

NAME

perl5113delta - what is new for perl v5.11.3

DESCRIPTION

This document describes differences between the 5.11.2 release and the 5.11.3 release.

If you are upgrading from an earlier release such as 5.11.1, first read the *perl5112delta*, which describes differences between 5.11.1 and 5.11.2

Incompatible Changes

Filehandles are blessed directly into `IO::Handle`, as `FileHandle` is merely a wrapper around `IO::Handle`.

The previous behaviour was to bless Filehandles into *FileHandle* (an empty proxy class) if it was loaded into memory and otherwise to bless them into `IO::Handle`.

Core Enhancements

Unicode version

Perl is shipped with the latest Unicode version, 5.2, dated October 2009. See <http://www.unicode.org/versions/Unicode5.2.0> for details about this release of Unicode. See *perlunicode* for instructions on installing and using older versions of Unicode.

Unicode properties

Perl can now handle every Unicode character property. A new pod, *perluniprops*, lists all available non-Unihan character properties. By default the Unihan properties and certain others (deprecated and Unicode internal-only ones) are not exposed. See below for more details on these; there is also a section in the pod listing them, and why they are not exposed.

Perl now fully supports the Unicode compound-style of using `=` and `:` in writing regular expressions: `\p{property=value}` and `\p{property:value}` (both of which mean the same thing).

Perl now fully supports the Unicode loose matching rules for text between the braces in `\p{...}` constructs. In addition, Perl also allows underscores between digits of numbers.

All the Unicode-defined synonyms for properties and property values are now accepted.

`\X`, which matches a Unicode logical character, has been expanded to work better with various Asian languages. It now is defined as an extended grapheme cluster. (See <http://www.unicode.org/reports/tr29/>). Anything matched previously that made sense will continue to be matched. But in addition:

- `\X` will now not break apart a CR LF sequence.
- `\X` will now match a sequence including the ZWJ and ZWNJ characters.
- `\X` will now always match at least one character, including an initial mark. Marks generally come after a base character, but it is possible in Unicode to have them in isolation, and `\X` will now handle that case, for example at the beginning of a line or after a ZWSP. And this is the part where `\X` doesn't match the things that it used to that don't make sense. Formerly, for example, you could have the nonsensical case of an accented LF.
- `\X` will now match a (Korean) Hangul syllable sequence, and the Thai and Lao exception cases.

Otherwise, this change should be transparent for the non-affected languages.

`\p{...}` matches using the Canonical_Combining_Class property were completely broken in previous Perls. This is now fixed.

In previous Perls, the Unicode `Decomposition_Type=Compat` property and a Perl extension had

the same name, which led to neither matching all the correct values (with more than 100 mistakes in one, and several thousand in the other). The Perl extension has now been renamed to be `Decomposition_Type=Noncanonical` (short: `dt=noncanon`). It has the same meaning as was previously intended, namely the union of all the non-canonical Decomposition types, with Unicode `Compat` being just one of those.

`\p{Uppercase}` and `\p{Lowercase}` have been brought into line with the Unicode definitions. This means they each match a few more characters than previously.

`\p{Cntrl}` now matches the same characters as `\p{Control}`. This means it no longer will match Private Use (`gc=co`), Surrogates (`gc=cs`), nor Format (`gc=cf`) code points. The Format code points represent the biggest possible problem. All but 36 of them are either officially deprecated or strongly discouraged from being used. Of those 36, likely the most widely used are the soft hyphen (U+00AD), and BOM, ZWSP, ZWNJ, WJ, and similar, plus Bi-directional controls.

`\p{Alpha}` now matches the same characters as `\p{Alphabetic}`. The Perl definition included a number of things that aren't really alpha (all marks), while omitting many that were. As a direct consequence, the definitions of `\p{Alnum}` and `\p{Word}` which depend on Alpha also change.

`\p{Word}` also now doesn't match certain characters it wasn't supposed to, such as fractions.

`\p{Print}` no longer matches the line control characters: Tab, LF, CR, FF, VT, and NEL. This brings it in line with the documentation.

`\p{Decomposition_Type=Canonical}` now includes the Hangul syllables.

The Numeric type property has been extended to include the Unihan characters.

There is a new Perl extension, the 'Present_In', or simply 'In', property. This is an extension of the Unicode Age property, but `\p{In=5.0}` matches any code point whose usage has been determined as of Unicode version 5.0. The `\p{Age=5.0}` only matches code points added in *precisely* version 5.0.

A number of properties did not have the correct values for unassigned code points. This is now fixed. The affected properties are `Bidi_Class`, `East_Asian_Width`, `Joining_Type`, `Decomposition_Type`, `Hangul_Syllable_Type`, `Numeric_Type`, and `Line_Break`.

The `Default_Ignorable_Code_Point`, `ID_Continue`, and `ID_Start` properties have been updated to their current Unicode definitions.

Certain properties that are supposed to be Unicode internal-only were erroneously exposed by previous Perls. Use of these in regular expressions will now generate, if enabled, a deprecated warning message. The properties are: `Other_Alphabetic`, `Other_Default_Ignorable_Code_Point`, `Other_Grapheme_Extend`, `Other_ID_Continue`, `Other_ID_Start`, `Other_Lowercase`, `Other_Math`, and `Other_Uppercase`.

An installation can now fairly easily change which Unicode properties Perl understands. As mentioned above, certain properties are by default turned off. These include all the Unihan properties (which should be accessible via the CPAN module `Unicode::Unihan`) and any deprecated or Unicode internal-only property that Perl has never exposed.

The generated files in the `lib/unicore/To` directory are now more clearly marked as being stable, directly usable by applications. New hash entries in them give the format of the normal entries, which allows for easier machine parsing. Perl can generate files in this directory for any property, though most are suppressed. An installation can choose to change which get written. Instructions are in `perluniprops`.

Regular Expressions

U+0FFFF is now a legal character in regular expressions.

Modules and Pragmata

Pragmata Changes

`constant`

Upgraded from version 1.19 to 1.20.

`diagnostics`

This pragma no longer suppresses `Use of uninitialized value in range (or flip) warnings`. [perl #71204]

`feature`

Upgraded from 1.13 to 1.14. Added the `unicode_strings` feature:

```
use feature "unicode_strings";
```

This pragma turns on Unicode semantics for the case-changing operations (`uc/lc/ucfirst/lcfirst`) on strings that don't have the internal UTF-8 flag set, but that contain single-byte characters between 128 and 255.

`legacy`

The experimental `legacy` pragma, introduced in 5.11.2, has been removed, and its functionality replaced by the new feature pragma, `use feature "unicode_strings"`.

`threads`

Upgraded from version 1.74 to 1.75.

`warnings`

Upgraded from 1.07 to 1.08. Added new `warnings::fatal_enabled()` function.

Updated Modules

`Archive::Extract`

Upgraded from version 0.34 to 0.36.

`CPAN`

Upgraded from version 1.94_51 to 1.94_5301, which is 1.94_53 on CPAN plus some local fixes for `bleadperl`.

Includes better `bzip2` support, improved `FirstTime` experience with auto-selection of CPAN mirrors, proper handling of modules removed from the Perl core, and an updated `'cpan'` utility script

`CPANPLUS`

Upgraded from version 0.89_09 to 0.90.

`Encode`

Upgraded from version 2.38 to 2.39.

`ExtUtils::MakeMaker`

Upgraded from version 6.55_02 to 6.56. Adds new `BUILD_REQUIRES` key to indicate build-only prerequisites. Also adds support for `mingw64` and the new `"package NAME VERSION"` syntax.

`File::Path`

Upgraded from version 2.08 to 2.08_01.

`Module::Build`

Upgraded from version 0.35_09 to 0.36. Compared to 0.35, this version has a new

'installdeps' action, supports the PERL_MB_OPT environment variable, adds a 'share_dir' property for *File::ShareDir* support, support the "package NAME VERSION" syntax and has many other enhancements and bug fixes. The 'passthrough' style of *Module::Build::Compat* has been deprecated.

Module::CoreList

Upgraded from version 2.23 to 2.24.

POSIX

Upgraded from version 1.18 to 1.19. Error codes for *getaddrinfo()* and *getnameinfo()* are now available.

Pod::Simple

Upgraded from version 3.10 to 3.13.

Safe

Upgraded from version 2.19 to 2.20.

Utility Changes

perlbug

No longer reports "Message sent" when it hasn't actually sent the message

Changes to Existing Documentation

The Pod specification (*perlpodspec*) has been updated to bring the specification in line with modern usage already supported by most Pod systems. A parameter string may now follow the format name in a "begin/end" region. Links to URIs with a text description are now allowed. The usage of `L<"section">` has been marked as deprecated.

if.pm has been documented in "use" in *perlfunc* as a means to get conditional loading of modules despite the implicit BEGIN block around *use*.

Installation and Configuration Improvements

Testing improvements

It's now possible to override `PERL5OPT` and friends in *t/TEST*

Platform Specific Changes

Win32

- Always add a manifest resource to *perl.exe* to specify the `trustInfo` settings for Windows Vista and later. Without this setting Windows will treat *perl.exe* as a legacy application and apply various heuristics like redirecting access to protected file system areas (like the "Program Files" folder) to the users "VirtualStore" instead of generating a proper "permission denied" error.

For VC8 and VC9 this manifest setting is automatically generated by the compiler/linker (together with the binding information for their respective runtime libraries); for all other compilers we need to embed the manifest resource explicitly in the external resource file.

This change also requests the Microsoft Common-Controls version 6.0 (themed controls introduced in Windows XP) via the dependency list in the assembly manifest. For VC8 and VC9 this is specified using the `/manifestdependency` linker commandline option instead.

cygwin

Enable IPv6 support on cygwin 1.7 and newer

OpenVMS

Make `-UDEBUGGING` the default on VMS for 5.12.0.

Like it has been everywhere else for ages and ages. Also make command-line selection of `-UDEBUGGING` and `-DDEBUGGING` work in `configure.com`; before the only way to turn it off was by saying no in answer to the interactive question.

Selected Bug Fixes

- Ensure that `pp_qr` returns a new regexp SV each time. Resolves RT #69852.
Instead of returning a(nother) reference to the (pre-compiled) regexp in the optree, use `reg_temp_copy()` to create a copy of it, and return a reference to that. This resolves issues about `Regexp::DESTROY` not being called in a timely fashion (the original bug tracked by RT #69852), as well as bugs related to blessing regexps, and of assigning to regexps, as described in correspondence added to the ticket.
It transpires that we also need to undo the `SvPVX()` sharing when ithreads cloning a `Regexp` SV, because `mother_re` is set to `NULL`, instead of a cloned copy of the `mother_re`. This change might fix bugs with regexps and threads in certain other situations, but as yet neither tests nor bug reports have indicated any problems, so it might not actually be an edge case that it's possible to reach.
- Several compilation errors and segfaults when perl was built with `-Dmad` were fixed.
- Fixes for lexer API changes in 5.11.2 which broke `NYTProf`'s `savesrc` option.
- `-t` should only return `TRUE` for file handles connected to a TTY
The Microsoft C version of `isatty()` returns `TRUE` for all character mode devices, including the `/dev/null` style "nul" device and printers like "lpt1".
- Fixed a regression caused by commit fafafba which caused a panic during parameter passing [perl #70171]
- On systems which in-place edits without backup files, `-i*` now works as the documentation says it does [perl #70802]
- Saving and restoring magic flags no longer loses `readonly` flag.
- The malformed syntax `grep EXPR LIST` (note the missing comma) no longer causes abrupt and total failure.
- Regular expressions compiled with `qr{ }` literals properly set `$'` when matching again.
- Using named subroutines with `sort` should no longer lead to bus errors [perl #71076]
- Numerous bugfixes catch small issues caused by the recently-added Lexer API.
- Smart match against `@_` sometimes gave false negatives. [perl #71078]
- `$@` may now be assigned a read-only value (without error or busting the stack).
- `sort` called recursively from within an active comparison subroutine no longer causes a bus error if run multiple times. [perl #71076]

New or Changed Diagnostics

- `split` now warns when called in void context
- `printf`-style functions called with too few arguments will now issue the warning "Missing argument in %s" [perl #71000]

New Tests

Many modules updated from CPAN incorporate new tests.

t/comp/final_line_num.t

See if line numbers are correct at EOF

t/comp/form_scope.t

See if format scoping works

t/comp/line_debug.t

See if @{"_<\$file"} works

t/op/filetest_t.t

See if -t file test works

t/op/qr.t

See if qr works

t/op/utf8cache.t

Tests malfunctions of utf8 cache

t/re/uniprops.t

Test unicode \p{} regex constructs

Deprecations

The following items are now deprecated.

Use of "goto" to jump into a construct is deprecated

Using `goto` to jump from an outer scope into an inner scope is now deprecated. This rare use case was causing problems in the implementation of scopes.

Acknowledgements

Perl 5.11.3 represents approximately one month of development since Perl 5.11.2 and contains 61407 lines of changes across 396 files from 40 authors and committers:

Abigail, Alex Davies, Alexandr Ciornii, Andrew Rodland, Andy Dougherty, Bram, brian d foy, Chip Salzenberg, Chris Williams, Craig A. Berry, Daniel Frederick Crisman, David Golden, Dennis Kaarsemaker, Eric Brine, Father Chrysostomos, Gene Sullivan, Gerard Goossen, H. Merijn Brand, Hugo van der Sanden, Jan Dubois, Jerry D. Hedden, Jesse Vincent, Jim Cromie, Karl Williamson, Leon Brocard, Max Maischein, Michael Breen, Moritz Lenz, Nicholas Clark, Rafael Garcia-Suarez, Reini Urban, Ricardo Signes, Stepan Kasal, Steve Hay, Steve Peters, Tim Bunce, Tony Cook, Vincent Pit and Zefram.

Many of the changes included in this version originated in the CPAN modules included in Perl's core. We're grateful to the entire CPAN community for helping Perl to flourish.

Reporting Bugs

If you find what you think is a bug, you might check the articles recently posted to the `comp.lang.perl.misc` newsgroup and the perl bug database at <http://rt.perl.org/perlbug/>. There may also be information at <http://www.perl.org/>, the Perl Home Page.

If you believe you have an unreported bug, please run the **perlbug** program included with your release. Be sure to trim your bug down to a tiny but sufficient test case. Your bug report, along with the output of `perl -V`, will be sent off to perlbug@perl.org to be analysed by the Perl porting team.

If the bug you are reporting has security implications, which make it inappropriate to send to a publicly archived mailing list, then please send it to perl5-security-report@perl.org. This points to a closed subscription unarchived mailing list, which includes all the core committers, who be able to help assess the impact of issues, figure out a resolution, and help co-ordinate the release of patches to mitigate or fix the problem across all platforms on which Perl is supported. Please only use this

address for security issues in the Perl core, not for modules independently distributed on CPAN.

SEE ALSO

The *Changes* file for an explanation of how to view exhaustive details on what changed.

The *INSTALL* file for how to build Perl.

The *README* file for general stuff.

The *Artistic* and *Copying* files for copyright information.