Changes in R

From version 3.2.2 to version 3.2.3

by the R Core Team

CHANGES IN R 3.3.1 patched

NEW FEATURES

- extSoftVersion() now reports the version (if any) of the readline library in use.

- Convenience function hasName() has been added; it is intended to replace the common idiom !is.null(x$name) without the usually unintended partial name matching.

- The version of LAPACK included in the sources has been updated to 3.6.1, a bug-fix release including a speedup for the non-symmetric case of eigen().

- Use options(deparse.max.lines) to limit the number of lines recorded in .Traceback and other deparsing activities.

INSTALLATION and INCLUDED SOFTWARE

- Versions of the readline library >= 6.3 had been changed so that terminal window resizes were not signalled to readline: code has been added using a explicit signal handler to work around that (when R is compiled against readline >= 6.3). (PR#16604)

- configure works better with Oracle Developer Studio 12.5.

UTILITIES

- R CMD check reports more dubious flags in files 'src/Makevars[.in]', including '-w' and '-g'.

- R CMD check has been set up to filter important warnings from recent versions of gfortran with '-Wall -pedantic': this now reports non-portable GNU extensions such as out-of-order declarations.

BUG FIXES

- The check for non-portable flags in R CMD check could be stymied by 'src/Makevars' files which contained targets.

- (Windows only) When using certain desktop themes in Windows 7 or higher, Alt-Tab could cause Rterm to stop accepting input. (PR#14406; patch submitted by Jan Gleixner.)

- pretty(d,..) behaves better for date-time d (PR#16923).

- When a class name matches multiple classes in the cache, perform a dynamic search in order to obey namespace imports. This should eliminate annoying messages about multiple hits in the class cache. Also, pass along the package from the ClassExtends object when looking up superclasses in the cache.

- sample(NA_real_) now works.

- Packages using non-ASCII encodings in their code did not install data properly on systems using different encodings.

- merge(df1,df2) now also works for data frames with column names "na.last", "decreasing", or "method". (PR#17119)

- contour() caused a segfault if the labels argument had length zero. (Reported by Bill Dunlap.)

- unique(warnings()) works more correctly, thanks to a new duplicated.warnings() method.

- findInterval(x,vec = numeric(),all.inside = TRUE) now returns 0s as documented. (Reported by Bill Dunlap.)

- (Windows only) R CMD SHLIB failed when a symbol in the resulting library had the same name as a keyword in the '.def' file. (PR#17130)

- pmax() and pmin() now work with (more ?) classed objects, such as "Matrix" from the **Matrix** package, as documented for a long time.

- axis(side,x = D) and hence Axis() and plot() now work correctly for "Date" and time objects D, even when "time goes backward", e.g., with decreasing xlim. (Reported by William May).

CHANGES IN R 3.3.1

BUG FIXES

- R CMD INSTALL and hence install.packages() gave an internal error installing a package called **description** from a tarball on a case-insensitive file system.

- match(x,t) (and hence x %in% t) failed when x was of length one, and either character and x and t only differed in their Encoding or when x and t where complex with NAs or NaNs. (PR#16885.)

- unloadNamespace(ns) also works again when ns is a 'namespace', as from getNamespace().

- rgamma(1,Inf) or rgamma(1,0,0) no longer give NaN but the correct limit.

- length(baseenv()) is correct now.

- pretty(d,..) for date-time d rarely failed when "halfmonth" time steps were tried (PR#16923) and on 'inaccurate' platforms such as 32-bit Windows or a configuration with --disable-long-double; see comment #15 of PR#16761.

- In text.default(x,y,labels), the rarely(?) used default for labels is now correct also for the case of a 2-column matrix x and missing y.

- as.factor(c(a = 1L)) preserves names() again as in R < 3.1.0.

- strtrim(""[0],0[0]) now works.

- Use of Ctrl-C to terminate a reverse incremental search started by Ctrl-R in the readline-based Unix terminal interface is now supported when R was compiled against readline >= 6.0 (Ctrl-G always worked). (PR#16603)

- diff(<difftime>) now keeps the "units" attribute, as subtraction already did, PR#16940.

CHANGES IN R 3.3.0

SIGNIFICANT USER-VISIBLE CHANGES

- `nchar(x,*)`'s argument `keepNA` governing how the result for NAs in x is determined, gets a new default `keepNA = NA` which returns NA where x is NA, except for `type = "width"` which still returns 2, the formatting / printing width of NA.

- All builds have support for '`https:`' URLs in the default methods for `download.file()`, `url()` and code making use of them.

  Unfortunately that cannot guarantee that any particular '`https:`' URL can be accessed. For example, server and client have to successfully negotiate a cryptographic protocol (TLS/SSL, . . .) and the server's identity has to be verifiable *via* the available certificates. Different access methods may allow different protocols or use private certificate bundles: we encountered a '`https:`' CRAN mirror which could be accessed by one browser but not by another nor by `download.file()` on the same Linux machine.

NEW FEATURES

- The `print` method for `methods()` gains a `byclass` argument.

- New functions `validEnc()` and `validUTF8()` to give access to the validity checks for inputs used by `grep()` and friends.

- Experimental new functionality for S3 method checking, notably `isS3method()`.

  Also, the names of the R 'language elements' are exported as character vector `tools::langElts`.

- `str(x)` now displays `"Time-Series"` also for matrix (multivariate) time-series, i.e. when `is.ts(x)` is true.

- (Windows only) The GUI menu item to install local packages now accepts '*.tar.gz' files as well as '*.zip' files (but defaults to the latter).

- New programmeR's utility function `chkDots()`.

- `D()` now signals an error when given invalid input, rather than silently returning NA. (Request of John Nash.)

- `formula` objects are slightly more "first class": e.g., `formula()` or `new("formula",y ~ x)` are now valid. Similarly, for `"table"`, `"ordered"` and `"summary.table"`. Packages defining S4 classes with the above S3/S4 classes as slots should be reinstalled.

- New function `strrep()` for repeating the elements of a character vector.

- `rapply()` preserves attributes on the list when `how = "replace"`.

- New S3 generic function `sigma()` with methods for extracting the estimated standard deviation aka "residual standard deviation" from a fitted model.

- `news()` now displays R and package news files within the HTML help system if it is available. If no news file is found, a visible NULL is returned to the console.

- `as.raster(x)` now also accepts `raw` arrays x assuming values in `0:255`.

- Subscripting of matrix/array objects of type `"expression"` is now supported.

- `type.convert("i")` now returns a factor instead of a complex value with zero real part and missing imaginary part.

- Graphics devices `cairo_pdf()` and `cairo_ps()` now allow non-default values of the cairographics 'fallback resolution' to be set.

  This now defaults to 300 on all platforms: that is the default documented by cairographics, but apparently was not used by all system installations.

- `file()` gains an explicit `method` argument rather than implicitly using `getOption("url.method","default")`.

- Thanks to a patch from Tomas Kalibera, `x[x != 0]` is now typically faster than `x[which(x != 0)]` (in the case where x has no NAs, the two are equivalent).

- `read.table()` now always uses the names for a named `colClasses` argument (previously names were only used when `colClasses` was too short). (In part, wish of PR#16478.)

- (Windows only) `download.file()` with default `method = "auto"` and a 'ftps://' URL chooses `"libcurl"` if that is available.

- The out-of-the box Bioconductor mirror has been changed to one using 'https://': use `chooseBioCmirror()` to choose a 'http://' mirror if required.

- The data frame and formula methods for `aggregate()` gain a `drop` argument.

- `available.packages()` gains a `repos` argument.

- The undocumented switching of methods for `url()` on 'https:' and 'ftps:' URLs is confined to `method = "default"` (and documented).

- `smoothScatter()` gains a `ret.selection` argument.

- `qr()` no longer has a `...` argument to pass additional arguments to methods.

- `[` has a method for class `"table"`.

- It is now possible (again) to `replayPlot()` a display list snapshot that was created by `recordPlot()` in a different R session.

  It is still not a good idea to use snapshots as a persistent storage format for R plots, but it is now not completely silly to use a snapshot as a format for transferring an R plot between two R sessions.

  The underlying changes mean that packages providing graphics devices (e.g., **Cairo**, **RSvgDevice**, **cairoDevice**, **tikzDevice**) will need to be reinstalled.

  Code for restoring snapshots was contributed by Jeroen Ooms and JJ Allaire.

  Some testing code is available at https://github.com/pmur002/R-display-list.

- `tools::undoc(dir = D)` and `codoc(dir = D)` now also work when D is a directory whose `normalizePath()`ed version does not end in the package name, e.g. from a symlink.

- `abbreviate()` has more support for multi-byte character sets – it no longer removes bytes within characters and knows about Latin vowels with accents. It is still only really suitable for (most) European languages, and still warns on non-ASCII input.

  `abbreviate(use.classes = FALSE)` is now implemented, and that is more suitable for non-European languages.

- `match(x,table)` is faster (sometimes by an order of magnitude) when x is of length one and `incomparables` is unchanged, thanks to Peter Haverty (PR#16491).

- More consistent, partly not back-compatible behavior of NA and NaN coercion to complex numbers, operations less often resulting in complex NA (`NA_complex_`).

- `lengths()` considers methods for `length` and `[[` on `x`, so it should work automatically on any objects for which appropriate methods on those generics are defined.

- The logic for selecting the default screen device on OS X has been simplified: it is now `quartz()` if that is available even if environment variable `DISPLAY` has been set by the user.

  The choice can easily be overridden *via* environment variable `R_INTERACTIVE_DEVICE`.

- On Unix-like platforms which support the `getline` C library function, `system(*,intern = TRUE)` no longer truncates (output) lines longer than 8192 characters, thanks to Karl Millar. (PR#16544)

- `rank()` gains a `ties.method = "last"` option, for convenience (and symmetry).

- `regmatches(invert = NA)` can now be used to extract both non-matched and matched substrings.

- `data.frame()` gains argument `fix.empty.names`; `as.data.frame.list()` gets new `cut.names`, `col.names` and `fix.empty.names`.

- `plot(x ~ x,*)` now warns that it is the same as `plot(x ~ 1,*)`.

- `recordPlot()` has new arguments `load` and `attach` to allow package names to be stored as part of a recorded plot. `replayPlot()` has new argument `reloadPkgs` to load/attach any package names that were stored as part of a recorded plot.

- S4 dispatch works within calls to `.Internal()`. This means explicit S4 generics are no longer needed for `unlist()` and `as.vector()`.

- Only font family names starting with '`"Hershey"`' (and not '`"Her"`' as before) are given special treatment by the graphics engine.

- S4 values are automatically coerced to vector (via `as.vector`) when subassigned into atomic vectors.

- `findInterval()` gets a `left.open` option.

- The version of LAPACK included in the sources has been updated to 3.6.0, including those 'deprecated' routines which were previously included. *Ca* 40 double-complex routines have been added at the request of a package maintainer.

  As before, the details of what is included are in 'src/modules/lapack/README' and this now gives information on earlier additions.

- `tapply()` has been made considerably more efficient without changing functionality, thanks to proposals from Peter Haverty and Suharto Anggono. (PR#16640)

- `match.arg(arg)` (the one-argument case) is faster; so is `sort.int()`. (PR#16640)

- The `format` method for `object_size` objects now also accepts "binary" units such as "KiB" and e.g., "Tb". (Partly from PR#16649.)

- Profiling now records calls of the form `foo::bar` and some similar cases directly rather than as calls to `<Anonymous>`. Contributed by Winston Chang.

- New string utilities `startsWith(x,prefix)` and `endsWith(x,suffix)`. Also provide speedups for some `grepl("^...",*)` uses (related to proposals in PR#16490).

- Reference class finalizers run at exit, as well as on garbage collection.

- Avoid **parallel** dependency on **stats** for port choice and random number seeds. (PR#16668)

- The radix sort algorithm and implementation from **data.table** (`forder`) replaces the previous radix (counting) sort and adds a new method for `order()`. Contributed by Matt Dowle and Arun Srinivasan, the new algorithm supports logical, integer (even with large values), real, and character vectors. It outperforms all other methods, but there are some caveats (see `?sort`).

- The `order()` function gains a `method` argument for choosing between `"shell"` and `"radix"`.

- New function `grouping()` returns a permutation that stably rearranges data so that identical values are adjacent. The return value includes extra partitioning information on the groups. The implementation came included with the new radix sort.

- `rhyper(nn,m,n,k)` no longer returns `NA` when one of the three parameters exceeds the maximal integer.

- `switch()` now warns when no alternatives are provided.

- `parallel::detectCores()` now has default `logical = TRUE` on all platforms – as this was the default on Windows, this change only affects Sparc Solaris.

  Option `logical = FALSE` is now supported on Linux and recent versions of OS X (for the latter, thanks to a suggestion of Kyaw Sint).

- `hist()` for `"Date"` or `"POSIXt"` objects would sometimes give misleading labels on the breaks, as they were set to the day before the start of the period being displayed. The display format has been changed, and the shift of the start day has been made conditional on `right = TRUE` (the default). (PR#16679)

- R now uses a new version of the logo (donated to the R Foundation by RStudio). It is defined in '.svg' format, so will resize without unnecessary degradation when displayed on HTML pages—there is also a vector PDF version. Thanks to Dirk Eddelbuettel for producing the corresponding X11 icon.

- New function `.traceback()` returns the stack trace which `traceback()` prints.

- `lengths()` dispatches internally.

- `dotchart()` gains a `pt.cex` argument to control the size of points separately from the size of plot labels. Thanks to Michael Friendly and Milan Bouchet-Valat for ideas and patches.

- `as.roman(ch)` now correctly deals with more diverse character vectors ch; also arithmetic with the resulting roman numbers works in more cases. (PR#16779)

- `prcomp()` gains a new option `rank.` allowing to directly aim for less than `min(n,p)` PC's. The `summary()` and its `print()` method have been amended, notably for this case.

- `gzcon()` gains a new option `text`, which marks the connection as text-oriented (so e.g. `pushBack()` works). It is still always opened in binary mode.

- The `import()` namespace directive now accepts an argument except which names symbols to exclude from the imports. The except expression should evaluate to a character vector (after substituting symbols for strings). See Writing R Extensions.

- New convenience function `Rcmd()` in package **tools** for invoking `R CMD` tools from within R.

- New functions `makevars_user()` and `makevars_site()` in package **tools** to determine the location of the user and site specific 'Makevars' files for customizing package compilation.

UTILITIES

- `R CMD check` has a new option '`--ignore-vignettes`' for use with non-Sweave vignettes whose '`VignetteBuilder`' package is not available.

- `R CMD check` now by default checks code usage (*via* **codetools**) with only the base package attached. Functions from default packages other than **base** which are used in the package code but not imported are reported as undefined globals, with a suggested addition to the NAMESPACE file.

- `R CMD check --as-cran` now also checks DOIs in package 'CITATION' and Rd files.

- `R CMD Rdconv` and `R CMD Rd2pdf` each have a new option '`--RdMacros=pkglist`' which allows Rd macros to be specified before processing.

DEPRECATED AND DEFUNCT

- The previously included versions of `zlib`, `bzip2`, `xz` and PCRE have been removed, so suitable external (usually system) versions are required (see the 'R Installation and Administration' manual).

- The unexported and undocumented Windows-only devices `cairo_bmp()`, `cairo_png()` and `cairo_tiff()` have been removed. (These devices should be used as e.g. `bmp(type = "cairo")`.)

- (Windows only) Function `setInternet2()` has no effect and will be removed in due course. The choice between methods `"internal"` and `"wininet"` is now made by the method arguments of `url()` and `download.file()` and their defaults can be set *via* options. The out-of-the-box default remains `"wininet"` (as it has been since R 3.2.2).

- `[<-` with an S4 value into a list currently embeds the S4 object into its own list such that the end result is roughly equivalent to using `[[<-`. That behavior is deprecated. In the future, the S4 value will be coerced to a list with `as.list()`.

- Package **tools**' functions `package.dependencies()`, `pkgDepends()`, etc are deprecated now, mostly in favor of `package_dependencies()` which is both more flexible and efficient.

INSTALLATION and INCLUDED SOFTWARE

- Support for very old versions of `valgrind` (e.g., 3.3.0) has been removed.

- The included `libtool` script (generated by `configure`) has been updated to version 2.4.6 (from 2.2.6a).

- `libcurl` version 7.28.0 or later with support for the `https` protocol is required for installation (except on Windows).

- BSD networking is now required (except on Windows) and so `capabilities("http/ftp")` is always true.

- `configure` uses `pkg-config` for PNG, TIFF and JPEG where this is available. This should work better with multiple installs and with those using static libraries.

- The minimum supported version of OS X is 10.6 ('Snow Leopard'): even that has been unsupported by Apple since 2012.

- The `configure` default on OS X is '`--disable-R-framework`': enable this if you intend to install under '/Library/Frameworks' and use with `R.app`.

- The minimum preferred version of PCRE has since R 3.0.0 been 8.32 (released in Nov 2012). Versions 8.10 to 8.31 are now deprecated (with warnings from `configure`), but will still be accepted until R 3.4.0.

- `configure` looks for C functions `__cospi`, `__sinpi` and `__tanpi` and uses these if `cospi` *etc* are not found. (OS X is the main instance.)

- (Windows) R is now built using `gcc` 4.9.3. This build will require recompilation of at least those packages that include C++ code, and possibly others. A build of R-devel using the older toolchain will be temporarily available for comparison purposes.

  During the transition, the environment variable `R_COMPILED_BY` has been defined to indicate which toolchain was used to compile R (and hence, which should be used to compile code in packages). The `COMPILED_BY` variable described below will be a permanent replacement for this.

- (Windows) A `make` and `R CMD config` variable named `COMPILED_BY` has been added. This indicates which toolchain was used to compile R (and hence, which should be used to compile code in packages).

PACKAGE INSTALLATION

- The `make` macro `AWK` which used to be made available to files such as 'src/Makefile' is no longer set.

C-LEVEL FACILITIES

- The API call `logspace_sum` introduced in R 3.2.0 is now remapped as an entry point to `Rf_logspace_sum`, and its first argument has gained a `const` qualifier. (PR#16470)

  Code using it will need to be reinstalled.

  Similarly, entry point `log1pexp` also defined in 'Rmath.h' is remapped there to `Rf_log1pexp`

- `R_GE_version` has been increased to 11.

- New API call `R_orderVector1`, a faster one-argument version of `R_orderVector`.

- When R headers such as 'R.h' and 'Rmath.h' are called from C++ code in packages they include the C++ versions of system headers such as '<cmath>' rather than the legacy headers such as '<math.h>'. (Headers 'Rinternals.h' and 'Rinterface.h' already did, and inclusion of system headers can still be circumvented by defining `NO_C_HEADERS`, including as from this version for those two headers.)

  The manual has long said that R headers should **not** be included within an `extern "C"` block, and almost all the packages affected by this change were doing so.

- Including header 'S.h' from C++ code would fail on some platforms, and so gives a compilation error on all.

- The deprecated header 'Rdefines.h' is now compatible with defining `R_NO_REMAP`.

- The connections API now includes a function `R_GetConnection()` which allows packages implementing connections to convert R connection objects to `Rconnection` handles used in the API. Code which previously used the low-level R-internal `getConnection()` entry point should switch to the official API.

BUG FIXES

- C-level asChar(x) is fixed for when x is not a vector, and it returns "TRUE"/"FALSE" instead of "T"/"F" for logical vectors.

- The first arguments of .colSums() etc (with an initial dot) are now named x rather than X (matching colSums()): thus error messages are corrected.

- A coef() method for class "maov" has been added to allow vcov() to work with multivariate results. (PR#16380)

- method = "libcurl" connections signal errors rather than retrieving HTTP error pages (where the ISP reports the error).

- xpdrows.data.frame() was not checking for unique row names; in particular, this affected assignment to non-existing rows via numerical indexing. (PR#16570)

- tail.matrix() did not work for zero rows matrices, and could produce row "labels" such as "[1e+05,]".

- Data frames with a column named "stringsAsFactors" now format and print correctly. (PR#16580)

- cor() is now guaranteed to return a value with absolute value less than or equal to 1. (PR#16638)

- Array subsetting now keeps names(dim(.)).

- Blocking socket connection selection recovers more gracefully on signal interrupts.

- The data.frame method of rbind() construction row.names works better in borderline integer cases, but may change the names assigned. (PR#16666)

- (X11 only) getGraphicsEvent() miscoded buttons and missed mouse motion events. (PR#16700)

- methods(round) now also lists round.POSIXt.

- tar() now works with the default files = NULL. (PR#16716)

- Jumps to outer contexts, for example in error recovery, now make intermediate jumps to contexts where on.exit() actions are established instead of trying to run all on.exit() actions before jumping to the final target. This unwinds the stack gradually, releases resources held on the stack, and significantly reduces the chance of a segfault when running out of C stack space. Error handlers established using withCallingHandlers() and options("error") specifications are ignored when handling a C stack overflow error as attempting one of these would trigger a cascade of C stack overflow errors. (These changes resolve PR#16753.)

- The spacing could be wrong when printing a complex array. (Report and patch by Lukas Stadler.)

- pretty(d,n,min.n,*) for date-time objects d works again in border cases with large min.n, returns a labels attribute also for small-range dates and in such cases its returned length is closer to the desired n. (PR#16761) Additionally, it finally does cover the range of d, as it always claimed.

- tsp(x) <-NULL did not handle correctly objects inheriting from both "ts" and "mts". (PR#16769)

- install.packages() could give false errors when options("pkgType") was "binary". (Reported by Jose Claudio Faria.)

- A bug fix in R 3.0.2 fixed problems with `locator()` in X11, but introduced problems in Windows. Now both should be fixed. (PR#15700)

- `download.file()` with `method = "wininet"` incorrectly warned of download file length difference when reported length was unknown. (PR#16805)

- `diag(NULL,1)` crashed because of missed type checking. (PR#16853)