

Changes in R

From 3.0.2 to 3.0.1

by the R Core Team

CHANGES IN R 3.0.2 patched

NEW FEATURES

- On Windows there is support for making '.texi' manuals using texinfo 5.0 or later: the setting is in file 'src/gnuwin32/MkRules.dist'.
A packaging of the Perl script and modules for texinfo 5.2 has been made available at <http://www.stats.ox.ac.uk/pub/Rtools/>.
- `write.table()` now handles matrices of 2^{31} or more elements, for those with large amounts of patience and disc space.
- There is a new function, `La_version()`, to report the version of LAPACK in use.
- The HTML version of 'An Introduction to R' now has links to PNG versions of the figures.
- There is some support to produce manuals in ebook formats. (See 'doc/manual/Makefile'. Suggested by Mauro Cavalcanti.)

PACKAGE INSTALLATION

- The new field 'SysDataCompression' in the 'DESCRIPTION' file allows user control over the compression used for 'sysdata.rda' objects in the lazy-load database.

C-LEVEL FACILITIES

- The long undocumented remapping of `rround()` to `Rf_fround()` in header 'Rmath.h' is now formally deprecated: use `fround()` directly.
- Remapping of `prec()` and `trunc()` in the 'Rmath.h' header has been disabled in C++ code (it has caused breakage with `libc++` headers).

BUG FIXES

- `getParseData()` truncated the imaginary part of complex number constants. (Reported by Yihui Xie.)
- `dbeta(x, a, b)` with `a` or `b` within a factor of 2 of the largest representable number could infinite-loop. (Reported by Ioannis Kosmidis.)
- `provideDimnames()` failed for arrays with a 0 dimension. (PR#15465)
- `rbind()` and `cbind()` did not handle list objects correctly. (PR#15468)
- `replayPlot()` now checks if it is replaying a plot from the same session.
- `rasterImage()` and `grid.raster()` now give error on an empty (zero-length) raster. (Reported by Ben North.)
- `plot.lm()` would sometimes scramble the labels in plot type 5. (PR#15458 and PR#14837)

- `min()` did not handle `NA_character_` values properly. (Reported by Magnus Thor Torfason.)
- (Windows only) `readRegistry()` would duplicate default values for keys. (PR#15455)
- `str(..., strict.width = "cut")` did not handle it properly when more than one line needed to be cut. (Reported by Gerrit Eichner.)
- Removing subclass back-references when S4 classes were removed or their namespace unloaded had several bugs (e.g., PR#15481).
- `aggregate()` could fail when there were too many levels present in the `by` argument. (PR#15004)
- `namespaceImportFrom()` needed to detect primitive functions when checking for duplicated imports (reported by Karl Forner).
- `getGraphicsEvent()` did not exit when a user closed the graphics window. (PR#15208)
- Errors in vignettes were not always captured and displayed properly. (PR#15495)
- `contour()` could fail when dealing with extremely small `z` values. (PR#15454)
- Several functions did not handle zero-length vectors properly, including `browseEnv()`, `format()`, `gl()`, `relist()` and `summary.data.frame()`. (E.g., PR#15499)
- `Sweave()` did not restore the R output to the console if it was interrupted by a user in the middle of evaluating a code chunk. (Reported by Michael Sumner.)
- Fake installs of packages with vignettes work again.
- Illegal characters in the input caused `parse()` (and thus `source()`) to segfault. (PR#15518)
- The nonsensical use of `nmax = 1` in `duplicated()` or `unique()` is now silently ignored.
- `qcauchy(p,*)` is now fully accurate even when `p` is very close to 1. (PR#15521)
- The `validmu()` and `valideta()` functions in the standard `glm()` families now also report non-finite values, rather than failing.
- Saved vignette results (in a `‘.Rout.save’` file) were not being compared to the new ones during R CMD check.
- Double-clicking outside of the list box (e.g. on the scrollbar) of a Tk listbox widget generated by `tk_select.list()` no longer causes the window to close. (PR#15407)
- Improved handling of edge cases in `parallel::splitindices()`. (PR#15552)
- HTML display of results from `help.search()` and `??` sometimes contained badly constructed links.
- `c()` and related functions such as `unlist()` converted raw vectors to invalid logical vectors. (PR#15535)
- (Windows only) When a call to `system2()` specified one of `stdin`, `stdout` or `stderr` to be a file, but the command was not found (e.g. it contained its arguments, or the program was not on the PATH), it left the file open and unusable until R terminated. (Reported by Mathew McLean.)
- The `bmp()` device was not recording `res = NA` correctly: it is now recorded as 72 ppi.
- Several potential problems with compiler-specific behaviour have been identified using the ‘Undefined Behaviour Sanitizer’ in conjunction with the `clang` compiler.

- `hcl()` now honours NA inputs (previously they were mapped to black).
- Some translations in base packages were being looked up in the main catalog rather than that for the package.
- As a result of the 3.0.2 change about ‘the last second before the epoch’, most conversions which should have given NA returned that time. (The platforms affected include Linux and OS X, but not Windows nor Solaris.)
- `rowsum` has more support for matrices and dataframes with 2^{31} or more elements. (PR#15587)
- `predict(<lm object>, interval = "confidence", scale = <something>)` now works. (PR#15564)
- The bug fix in 3.0.2 for PR#15411 was too aggressive, and sometimes removed spaces that should not have been removed. (PR#15583)
- Running R code in a `tcltk` callback failed to set the busy flag, which will be needed to tell OS X not to ‘App Nap’.

CHANGES IN R 3.0.2

NEW FEATURES

- The ‘NEWS’ files have been re-organized.
This file contains news for R \geq 3.0.0: news for the 0.x.y, 1.x.y and 2.x.y releases is in files ‘NEWS.0’, ‘NEWS.1’ and ‘NEWS.2’. The latter files are now installed when R is installed. An HTML version of news from 2.10.0 to 2.15.3 is available as ‘doc/html/NEWS.2.html’.
- `sum()` for integer arguments now uses an integer accumulator of at least 64 bits and so will be more accurate in the very rare case that a cumulative sum exceeds 2^{53} (necessarily summing more than 4 million elements).
- The `example()` and `tools::Rd2ex()` functions now have parameters to allow them to ignore `\dontrun` markup in examples. (Suggested by Peter Solymos.)
- `str(x)` is considerably faster for very large lists, or factors with 100,000 levels, the latter as in PR#15337.
- `col2rgb()` now converts factors to character strings not integer codes (suggested by Bryan Hanson).
- `tail(warnings())` now works, via the new ‘[’ method.
- There is now support for the LaTeX style file ‘zi4.sty’ which has in some distributions replaced ‘inconsolata.sty’.
- `unlist(x)` now typically returns all non-list xs unchanged, not just the “vector” ones. Consequently, `format(lst)` now also works when the list `lst` has non-vector elements.
- The `tools::getVignetteInfo()` function has been added to give information about installed vignettes.
- New `assertCondition()`, etc. utilities in **tools**, useful for testing.
- Profiling now records non-inlined calls from byte-compiled code to BUILTIN functions.
- Various functions in **stats** and elsewhere that use non-standard evaluation are now more careful to follow the namespace scoping rules. E.g. `stats::lm()` can now find `stats::model.frame()` even if **stats** is not on the search path or if some package defines a function of that name.

- If an invalid/corrupt `.Random.seed` object is encountered in the workspace it is ignored with a warning rather than giving an error. (This allows R itself to rely on a working RNG, e.g. to choose a random port.)
- `seq()` and `seq.int()` give more explicit error messages if called with invalid (e.g. NaN) inputs.
- When `parse()` finds a syntax error, it now makes partial parse information available up to the location of the error. (Request of Reijo Sund.)
- Methods invoked by `NextMethod()` had a different dynamic parent to the generic. This was causing trouble where S3 methods invoked via lazy evaluation could lose track of their generic. (PR#15267)
- Code for the negative binomial distribution now treats the case `size == 0` as a one-point distribution at zero.
- `abbreviate()` handles without warning non-ASCII input strings which require no abbreviation.
- `read.dcf()` no longer has a limit of 8191 bytes per line. (Wish of PR#15250.)
- `formatC(x)` no longer copies the class of `x` to the result, to avoid misuse creating invalid objects as in PR#15303. A warning is given if a class is discarded.
- Dataset `npk` has been copied from [MASS](#) to allow more tests to be run without recommended packages being installed.
- The initialization of the regression coefficients for non-degenerate differenced models in `arima()` has been changed and in some examples avoids a local maximum. (PR#15396)
- `termplot()` now has an argument `transform.x` to control the display of individual terms in the plot. (PR#15329)
- `format()` now supports `digits = 0`, to display `nsmall` decimal places.
- There is a new read-only `par()` parameter called "page", which returns a logical value indicating whether the next `plot.new()` call will start a new page.
- Processing Sweave and Rd documents to PDF now renders backticks and single quotes better in several instances, including in `'\code'` and `'\samp'` expressions.
- `utils::modifyList()` gets a new argument `keep.null` allowing NULL components in the replacement to be retained, instead of causing corresponding components to be deleted.
- `tools::pkgVignettes()` gains argument check; if set to TRUE, it will warn when it appears a vignette requests a non-existent vignette engine.

UTILITIES

- R CMD check `--as-cran` checks the line widths in usage and examples sections of the package Rd files.
- R CMD check `--as-cran` now implies `'--timings'`.
- R CMD check looks for command `gfile` if a suitable file is not found. (Although file is not from GNU, OpenCSW on Solaris installs it as `gfile`.)
- R CMD build (with the internal tar) checks the permissions of `'configure'` and `'cleanup'` files and adds execute permission to the recorded permissions for these files if needed, with a warning. This is useful on OSes and file systems which do not support execute permissions (notably, on Windows).

- R CMD build now weaves and tangles all vignettes, so suggested packages are not required during package installation if the source tarball was prepared with current R CMD build.
- checkFF() (used by R CMD check) does a better job of detecting calls from other packages, including not reporting those where a function has been copied from another namespace (e.g. as a default method). It now reports calls where .NAME is a symbol registered in another package.
- On Unix-alike systems, R CMD INSTALL now installs packages group writably whenever the library (lib.loc) is group writable. Hence, update.packages() works for other group members (suggested originally and from a patch by Dirk Eddelbuettel).
- R CMD javareconf now supports the use of symbolic links for JAVA_HOME on platforms which have realpath. So it is now possible to use

```
R CMD javareconf JAVA_HOME=/usr/lib/jvm/java-1.7.0
```

on a Linux system and record that value rather than the frequently-changing full path such as `/usr/lib/jvm/java-1.7.0-openjdk-1.7.0.25.x86_64`.

- (Windows only.) Rscript -e requires a non-empty argument for consistency with Unix versions of R. (Also Rterm -e and R -e.)
- R CMD check does more thorough checking of declared packages and namespaces. It reports
 - packages declared in more than one of the ‘Depends’, ‘Imports’, ‘Suggests’ and ‘Enhances’ fields of the ‘DESCRIPTION’ file.
 - namespaces declared in ‘Imports’ but not imported from, neither in the ‘NAMESPACE’ file nor using the :: nor ::: operators.
 - packages which are used in library() or requires() calls in the R code but were already put on the search path *via* ‘Depends’.
 - packages declared in ‘Depends’ not imported *via* the ‘NAMESPACE’ file (except the standard packages). Objects used from ‘Depends’ packages should be imported to avoid conflicts and to allow correct operation when the namespace is loaded but not attached.
 - objects imported *via* ::: calls where :: would do.
 - objects imported by :: which are not exported.
 - objects imported by ::: calls which do not exist.

See ‘Writing R Extensions’ for good practice.

- R CMD check optionally checks for non-standard top-level files and directories (which are often mistakes): this is enabled for ‘--as-cran’.
- LaTeX style file upquote.sty is no longer included (the version was several years old): it is no longer used in R. A much later version is commonly included in LaTeX distributions but does not play well with the ae fonts which are the default for Sweave vignettes.
- R CMD build makes more use of the ‘build’ sub-directory of package sources, for example to record information about the vignettes.
- R CMD check analyses ::: calls.

INSTALLATION and INCLUDED SOFTWARE

- The macros used for the texinfo manuals have been changed to work better with the incompatible changes made in texinfo 5.x.
- The minimum version for a system xz library is now 5.0.3 (was 4.999). This is in part to avoid 5.0.2, which can compress in ways other versions cannot decompress.
- The included version of PCRE has been updated to 8.33.
- The included version of zlib has been updated to 1.2.8, a bug-fix release.
- The included version of xz utils's liblzma has been updated to 5.0.5.
- Since javareconf (see above) is used when R is installed, a stable link for JAVA_HOME can be supplied then.
- Configuring with '--disable-byte-compilation' will override the 'DESCRIPTION' files of recommended packages, which typically require byte-compilation.
- More of the installation and checking process will work even when TMPDIR is set to a path containing spaces, but this is not recommended and external software (such as texi2dvi) may fail.

PACKAGE INSTALLATION

- Installation is aborted immediately if a LinkingTo package is not installed.
- R CMD INSTALL has a new option --no-byte-compile which will override a 'ByteCompile' field in the package's 'DESCRIPTION' file.
- License 'BSD' is deprecated: use 'BSD_3_clause' or 'BSD_2_clause' instead.
License 'X11' is deprecated: use 'MIT' or 'BSD_2_clause' instead.
- Version requirements for LinkingTo packages are now recognized: they are checked at installation. (Fields with version requirements were previously silently ignored.)
- The limit of 500 S3method entries in a NAMESPACE file has been removed.
- The default 'version' of Bioconductor for its packages has been changed to the upcoming '2.13', but this can be set by the environment variable R_BIOC_VERSION, e.g. in file 'Renviron.site'.

C-LEVEL FACILITIES

- 'Rdefines.h' has been tweaked so it can be included in C++ code after 'R_ext/Boolean.h' (which is included by 'R.h').
Note that 'Rdefines.h' is not kept up-to-date, and 'Rinternals.h' is preferred for new code.
- eval and applyClosure are now protected against package code supplying an invalid rho.

DEPRECATED AND DEFUNCT

- The unused namespace argument to package.skeleton() is now formally deprecated and will be removed in R 3.1.0.
- plclust() is deprecated: use the plot() method for class "hclust" instead.
- Functions readNEWS() and checkNEWS() in package **tools** are deprecated (and they have not worked with current 'NEWS' files for a long time).

DOCUMENTATION

- ‘An Introduction to R’ has a new chapter on using R as a scripting language including interacting with the OS.

BUG FIXES

- `help.request()` could not determine the current version of R on CRAN. (PR#15241)
- On Windows, `file.info()` failed on root directories unless the path was terminated with an explicit `"."`. (PR#15302)
- The `regmatches<-()` replacement function mishandled results coming from `regexpr()`. (PR#15311)
- The help for `setClass()` and `representation()` still suggested the deprecated argument `representation=`. (PR#15312)
- R CMD config failed in an installed build of R 3.0.1 (only) when a sub-architecture was used. (Reported by Berwin Turlach.)
- On Windows, the installer modified the ‘etc/Rconsole’ and ‘etc/Rprofile.site’ files even when default options were chosen, so the MD5 sums did not refer to the installed versions. (Reported by Tal Galili.)
- `plot(hclust(), cex =)` respects `cex` again (and possibly others similarly). (Reported by Peter Langfelder.)
- If multiple packages were checked by R CMD check, and one was written for a different OS, it would set `--no-install` for all following packages as well as itself.
- `qr.coef()` and related functions did not properly coerce real vectors to complex when necessary. (PR#15332)
- `fTable(a)` now fixes up empty `dimnames` such that the result is printable.
- `package.skeleton()` was not starting its search for function objects in the correct place if environment was supplied. (Reported by Karl Forner.)
- Parsing code was changing the length field of vectors and confusing the memory manager. (PR#15345)
- The Fortran routine ZHER2K in the reference BLAS had a comment-out bug in two places. This caused trouble with `eigen()` for Hermitian matrices. (PR#15345 and report from Robin Hankin)
- `vignette()` and `browseVignettes()` did not display non-Sweave vignettes properly.
- Two warning/error messages have been corrected: the (optional) warning produced by a partial name match with a pairlist, the error message from a zero-length argument to the `:` operator. (Found by Radford Neal; PR#15358, PR#15356)
- `svd()` returned `NULL` rather than omitting components as documented. (Found by Radford Neal; PR#15360)
- `mclapply()` and `mcpParallel()` with `silent = TRUE` could break a process that uses `stdout` output unguarded against broken pipes (e.g., `zip` will fail silently). To work around such issues, they now replace `stdout` with a descriptor pointed to ‘/dev/null’ instead. For this purpose, internal `closeStdout` and `closeStderr` functions have gained the `to.null` flag.
- `log()`, `signif()` and `round()` now raise an error if a single named argument is not named `x`. (PR#15361)

- `deparse()` now deparses raw vectors in a form that is syntactically correct. (PR#15369)
- The jpeg driver in Sweave created a JPEG file, but gave it a `'png'` extension. (PR#15370)
- Deparsing of infix operators with named arguments is improved. (PR#15350)
- `mget()`, `seq.int()` and `numericDeriv()` did not duplicate arguments properly. (PR#15352, PR#15353, PR#15354)
- `kmeans(algorithm = "Hartigan-Wong")` now always stops iterating in the QTran stage. (PR#15364).
- `read.dcf()` re-allocated incorrectly and so could segfault when called on a file with lines of more than 100 bytes.
- On systems where `mktime()` does not set `errno`, the last second before the epoch could not be converted from `POSIXlt` to `POSIXct`. (Reported by Bill Dunlap.)
- `add1.glm()` miscalculated F-statistics when `df > 1`. (Bill Dunlap, PR#15386).
- `stem()` now discards infinite inputs rather than hanging. (PR#15376)
- The parser now enforces C99 syntax for floating point hexadecimal constants (e.g. `0x1.1p0`), rather than returning unintended values for malformed constants. (PR#15234)
- `model.matrix()` now works with very long LHS names (more than 500 bytes). (PR#15377)
- `integrate()` reverts to the pre-2.12.0 behaviour: from 2.12.0 to 3.0.1 it sometimes failed to achieve the requested tolerance and reported error estimates that were exceeded. (PR#15219)
- `strptime()` now handles `'%W'` fields with value 0. (PR#15915)
- R is now better protected against people trying to interact with the console in startup code. (PR#15325)
- Subsetting 1D arrays often lost `dimnames` (PR#15301).
- Unary `+` on a logical vector did not coerce to integer, although unary `-` did.
- `na.omit()` and `na.exclude()` added a row to a zero-row data frame. (PR#15399)
- All the (where necessary cut-down) vignettes are installed if R was configured with `'--without-recommended-packages'`.
- `source()` did not display filenames when reporting syntax errors.
- Syntax error reports misplaced the caret pointing out the bad token.
- (Windows only) Starting R with `R` (instead of `Rterm` or `Rgui`) would lose any zero-length strings from the command line arguments. (PR#15406)
- Errors in the encoding specified on the command line via `--encoding=foo` were not handled properly. (PR#15405)
- If `x` is a symbol, `is.vector(x, "name")` now returns `TRUE`, since `"name"` and `"symbol"` should be synonyms. (Reported by Hervé Pagès.)
- `R CMD rtags` works on platforms (such as OS X) with a XSI-conformant shell command `echo`. (PR#15231)
- `is.unsorted(NA)` returns `false` as documented (rather than `NA`).
- `R CMD LINK` did not know about sub-architectures.