

Changes in R

From 3.0.3 to 3.1.1

by the R Core Team

CHANGES IN R 3.1.1

NEW FEATURES

- When `attach()` reports conflicts, it does so compatibly with `library()` by using `message()`.
- R CMD Sweave no longer cleans any files by default, compatibly with versions of R prior to 3.1.0. There are new options `'--clean'`, `'--clean=default'` and `'--clean=keepOuts'`.
- `tools::buildVignette()` and `tools::buildVignettes()` with `clean = FALSE` no longer remove any created files. `buildvignette()` gains a `keep` argument for more cleaning customization.
- The Bioconductor `'version'` used by `setRepositories()` can now be set by environment variable `R_BIOC_VERSION` at runtime, not just when R is installed. (It has been stated that Bioconductor will switch from `'version'` 2.14 to `'version'` 3.0 during the lifetime of the R 3.1 series.)
- Error messages from bugs in embedded `'Sexpr'` code in Sweave documents now report the source location.
- `type.convert()`, `read.table()` and similar `read.*()` functions get a new `numerals` argument, specifying how numeric input is converted when its conversion to double precision loses accuracy. The default value, `"allow.loss"` allows accuracy loss, as in R versions before 3.1.0.
- For some compilers, integer addition could overflow without a warning. R's internal code for both integer addition and subtraction is more robust now. (PR#15774)
- The function determining the default number of knots for `smooth.spline()` is now exported, as `.nknots.smspl()`.
- `dbeta(, a, b)`, `pbeta()`, `qbeta()` and `rbeta()` are now defined also for $a = 0$, $b = 0$, or infinite a and b (where they typically returned `NaN` before).
- Many package authors report that the RStudio graphics device does not work correctly with their package's use of `dev.new()`. The new option `dev.new(noRStudioGD = TRUE)` replaces the RStudio override by the default device as selected by R itself, still respecting environment variables `R_INTERACTIVE_DEVICE` and `R_DEFAULT_DEVICE`.
- `readRDS()` now returns visibly.
- Modifying internal logical scalar constants now results in an error instead of a warning.
- `install.packages(repos = NULL)` now accepts `http://` or `ftp://` URLs of package archives as well as file paths, and will download as required. In most cases `repos = NULL` can be deduced from the extension of the URL.
- The warning when using partial matching with the `$` operator on data frames is now only given when options(`"warnPartialMatchDollar"`) is `TRUE`.
- Package help requests like `package?foo` now try the package `foo` whether loaded or not.

- General help requests now default to trying all loaded packages, not just those on the search path.
- Added a new function `promptImport()`, to generate a help page for a function that was imported from another package (and presumably re-exported, or help would not be needed).

INSTALLATION and INCLUDED SOFTWARE

- configure option `'--with-internal-tzcode'` can now be used with variable `rsharedir`.
- The included version of PCRE has been updated to 8.35.
- There is a new target `make uninstall-libR` to remove an installed shared/static 'libR'. `make install-libR` now works if a sub-architecture is used, although the user will need to specify `libdir` differently for different sub-architectures.
- There is more extensive advice on which LaTeX packages are required to install R or to make package manuals (as done by `R CMD check`) in the 'Writing R Extensions' manual.
- Compilers/linkers were handling the visibility control in `'src/extra/xz'` inconsistently (and apparently in some cases incorrectly), so it has been simplified. (PR#15327)
- (Windows) There is updated support for the use of ICU for collation: see the 'R Installation and Administration Manual'.

BUG FIXES

- `dbinom(x,n)`, `pbinom()`, `dpois()`, etc, are slightly less restrictive in checking if `n` is integer-valued. (Wish of PR#15734.)
- `pchisq(x,df,ncp,log.p = TRUE)` is more accurate and no longer underflows for small `x` and `ncp < 80`, e.g, for `pchisq(1e-5,df = 100,ncp = 1,log = TRUE)`. (Based on PR#15635 and a suggestion by Roby Joehanes.)
- The `s` ("step into") command in the debugger would cause R to step into expressions evaluated there, not just into functions being debugged. (PR#15770)
- The C code used by `strptime()` rejected time-zone offsets of more than +1200 (+1245, +1300 and +1400 can occur). (PR#15768)
- (Windows only.) `png(type = "cairo",antialias = "gray")` was not accepted. (PR#15760)
- Use of `save(...,envir=)` with named objects could fail. (PR#15758)
- `Sweave()` mis-parsed 'Sexpr' expressions that contained backslashes. (PR#15779)
- The return value from `options(foo = NULL)` was not the previous value of the option. (PR#15781)
- `enc2utf8()` and `enc2native()` did not always mark the encoding of the return values when it was known.
- `dnbinom(x,size = <large>,mu,log = TRUE)` no longer underflows to `-Inf` for large `mu`, thanks to a suggestion from Alessandro Mammana (MPI MolGen, Berlin).
- `pbeta(x,a,b,log = TRUE)` no longer behaves discontinuously (in a small `x`-region) because of denormalized numbers. Also, `pbeta(1-1e-12,1e30,1.001,log=TRUE)` now terminates "in real time".

- The "CRAN" filter (see `available.packages()`) no longer removes duplicates other than of packages on CRAN, and does not fail if there is no CRAN repository in `getOption("repos")`.
- The device listing from `dev2bitmap()` and `bitmap()` was truncated to 1000 characters: modern versions of GhostScript on most platforms have many more devices.
- (Windows.) Commands such as `Sys.which()` and `pipe()` which needed to find the full path to a command could segfault if the 'long' path name was much longer than the 'short' path name (which `Sys.which()` returns), as the behaviour of the Windows API call had changed.
- R CMD build will fail with an error if one of the packages specified in the 'VignetteBuilder' field is not installed. (Without loading those packages it cannot be ascertained which files are intended to be vignettes. This means that the 'VignetteBuilder' packages have to be installed for package checking too.) (Wish of PR#15775.)
- Misguided attempts to use `chull()` with non-finite points now give an error (related to PR#15777).
- For a formula with exactly 32 variables the 32nd variable was aliased to the intercept in some C-level computations of terms, so that for example attempting to remove it would remove the intercept instead (and leave a corrupt internal structure). (PR#15735)
- `anyDuplicated()` silently returned wrong values when the first duplicate was at an index which was too large to be stored in an integer vector (although a lot of RAM and patience would have been needed to encounter this).
- `tools::Rd2ex(commentDontrun = FALSE)` failed if the block had only one line.
- Hexadecimal constants such as `0x110p-5L` which were incorrectly qualified by `L` were parsed incorrectly since R 3.0.0, with a slightly garbled warning. (PR#15753)
- `system()` returned success on some platforms even if the system was unable to launch a process. (PR#15796)
- (Windows Rgui console.) Unbuffered output was sometimes not output immediately if the prompt was not on the last line of the console.
- The built-in help server did not declare the encoding for the 'DESCRIPTION' or other text files to be the package encoding, so non-ASCII characters could be displayed incorrectly.
- R is now trying harder to not cleanup child processes that were not spawned by `mcpipeline()` on platforms that provide information about the source process of the SIGCHLD signal. This allows 3rd party libraries to manage the exit status of children that they spawn without R interfering.
- `mcmapply()` was only parallelizing if the number of jobs was bigger than the number of cores. It now parallelizes if the number of jobs is more than one.
- Auto-printing would re-evaluate its argument when trying to dispatch to a print method. This is now avoided when possible.
- Unserializing (including `load()` and `readRDS()`) could silently return incorrect numeric values from ASCII saves if there was a read error.
- `getParseData()` could return incorrect values for the parents of some elements. (Reported by Andrew Redd.)
- Attempting to use data frames of 2^{31} or more rows with `merge()` or to create a merged data frame of that size now gives a clearer error message.

- `parse()` did not check its `file` argument was a connection if it was not a character string, so e.g. `parse(FALSE)` attempted to read from `stdin`.
Nor did `dump()` and `dput()`.
- The `"help.try.all.packages"` option was ignored when the shortcut syntax for help was used, e.g. `?foo`.
- A potential segfault in string allocation has been fixed. (Found by Radford Neal.)
- Potential memory protection errors in `sort()` and `D()` have been fixed. (Found by Radford Neal.)
- Fixed a lack of error checking in graphics event functions. (Found by Radford Neal; a different patch used here than the one in `pkgR`.)
- `numericDeriv()` sometimes miscalculated the gradient. (PR#15849, reported originally by Radford Neal)

CHANGES IN R 3.1.0

NEW FEATURES

- `type.convert()` (and hence by default `read.table()`) returns a character vector or factor when representing a numeric input as a double would lose accuracy. Similarly for complex inputs.
If a file contains numeric data with unrepresentable numbers of decimal places that are intended to be read as numeric, specify `colClasses` in `read.table()` to be `"numeric"`.
- `tools::Rdiff(useDiff = FALSE)` is closer to the POSIX definition of `diff -b` (as distinct from the description in the man pages of most systems).
- New function `anyNA()`, a version of `any(is.na(.))` which is fast for atomic vectors, based on a proposal by Tim Hesterberg. (Wish of PR#15239.)
- `arrayInd(*, useNames = TRUE)` and, analogously, `which(*, arr.ind = TRUE)` now make use of `names(.dimnames)` when available.
- `is.unsorted()` now also works for raw vectors.
- The `"table"` method for `as.data.frame()` (also useful as `as.data.frame.table()`) now passes `sep` and `base` arguments to `provideDimnames()`.
- `uniroot()` gets new optional arguments, notably `extendInt`, allowing to auto-extend the search interval when needed. The return value has an extra component, `init.it`.
- `switch(f, ...)` now warns when `f` is a factor, as this typically happens accidentally where the user meant to pass a character string, but `f` is treated as integer (as always documented).
- The parser has been modified to use less memory.
- The way the unary operators (`+ -!`) handle attributes is now more consistent. If there is no coercion, all attributes (including class) are copied from the input to the result: otherwise only names, dims and dimnames are.
- `colorRamp()` and `colorRampPalette()` now allow non-opaque colours and a ramp in opacity via the new argument `alpha = TRUE`. (Suggested by Alberto Krone-Martins, but optionally as there are existing uses which expect only RGB values.)
- `grid.show.layout()` and `grid.show.viewport()` get an optional `vp.ex` argument.

- There is a new function `find_gs_cmd()` in the **tools** package to locate a GhostScript executable. (This is an enhanced version of a previously internal function there.)
- `object.size()` gains a `format()` method.
- There is a new family, "ArialMT", for the `pdf()` and `postscript()` devices. This will only be rendered correctly on viewers which have access to Monotype TrueType fonts (which are sometimes requested by journals).
- The text and PDF news files, including 'NEWS' and 'NEWS.2', have been moved to the 'doc' directory.
- `combn(x, simplify = TRUE)` now gives a factor result for factor input `x` (previously user error). (Related to PR#15442.)
- Added `utils::fileSnapshot()` and `utils::changedFiles()` functions to allow snapshots and comparison of directories of files.
- `make.names(names, unique=TRUE)` now tries to preserve existing names. (Suggestion of PR#15452.)
- New functions `cospi(x)`, `sinpi(x)`, and `tanpi(x)`, for more accurate computation of `cos(pi*x)`, etc, both in R and the C API. Using these gains accuracy in some cases, e.g., inside `lgamma()` or `besselI()`. (Suggested by Morten Welinder in PR#15529.)
- `print.table(x, zero.print = ".")` now also has an effect when `x` is not integer-valued.
- There is more support to explore the system's idea of time-zone names. `Sys.timezone()` tries to give the current system setting by name (and succeeds at least on Linux, OS X, Solaris and Windows), and `OlsonNames()` lists the names in the system's Olson database. `Sys.timezone(location = FALSE)` gives the previous behaviour.
- Platforms with a 64-bit `time_t` type are allowed to handle conversions between the "POSIXct" and "POSIXlt" classes for date-times outside the 32-bit range (before 1902 or after 2037): the existing workarounds are used on other platforms. (Note that time-zone information for post-2037 is speculative at best, and the OS services are tested for known errors and so not used on OS X.)
Currently `time_t` is usually `long` and hence 64-bit on Unix-alike 64-bit platforms: however in several cases the time-zone database is 32-bit. For R for Windows it is 64-bit (for both architectures as from this version).
- The "save.defaults" option can include a value for `compression_level`. (Wish of PR#15579.)
- `colSums()` and friends now have support for arrays and data-frame columns with 2^{31} or more elements.
- `as.factor()` is faster when `f` is an unclassed integer vector (for example, when called from `tapply()`).
- `fft()` now works with longer inputs, from the 12 million previously supported up to 2 billion. (PR#15593)
- Complex `svd()` now uses LAPACK subroutine ZGESDD, the complex analogue of the routine used for the real case.
- Sweave now outputs '.tex' files in UTF-8 if the input encoding is declared to be UTF-8, regardless of the local encoding. The UTF-8 encoding may now be declared using a LaTeX comment containing the string `%\SweaveUTF8` on a line by itself.
- `file.copy()` gains a `copy.date` argument.

- Printing of date-times will make use of the time-zone abbreviation in use at the time, if known. For example, for Paris pre-1940 this could be 'LMT', 'PMT', 'WET' or 'WEST'. To enable this, the "POSIXt" class has an optional component "zone" recording the abbreviation for each element.

For platforms which support it, there is also a component "gmtoff" recording the offset from GMT where known.

- (On Windows, by default on OS X and optionally elsewhere.) The system C function `strftime` has been replaced by a more comprehensive version with closer conformance to the POSIX 2008 standard.
- `dnorm(x, log = FALSE)` is more accurate (but somewhat slower) for $|x| > 5$; as suggested in PR#15620.
- Some versions of the `tiff()` device have further compression options.
- `read.table()`, `readLines()` and `scan()` have a new argument to influence the treatment of embedded nuls.
- Avoid duplicating the right hand side values in complex assignments when possible. This reduces copying of replacement values in expressions such as `Z$a <-a0` and `ans[[i]] <-tmp`: some package code has relied on there being copies.

Also, a number of other changes to reduce copying of objects; all contributed by or based on suggestions by Michael Lawrence.

- The `fast` argument of `KalmanLike()`, `KalmanRun()` and `KalmanForecast()` has been replaced by `update`, which instead of updating `mod` in place, optionally returns the updated model in an attribute "mod" of the return value.
- `arma()` and `makeARIMA()` get a new optional argument `SSinit`, allowing the choice of a different state space initialization which has been observed to be more reliable close to non-stationarity: see PR#14682.
- `warning()` has a new argument `noBreaks.`, to simplify post-processing of output with `options(warn = 1)`.
- `pushBack()` gains an argument `encoding`, to support reading of UTF-8 characters using `scan()`, `read.table()` and related functions in a non-UTF-8 locale.
- `all.equal.list()` gets a new argument `use.names` which by default labels differing components by names (if they match) rather than by integer index. Saved R output in packages may need to be updated.
- The methods for `all.equal()` and `attr.all.equal()` now have argument `check.attributes` after `...` so it cannot be partially nor positionally matched (as it has been, unintentionally).
A side effect is that some previously undetected errors of passing empty arguments (no object between commas) to `all.equal()` are detected and reported.
There are explicit checks that `check.attributes` is logical, `tolerance` is numeric and `scale` is NULL or numeric. This catches some unintended positional matching.
The message for `all.equal.numeric()` reports a "scaled difference" only for `scale != 1`.
- `all.equal()` now has a "POSIXt" method replacing the "POSIXct" method.
- The "Date" and "POSIXt" methods of `seq()` allows `by = "quarter"` for completeness (by = "3 months" always worked).

- `file.path()` removes any trailing separator on Windows, where they are invalid (although sometimes accepted). This is intended to enhance the portability of code written by those using POSIX file systems (where a trailing `/` can be used to confine path matching to directories).
- New function `agrep1()` which like `grep1()` returns a logical vector.
- `fifo()` is now supported on Windows. (PR#15600)
- `sort.list(method = "radix")` now allows negative integers (wish of PR#15644).
- Some functionality of `print.ts()` is now available in `.preformat.ts()` for more modularity.
- `mcpipeline()` gains an option `detach = TRUE` which allows execution of code independently of the current session. It is based on a new `estranged = TRUE` argument to `mcfork()` which forks child processes such that they become independent of the parent process.
- The `pdf()` device omits circles and text at extremely small sizes, since some viewers were failing on such files.
- The rightmost break for the "months", "quarters" and "years" cases of `hist.POSIXlt()` has been increased by a day. (Inter alia, fixes PR#15717.)
- The handling of `DF[i,] <-a` where `i` is of length 0 is improved. (Inter alia, fixes PR#15718.)
- `hclust()` gains a new method "ward.D2" which implements Ward's method correctly. The previous "ward" method is "ward.D" now, with the old name still working. Thanks to research and proposals by Pierre Legendre.
- The `sunspot.month` dataset has been amended and updated from the official source, whereas the `sunspots` and `sunspot.year` datasets will remain immutable. The documentation and source links have been updated correspondingly.
- The `summary()` method for "lm" fits warns if the fit is essentially perfect, as most of the summary may be computed inaccurately (and with platform-dependent values). Programmers who use `summary()` in order to extract just a component which will be reliable (e.g. `$cov.unscaled`) should wrap their calls in `suppressWarnings()`.

INSTALLATION and INCLUDED SOFTWARE

- The included version of LAPACK has been updated to 3.5.0.
- There is some support for parallel testing of an installation, by setting `TEST_MC_CORES` to an integer greater than one to indicate the maximum number of cores to be used in parallel. (It is worth specifying at least 8 cores if available.) Most of these require a make program (such as GNU make and dmake) which supports the `$MAKE -j nproc` syntax.
 Except on Windows: the tests of standard package examples in `make check` are done in parallel. This also applies to running `tools::testInstalledPackages()`.
 The more time-consuming regression tests are done in parallel.
 The package checks in `make check-devel` and `make check-recommended` are done in parallel.
- More of `make check` will work if recommended packages are not installed: but recommended packages remain needed for thorough checking of an R build.
- The version of 'tzcode' included in 'src/extra/tzone' has been updated. (Formerly used only on Windows.)

- The included (64-bit) time-zone conversion code and Olson time-zone database can be used instead of the system version: use configure option `'--with-internal-tzcode'`. This is the default on Windows and OS X. (Note that this does not currently work if a non-default `rsharedir` configure variable is used.)

(It might be necessary to set environment variable `TZ` on OSes where this is not already set, although the system timezone is deduced correctly on at least Linux, OS X and Windows.)

This option also switches to the version of `strftime` included in directory `'src/extra/tzone'`.

- configure now tests for a C++11-compliant compiler by testing some basic features. This by default tries flags for the compiler specified by `'CXX'`, but an alternative compiler, options and standard can be specified by variables `'CXX1X'`, `'CXX1XFLAGS'` and `'CXX1XSTD'` (e.g. `'-std=gnu++11'`).
- R can now optionally be compiled to use reference counting instead of the NAMED mechanism by defining `SWITCH_TO_REFCNT` in `'Rinternals.h'`. This may become the default in the future.
- There is a new option `'--use-system-tre'` to use a suitable system `tre` library: at present this means a version from their git repository, after corrections. (Wish of PR#15660.)

PACKAGE INSTALLATION

- The CRANextra repository is no longer a default repository on Windows: all the binary versions of packages from CRAN are now on CRAN, although CRANextra contains packages from Omegahat and elsewhere used by CRAN packages.
- Only vignettes sources in directory `'vignettes'` are considered to be vignettes and hence indexed as such.
- In the `'DESCRIPTION'` file,

```
License: X11
```

is no longer recognized as valid. Use `'MIT'` or `'BSD_2_clause'` instead, both of which need `'+ file LICENSE'`.

- For consistency, entries in `'.Rinstignore'` are now matched case-insensitively on all platforms.
- Help for S4 methods with very long signatures now tries harder to split the description in the `'Usage'` field to no more than 80 characters per line (some packages had over 120 characters).
- `R CMD INSTALL --build` (not Windows) now defaults to the internal `tar()` unless `R_INSTALL_TAR` is set.
- There is support for compiling C++11 code in packages on suitable platforms: see `'Writing R Extensions'`.
- Fake installs now install the contents of directory `'inst'`: some packages use this to install e.g. C++ headers for use by other packages that are independent of the package itself. Option `'--no-inst'` can be used to get the previous behaviour.

DEBUGGING

- The behaviour of the code browser has been made more consistent, in part following the suggestions in PR#14985.
- Calls to `browser()` are now consistent with calls to the browser triggered by `debug()`, in that Enter will default to `n` rather than `c`.
- A new browser command `s` has been added, to “step into” function calls.
- A new browser command `f` has been added, to “finish” the current loop or function.
- Within the browser, the command `help` will display a short list of available commands.

UTILITIES

- Only vignettes sources in directory ‘vignettes’ are considered to be vignettes by `R CMD check`. That has been the preferred location since R 2.14.0 and is now obligatory.
- For consistency, `R CMD build` now matches entries in ‘.Rbuildignore’ and ‘vignettes/install_extras’ case-insensitively on all platforms (not just on Windows).
- `checkFF()` (called by `R CMD check` by default) can optionally check foreign function calls for consistency with the registered type and argument count. This is the default for `R CMD check --as-cran` or can be enabled by setting environment variable `_R_CHECK_FF_CALLS_` to ‘registration’ (but is in any case suppressed by ‘--install=no’). Because this checks calls in which `.NAME` is an R object and not just a literal character string, some other problems are detected for such calls.

Functions `suppressForeignCheck()` and `dontCheck()` have been added to allow package authors to suppress false positive reports.

- `R CMD check --as-cran` warns about a false value of the ‘DESCRIPTION’ field ‘BuildVignettes’ for Open Source packages, and ignores it. (An Open Source package needs to have complete sources for its vignettes which should be usable on a suitably well-equipped system).
- `R CMD check --no-rebuild-vignettes` is defunct:
`R CMD check --no-build-vignettes` has been preferred since R 3.0.0.
- `R CMD build --no-vignettes` is defunct:
`R CMD build --no-build-vignettes` has been preferred since R 3.0.0.
- `R CMD Sweave` and `R CMD Stangle` now process both Sweave and non-Sweave vignettes. The `tools::buildVignette()` function has been added to do the same tasks from within R.
- The flags returned by `R CMD config --ldflags` and (where installed) `pkg-config --libs libR` are now those needed to link a front-end against the (shared or static) R library.
- ‘Sweave.sty’ has a new option ‘[inconsolata]’.
- `R CMD check` customizations such as `_R_CHECK_DEPENDS_ONLY_` make available packages only in ‘LinkingTo’ only for installation, and not for loading/runtime tests.
- `tools::checkFF()` reports on `.C` and `.Fortran` calls with `DUP = FALSE` if argument `check_DUP` is true. This is selected by `R CMD check` by default.
- `R CMD check --use-gct` can be tuned to garbage-collect less frequently using `gctorture2()` via the setting of environment variable `_R_CHECK_GCT_N_`.
- Where supported, `tools::texi2dvi()` limits the number of passes tried to 20.

C-LEVEL FACILITIES

- (Windows only) A function `R_WaitEvent()` has been added (with declaration in header `R.h`) to block execution until the next event is received by R.
- Remapping in the `Rmath.h` header can be suppressed by defining `R_NO_REMAP_RMATH`.
- The remapping of `rround()` in header `Rmath.h` has been removed: use `fround()` instead.
- `ftrunc()` in header `Rmath.h` is now a wrapper for the C99 function `trunc()`, which might as well be used in C code: `ftrunc()` is still needed for portable C++ code.
- The never-documented remapping of `prec()` to `fprec()` in header `Rmath.h` has been removed.
- The included LAPACK subset now contains ZGESDD and ZGELSD.
- The function `LENGTH()` now checks that it is only applied to vector arguments. However, in packages `length()` should be used. (In R itself `LENGTH()` is a macro without the function overhead of `length()`.)
- Calls to `SET_VECTOR_ELT()` and `SET_STRING_ELT()` are now checked for indices which are in-range: several packages were writing one element beyond the allocated length.
- `allocVector3` has been added which allows custom allocators to be used for individual vector allocations.

DEPRECATED AND DEFUNCT

- `chol(pivot = TRUE, LINPACK = TRUE)` is defunct.
Arguments `EISPACK` for `eigen()` and `LINPACK` for `chol()`, `chol2inv()`, `solve()` and `svd()` are ignored: LAPACK is always used.
- `.find.package()` and `.path.package()` are defunct: only the versions without the initial dot introduced in R 2.13.0 have ever been in the API.
- Partial matching when using the `$` operator *on data frames* now throws a warning and may become defunct in the future. If partial matching is intended, replace `foo$bar` by `foo[["bar", exact = FALSE]]`.
- The long-deprecated use of `\synopsis` in the `'Usage'` section of `'Rd'` files has been removed: such sections are now ignored (with a warning).
- `package.skeleton()`'s deprecated argument `namespace` has been removed.
- Many methods are no longer exported by package `stats`. They are all registered on their generic, which should be called rather than calling a method directly.
- Functions `readNEWS()` and `checkNEWS()` in package `tools` are defunct.
- `download.file(method = "lynx")` is deprecated.
- `.C(DUP = FALSE)` and `.Fortran(DUP = FALSE)` are now deprecated, and may be disabled in future versions of R. As their help has long said, `.Call()` is much preferred. R CMD check notes such usages (by default).
- The workaround of setting `R_OSX_VALGRIND` has been removed: it is not needed in current `valgrind`.

BUG FIXES

- Calling `lm.wfit()` with no non-zero weights gave an array-overflow in the Fortran code and a not very sensible answer. It is now special-cased with a simpler answer (no qr component).
- Error messages involving non-syntactic names (e.g. as produced by ``\r`` when that object does not exist) now encode the control characters. (Reported by Hadley Wickham.)
- `getGraphicsEvent()` caused 100% usage of one CPU in Windows. (PR#15500)
- `nls()` with no `start` argument may now work inside another function (scoping issue).
- `pbeta()` and similar work better for very large (billions) `ncp`.
- Where time zones have changed abbreviations over the years, the software tries to more consistently use the abbreviation appropriate to the time or if that is unknown, the current abbreviation. On some platforms where the C function `localtime` changed the `tzname` variables the reported abbreviation could have been that of the last time converted.
- `all.equal(list(1), identity)` now works.
- Bug fix for pushing viewports in **grid** (reported by JJ Allaire and Kevin Ushey).
NOTE for anyone poking around within the graphics engine display list (despite the warnings not to) that this changes what is recorded by **grid** on the graphics engine display list.
- Extra checks have been added for unit resolution and conversion in **grid**, to catch instances of division-by-zero. This may introduce error messages in existing code and/or produce a different result in existing code (but only where a non-finite location or dimension may now become zero).
- Some bugs in TRE have been corrected by updating from the git repository. This allows R to be installed on some platforms for which this was a blocker (PR#15087 suggests Linux on ARM and HP-UX).
- `?` applied to a call to an S4 generic failed in several cases. (PR#15680)
- The implicit S4 generics for primitives with `...` in their argument list were incorrect. (PR#15690)
- Bug fixes to `methods::callGeneric()`. (PR#15691)
- The bug fix to `aggregate()` in PR#15004 introduced a new bug in the case of no grouping variables. (PR#15699)
- In rare cases printing deeply nested lists overran a buffer by one byte and on a few platforms segfaulted. (PR#15679)
- The dendrogram method of `as.dendrogram()` was hidden accidentally, (PR#15703), and `order.dendrogram(d)` gave too much for a leaf `d`. (PR#15702)
- R would try to kill processes on exit that have pids ever used by a child process spawned by `mcpParallel` even though the current process with that pid was not actually its child.
- `cophenetic()` applied to a "dendrogram" object sometimes incorrectly returned a "Labels" attribute with dimensions. (PR#15706)
- `printCoefmat()` called from quite a few `print()` methods now obeys small `getOption("width")` settings, line wrapping the "signif. codes" legend appropriately. (PR#15708)

- `model.matrix()` assumed that the stored dimnames for a matrix was NULL or length 2, but length 1 occurred.
- The clipping region for a device was sometimes used in base graphics before it was set.

CHANGES IN R 3.0.3

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.)
- On a Unix-alike `Sys.timezone()` returns NA if the environment variable TZ is unset, to distinguish it from an empty string which on some OSes means the 'UTC' time zone.
- The backtick may now be escaped in strings, to allow names containing them to be constructed, e.g. ````. (PR#15621)`
- `read.table()`, `readLines()` and `scan()` now warn when an embedded nul is found in the input. (Related to PR#15625 which was puzzled by the behaviour in this unsupported case.)
- (Windows only.) `file.smlink()` works around the undocumented restriction of the Windows system call to backslashes. (Wish of PR#15631.)
- `KalmanForecast(fast = FALSE)` is now the default, and the help contains an example of how `fast = TRUE` can be used in this version. (The usage will change in 3.1.0.)
- `strptime()` now checks the locale only when locale-specific formats are used and caches the locale in use: this can halve the time taken on OSes with slow system functions (e.g. OS X).
- `strptime()` and the `format()` methods for classes "POSIXct", "POSIXlt" and "Date" recognize strings with marked encodings: this allows, for example, UTF-8 French month names to be read on (French) Windows.
- `iconv(to = "utf8")` is now accepted on all platforms (some implementations did already, but GNU **libiconv** did not: however converted strings were not marked as being in UTF-8). The official name, "UTF-8" is still preferred.
- `available.packages()` is better protected against corrupt metadata files. (A recurring problem with Debian package **shogun-r**: PR#14713.)
- Finalizers are marked to be run at garbage collection, but run only at a somewhat safer later time (when interrupts are checked). This circumvents some problems with finalizers running arbitrary code during garbage collection (the known instances being `running.options()` and (C-level) `path.expand()` re-entrantly).

INSTALLATION and INCLUDED SOFTWARE

- The included version of PCRE has been updated to 8.34. This fixes bugs and makes the behaviour closer to Perl 5.18. In particular, the concept of 'space' includes 'VT' and hence agrees with POSIX's.

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.
- `install.packages(dependencies = value)` for `value = NA` (the default) or `value = TRUE` omits packages only in LinkingTo for binary package installs.

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'.

- The code for date-times before 1902 assumed that the offset from GMT in 1902 was a whole number of minutes: that was not true of Paris (as recorded on some platforms).
- Using `Sys.setLocale` to set `LC_NUMERIC` to "C" (to restore the sane behavior) no longer gives a warning.
- `deparse()` now deparses complex vectors in a way that re-parses to the original values. (PR#15534, patch based on code submitted by Alex Bertram.)
- In some extreme cases (more than 10^{15}) integer inputs to `dpqrxxx()` functions might have been rounded up by one (with a warning about being non-integer). (PR#15624)
- Plotting symbol `pch = 14` had the triangle upside down on some devices (typically screen devices). The triangle is supposed to be point up. (Reported by Bill Venables.)
- `getSrcref()` did not work on method definitions if `rematchDefinition()` had been used.
- `KalmanForecast(fast = FALSE)` reported a (harmless) stack imbalance.
- The count of observations used by `KalmanRun()` did not take missing values into account.
- In locales where the abbreviated name of one month is a partial match for the full name of a later one, the `%B` format in `strptime()` could fail. An example was French on OS X, where 'juin' is abbreviated to 'jui' and partially matches juillet. Similarly for weekday names.
- `pbeta(x, a, b, log.p = TRUE)` sometimes underflowed to zero for very small and very differently sized `a, b`. (PR#15641)
- `approx()` and `approxfun()` now handle infinite values with the "constant" method. (PR#15655)
- `stripchart()` again respects reversed limits in `xlim` and `ylim`. (PR#15664)