | • %{^CAPTURE_ALL} • %- | |
| Last successful pattern | \${^LAST_SUCESSFUL_PA | TTERN} | Result of last successful regexp assertion | • \$LAST_REGEXP_ • \$^R | _CODE_RESULT | |
| regexp debug flag | \${^RE_DEBUG_FLAG} | | regexp internal optimization/mem | nory \${^RE_TRIE_N | MAXBUF} | |
| Format Variables | | | | | | |
| | • \$ACCUMULATOR • \$^A | | | | | |
| defaults to \f | IO::Handle->format_form \$FORMAT_FORMFEED \$^L | | Set of characters after which a string may be broken to fill continuation fields | | at_line_break_characters EXPR _BREAK_CHARACTERS | |
| the page on currently | HANDLE->format_lines_\$FORMAT_LINES_LEF\$- | | Current page length of current output channel | HANDLE->format_lines_per_page(EXPR)\$FORMAT_LINES_PER_PAGE\$= | | |
| page format of output | HANDLE->format_top_n\$FORMAT_TOP_NAME\$^ | | Report format name of output channel | HANDLE->format_name(EXPR) \$FORMAT_NAME \$~ | | |
| | The variables \$0, \$1, \$^E, and \$? contain information about different types of error conditions that may appear during execution of a Perl program. They correspond to errors detected by the Perl interpreter, C library, operating system, or an external program, respectively. | | | | | |
| | • \$EVAL_ERROR • \$@ | | Current state of interpreter | • \$EXCEPTIONS_E • \$^S | BEING_CAUGHT | |
| integer variable | • \$OS_ERROR • \$ERRNO • \$! | \$1 returns the system variable <u>errno</u> when used in a numeric context, but returns the string from <u>perror()</u> when used in string context. | Hash of error names to 0 or 1, set to 1 if current error is this error. | • %OS_ERROR • %ERRNO • %! | | |
| | • \$EXTENDED_OS_ERRO • \$^E | | | | | |
| Status returned by last | • \$CHILD_ERROR • \$? | | native status returned by last pipe close , backtick command, wait() or waitpid() or system() call | \${^CHILD_ERROR_ | NATIVE} | |

| Current value of warning switch | • \$^W | | | Current set of warning checks enabled by the use warnings pragma | \${^WARNING_BITS | 5} | |
|---|--|--|---------------------------|---|---|---|--|
| Variables related to the interpreter state | These variables provide inform | ation about the c | urrent interpreter state. | | | | |
| Flag associated with the -c switch | | | | The current value of the debugging flags | • \$DEBUGGING • \$^D | | |
| Current phase of the perl interpreter | \${^GLOBAL_PHASE} | | | Debugging support. Internal variable. | • \$PERLDB • \$^P | | |
| Compile-time hints for the perl interpreter. Internal use only | \$^H | | | Values of compiled statements | %^H | %^H | |
| Taint mode | \${^TAINT} | | | Safe locale operations availability | \${^SAFE_LOCALES | 5} | |
| Input/Output Layers. Internal use by PerlIO only. | \${^OPEN} | | | Unicode Settings of Perl | \${^UNICODE} | | |
| Internal UTF-8 offset caching code state | \${^UTF8CACHE} | | | State of UTF-8 locale detected by perl at startup. | \${^UTF8LOCALE} | | |
| File handle Variables | See also: Perl File Handles The following variables a | | | are used in the Input/Output handling as well as program arguments. | | | |
| Name of current file read from <> | \$ARGV | Command line arguments of the script ← See diamond operator <>. → | | @ARGV | Number of arguments minus one | \$#ARGV | |
| Special file handle that iterates over command-line filenames in @ARGV | ARGV Special file handle that points to currently open output file when doing edit-in-place processing | | ARGVOUT | | | | |
| Output field separator for the print operator | IO::Handle->output_field_separator(EXPR) \$OUTPUT_FIELD_SEPARATOR \$OFS \$, | | | Current line number for the last file handled accessed | HANDLE->input_\$INPUT_LINE_N\$NR\$. | | |
| Input record separator (newline by default) | IO::Handle->input_record_separator(EXPR) \$INPUT_RECORD_SEPARATOR \$RS \$/ | | | Output record separator | • IO::Handle->outpu • \$OUTPUT_RECO • \$ORS • \$\ | t_record_separator(EXPR) RD_SEPARATOR | |
| Auto-flush control order of output @ Perl Maven Suffering from Buffering? | HANDLE->autoflush(EXPR) SOUTPUT_AUTOFLUSH IN SI Perl activates file buffering by default. Assign 1 to \$1 to activate auto-flush. | | | Last read file handle | \${^LAST_FH} | | |

Perl 5 Input/Output

| References | Writing to | oerldoc browser o files with Perl @ Pe in-memory @ stack | erl Maven • | upid open() tricks @ No explicit filenam create an anonymo | | print to a stringread lines from | |
|--|--|---|---------------|--|---|---|---|
| print, printf, sprintf | | print, printf, sprintf (which describes the format). Note: print is more efficient than printf. print and printf output to stdout by default, but accept a file handle as the first argument if it is NOT followed by a separating comma! (a ',' puts it in the list to print!) | | | | | |
| diamond operator <> | | th <> and <<>> operators read the content of files listed on the command line via @ARGV. Nothing or - on the command line identifies stdin. e <> operator supports shell redirection and pipe operations which <<>> does not allow (for security reasons). | | | | | |
| The double diamond, a more secure <> (Perl >= v5.22) | - | <pre>print <>;</pre> | | | <pre>print <<>>; print sort <<>>;</pre> | ← safer one ← safer one | Redirection cannot be forced via file names embedding them |
| In-place-editing of The <> operator tries to duplicate the original file's permission and ownership. | Set \$^I to change the In a while renames opens a prints int | to a backup file extension (such as Emacs "~" or ".bak") to he behaviour of the <> and <<>> operators and print. le (<>) {} loop, when \$^T is not undef (its default), Perl: es currently processed file with the specified extension added, a new file with the original name into the new file. odification goes into the new file: in-place-editing it! print sort <<>>; | | | | | |
| perl -i cmdline option | It's also pos | s also possible to do this on the command line! For example: perl _p _i ~ _w _e 's/something/Something else/g' data*.dat | | | | | |
| Special filehandle names | ARGV | The special filehandle that iterates over command-line filenames in @ARGV. Usually written as the null filehandle in the angle operator <> (or <<>>) | | | | | |
| Also See: • File handle Variables | ARGVOUT | The special filehandle that points to the currently open output file when doing edit-in-place processing with <u>-i</u> . • Useful when you have to do a lot of inserting and don't want to keep modifying \$ | | | | | |
| section above. | STDIN | <pre> <stdin> : line input operator for the STDIN filehandle (for the <u>standard input</u>). • Each time <stdin> is used in scalar context, Perl reads 1 complete line of the standard input and uses it as the value of <stdin>. • The string includes a line termination character. Use the <u>chomp()</u> built-in function to strip it off the variable. • If <stdin> is read in list context, it returns all lines inside a list! For example, foreach (<stdin>) { } reads the entire stdin in 1 step: \$_ holds it all! while (<stdin>) { # print all while (defined(\$_ = <stdin>)) {</stdin></stdin></stdin></stdin></stdin></stdin></stdin></pre> | | | | | |
| | | } | # stdin | } | | stored in the defaul | t variable \$_ and the loop stops on :STDIN> returns undef. |
| | STDOUT | standard output | | | | | |
| | STDERR | standard error Note: generally STDERR is not buffered, while STDOUT is buffered by default. Text sent on STDERR may show up before STDOUT. • Print a new line on STDOUT to help flushing it or assign 1 to \$ to activate auto-flush. | | | | | |
| | DATA | | | | | | |
| <u>say</u> | • <u>say</u> | use fea | ture qw(say); | or use v5. | 10; (or higher). Like p | orint, but implicitly appends | a newline at the end of the list. |
| <u>open</u> | | | | | | | |

Perl 5 Statements

| Loop control | See <u>perlsyn</u> for more information on Perl syntax which includes declarations, blocks, loops, labels, subroutines, etc | | | | |
|--|---|-----------------------------------|--|--|--|
| Use the <u>last</u> and <u>redo</u> inside a naked block of code to control looping. | loop control keywords: • last of exits the loop. • next of exterts the next iteration of the loop. • redo of restarts the loop block without evaluating the condition again. | | The last, next, and redo loop control keywords work in the following constructs: • while (condition) { } • until (condition) { } • for (init; condition; continue) { } • foreach array { } • naked block: { } | Notes: • The while and foreach loops may have a <u>continue</u> <u>block</u> : executed before evaluating condition again, which corresponds to the 3rd part of a for loop statement. See <u>this</u> @ stackOverflow. • Blocks can be labelled <u>o</u> as targets to <u>last</u> , <u>next</u> , and <u>redo</u> | |
| Statement modifiers | • if EXPR • unless EXPR • while EXPR • until EXPR • for LIST • foreach LIST • when EXPR • do block | processed. The has "END" of forea | ach statements impose a list context; the complete list is refore a loop like the following trying to stop on a line that on it will not work since it reads all of STDIN: ch (<stdin>) { if ?END/;</stdin> | The while statement imposes a scalar context; it takes one line at a time from <stdin> and the following code works properly: while (<stdin>) { last if /_END/; ; }</stdin></stdin> | |
| Conditional statements | | | | | |

Perl 5 Subroutines

| ı | | | | | | | | |
|---|--|--|--------------------|-----------------------------------|---|--------------------------------|---------------------------------|--|
| Perl subroutines | | | | | | | | |
| subroutine & | | Why we teach the subroutine ampersand Another point of view: Subroutines and Ampersands Why should I use the & to call a Perl subroutine? @ StackOverflow | | | | | | |
| Subroutine Prototypes | An older Po | erl feature. Clashes v | with subroutine si | ignatures as of Perl v5.20 | 0. In <i>Perl</i> >= <i>v</i> 5.20 put the :protot | type attribute before sul | broutine prototype parenthesis. | |
| Subroutine signatures | Exactly zero arguments | | () | Zero or 1 argument, no default, u | ınnamed: | (\$=) | | |
| Perl >=5.36: StablePerl >= 5.20: | Zero or 1 a | argument, no default, | , named | (\$val=) | Zero or 1 argument, named, with | ı default | (\$val=1) | |
| Experimental See: <u>Use v5.20</u> | exactly 1 n | named argument: | | (\$val) | Exactly 2 arguments | | (\$v1, \$v2) | |
| subroutine signatures | 2, 3 or 4 arguments no defaults: (\$v1, | | , \$v2, \$=, \$=) | 2,3 or 4 arguments, 1 default: | | (\$v1, \$v2, \$v3='a', \$=) | | |
| | Two or mor | re, any number of arg | guments. | (\$v1, \$v2, @) | Two or more arguments, remaind | ders into a named array: | (\$v1, \$v2, @rest) | |
| | Two or more arguments: an even number | | | (\$v1, \$v2, %) | Two or more arguments, remainders into a named hash: | | (\$v1, \$v2, %rest) | |
| | Class meti | hod | | (\$class,) | Object method | | (\$self,) | |
| Variables in subroutines | global by o | default | | | | | | |
| | <u>my</u> | local, lexical scope | e, non persistent | | | | | |
| | <u>state</u> | Local, lexical scope | e, persistent | Perl >= v5.10 | Restriction: in Perl < v5.28: array | and hashes state cannot | be initialized in list context. | |
| | our | creates a lexical sc | coped alias to a p | package variable | | | | |
| | Localizes an existing package variable to the current scope. It's not a declaration. The variable previous value is restored when leaving the scope. | | | | | | | |
| Returned value | The retur The subr | The result of the last evaluated expression is implicitly returned The return operator can be used but it's not required unless used to change execution flow (return immediately from the subroutine). The subroutine can return a scalar in scalar context or a list if called in list context. Inside the subroutine, use the wantarray function to determine the context of the subroutine call. | | | | | | |

Perl 5 Built-in Functions

| Perl Functions Perl syntax | To get information about a Perl function from the command line use the perldoc -f command. To get information about print use: perldoc -f print |
|---|--|
| ! Cautionary notes | |
| each keyword is broken Use <u>Var::Pairs</u> instead. | Do NOT use the built-in each. It is broken, as described by <u>Damian Conway</u> in his <u>Modern Perl Best Practice O'Reilly course</u> , section control structure. • each is not re-entrant: • nested loops of each over the same hash does not work as expected and will create infinite loop since the nested loop each juts iterates from where the first loop each left it. • Exiting the loop leaves the state of the each internal pointer at the current location. • If you use each on the same hash later it will resume from where it left, it will not start form the beginning. |
| | |

Perl 5 Modules

| Peri Wodules | | |
|--|-----------|--|
| Perl core modules | | letect where a module is installed : perldoc -1 Module check if a module is part of Perl core : corelist Module (Perl >= v5.9.2) |
| Modules @perltutorial Modules Using simple modules © | <u>do</u> | Looks for the module file by searching the @INC path. Performed at run time (and therefore can be done conditionally). • If Perl finds the file, it places the code inside the calling program and executes it. Otherwise, Perl will skip the do statement silently. • The "included" code does not have access to the lexical variables from the main program. • Skip the @INC path lookup if given a file path starting with ./,/, or / |
| | require | Loads the module file once, also teaching the @INC path. Performed at run time (and therefore can be done conditionally). • If the require for the same file appears twice, Perl ignores it. Perl will issue an error message if it cannot find the file (as opposed to do) • Skip the @INC path lookup if given a file path starting with ./,/, or / |
| The normal way to access Perl modules ➡ | use | Similar to require except that Perl applies it before the program starts: it's done at compile time. • Therefore the use, statement cannot be invoked inside conditional statements such as if-else. Used often to include a module in a program. |

Topic - Directory Operations

| | | | • • | |
|--|--|---|---|---|
| Directory Operations | In Books: LPo | | | |
| Opening Files | All file open operations are relative to the <u>current workin</u> relative file names) | ng directory (for | open my \$filehandle, ' | <:utf8', 'a_relative/path.txt' |
| Built-in Functions | Related Functions/Packages / Descriptions | | | Notes |
| Getting file names by: Globbing: with glob | File::Glob (Perl >= v5.6.0) - provides more control. | Example: | <pre>my @all_files = glob my @perl_files = glob</pre> | '*'; '*.pm *.pl'; # 2 globs, space-separated |
| with the glob operator <> | The <> operator is identifying: • a filehandle, when: the item inside <> is a Perl identifier or an indirect file handle read scalar, • a glob expression otherwise. | Glob examples: | my @perl_files = <'*.pr | ; # 1 glob: no space, no need for string m *.pl'>; # 2 globs, space-separated |
| | | | <pre>my \$etc_dir = '/etc'; my @etc_dir_files = <\$</pre> | etc_dir/* \$etc_dir/.*>; |
| | | | my @files = <larry *="">;</larry> | # a glob |
| | See: readline | Filehandle | my @his_lines = <larry< th=""><th>>; # a filehandle read</th></larry<> | >; # a filehandle read |
| | | examples: | | >; # indirect filehandle read of LARRY handle ine LARRY; # another way to write above ine \$name; |
| with a directory handle LPo | opendir: open a directory: get a directory handle readdir: read the directory handle. But see this. closedir: close the directory handle. DirHandle (<= Perl 5.5) File::Spec::Functions (>= Perl v5.5.4) Path::Class | Example: iterate explicitly over a list of file names extracted from the directory using these 3 functions. | <pre>opendir my \$dh, \$dir or foreach \$file (readdir</pre> | r die "Failed opening \$dir: \$!"; \$dh) { inside \$dir\n"; # 1 no path in name! |
| Removing files | • unlink a list or \$_ | | unlink 'file1.txt', 'funlink qw(file1.txt funlink glob 'file?.txt | ile2.txt); |
| Renaming files | Tename an old file name to a new one. The fat comma operator is sometimes used to highlight what is the old and the new name. | As in here: | rename 'old_name' , 'rename 'old_name' => 'r | new_name'; new_name'; # using fat comma |
| Changing permissions | chmod changes file permissions | | | |
| Changing ownership | chown changes file ownership | | | |
| Creating Hard link | • <u>link</u> to create a hard link | | | |
| Creating symbolic link | symlink to create a symbolic link | | | |
| chdir Change current working directory | File::chdir File::HomeDir | • chdir without \$ENV{LOGDIR | environment values if they ar | the result of the serve of the |
| Modules | Functions Legend: Exported by default, exported on request, W. | in32 specific | | Extra Information |
| Cwd | getcwd, cwd, fastcwd, fastgetcwd, getdcwd abs_path, realpath, fast_abs_path | | | <pre>use Cwd; my \$curdir = getcwd; print "cwd is \$curdir\n";</pre> |
| | | | | |

PerlTidy formatting control

| perItidy option | Option | Impact |
|-------------------|---|---|
| indentation style | -bl, opening-brace-on-new-line brace-left | Without this option (the default) the code indentation style selected is <u>K&R style</u>. With this option, the indentation style is <u>Allman/BSD style</u>. |
| | | |