

NEWS for R version 2.12.0 (2010-10-15)

NEWS

R News

CHANGES IN R VERSION 2.12.0

NEW FEATURES:

- Reading a packages's 'CITATION' file now defaults to ASCII rather than Latin-1: a package with a non-ASCII 'CITATION' file should declare an encoding in its 'DESCRIPTION' file and use that encoding for the 'CITATION' file.
- `difftime()` now defaults to the "tzzone" attribute of "POSIXlt" objects rather than to the current timezone as set by the default for the `tz` argument. (Wish of [PR#14182](#).)
- `pretty()` is now generic, with new methods for "Date" and "POSIXt" classes (based on code contributed by Felix Andrews).
- `unique()` and `match()` are now faster on character vectors where all elements are in the global CHARSXP cache and have unmarked encoding (ASCII). Thanks to Matthew Dowle for suggesting improvements to the way the hash code is generated in 'unique.c'.
- The `enquote()` utility, in use internally, is exported now.
- `.C()` and `.Fortran()` now map non-zero return values (other than `NA_LOGICAL`) for logical vectors to `TRUE`: it has been an implicit assumption that they are treated as true.
- The `print()` methods for "glm" and "lm" objects now insert linebreaks in long calls in the same way that the `print()` methods for "summary.[g]lm" objects have long done. This does change the layout of the examples for a number of packages, e.g. MASS. ([PR#14250](#))
- `constrOptim()` can now be used with method "SANN". ([PR#14245](#))
It gains an argument `hessian` to be passed to `optim()`, which allows all the ... arguments to be intended for `f()` and `grad()`. ([PR#14071](#))
- `curve()` now allows `expr` to be an object of mode "expression" as well as "call" and "function".
- The "POSIX[cl]t" methods for `Axis()` have been replaced by a single method for "POSIXt".
There are no longer separate `plot()` methods for "POSIX[cl]t" and "Date": the default method has been able to handle those classes for a long time. This *inter alia* allows a single date-time object to be supplied, the wish of [PR#14016](#).
The methods had a different default ("") for `xlab`.

- Classes "POSIXct", "POSIXlt" and "difftime" have generators .POSIXct(), .POSIXlt() and .difftime(). Package authors are advised to make use of them (they are available from R 2.11.0) to proof against planned future changes to the classes.
The ordering of the classes has been changed, so "POSIXt" is now the second class. See the document 'Updating packages for changes in R 2.12.x' on <http://developer.r-project.org> for the consequences for a handful of CRAN packages.
- The "POSIXct" method of as.Date() allows a timezone to be specified (but still defaults to UTC).
- New list2env() utility function as an inverse of as.list(<environment>) and for fast multi-assign() to existing environment. as.environment() is now generic and uses list2env() as list method.
- There are several small changes to output which 'zap' small numbers, e.g. in printing quantiles of residuals in summaries from "lm" and "glm" fits, and in test statistics in print.anova().
- Special names such as "dim", "names", etc, are now allowed as slot names of S4 classes, with "class" the only remaining exception.
- File '.Renviron' can have architecture-specific versions such as '.Renviron.i386' on systems with sub-architectures.
- installed.packages() has a new argument subarch to filter on sub-architecture.
- The summary() method for packageStatus() now has a separate print() method.
- The default summary() method returns an object inheriting from class "summaryDefault" which has a separate print() method that calls zapsmall() for numeric/complex values.
- The startup message now includes the platform and if used, sub-architecture: this is useful where different (sub-)architectures run on the same OS.
- The getGraphicsEvent() mechanism now allows multiple windows to return graphics events, through the new functions setGraphicsEventHandlers(), setGraphicsEventEnv(), and getGraphicsEventEnv(). (Currently implemented in the windows() and X11() devices.)
- tools::texi2dvi() gains an index argument, mainly for use by R CMD Rd2pdf. It avoids the use of texindy by texinfo's texi2dvi >= 1.157, since that does not emulate 'makeindex' well enough to avoid problems with special characters (such as (, {, !) in indices.
- The ability of readLines() and scan() to re-encode inputs to marked UTF-8 strings on Windows since R 2.7.0 is extended to non-UTF-8 locales on other OSes.
- scan() gains a fileEncoding argument to match read.table().
- points() and lines() gain "table" methods to match plot(). (Wish of [PR#10472](#).)
- Sys.chmod() allows argument mode to be a vector, recycled along paths.
- There are |, & and xor() methods for classes "octmode" and "hexmode", which work bitwise.
- Environment variables R_DVIPSCMD, R_LATEXCMD, R_MAKEINDEXCMD, R_PDFLATEXCMD are no longer used nor set in an R session. (With the move to tools::texi2dvi(), the conventional environment variables LATEX, MAKEINDEX and PDFLATEX will be used. options("dvipscmd") defaults to the value of DVIPS, then to "dvips".)
- New function isatty() to see if terminal connections are redirected.
- summaryRprof() returns the sampling interval in component sample.interval and only returns in by.self data for functions with non-zero self times.
- print(x) and str(x) now indicate if an empty list x is named.

- `install.packages()` and `remove.packages()` with `lib` unspecified and multiple libraries in `.libPaths()` inform the user of the library location used with a message rather than a warning.
- There is limited support for multiple compressed streams on a file: all of `[bgx]zfile()` allow streams to be appended to an existing file, but `bzfile()` reads only the first stream.
- Function `person()` in package **utils** now uses a given/family scheme in preference to first/middle/last, is vectorized to handle an arbitrary number of persons, and gains a `role` argument to specify person roles using a controlled vocabulary (the MARC relator terms).
- Package **utils** adds a new "bibentry" class for representing and manipulating bibliographic information in enhanced BibTeX style, unifying and enhancing the previously existing mechanisms.
- A `bibstyle()` function has been added to the **tools** package with default JSS style for rendering "bibentry" objects, and a mechanism for registering other rendering styles.
- Several aspects of the display of text help are now customizable using the new `Rd2txt_options()` function. `options("help_text_width")` is no longer used.
- Added `\href` tag to the Rd format, to allow hyperlinks to URLs without displaying the full URL.
- Added `\newcommand` and `\renewcommand` tags to the Rd format, to allow user-defined macros.
- New `toRd()` generic in the **tools** package to convert objects to fragments of Rd code, and added "fragment" argument to `Rd2txt()`, `Rd2HTML()`, and `Rd2latex()` to support it.
- Directory `'R_HOME/share/texmf'` now follows the TDS conventions, so can be set as a `texmf` tree ('root directory' in MiKTeX parlance).
- S3 generic functions now use correct S4 inheritance when dispatching on an S4 object. See `?Methods`, section on "Methods for S3 Generic Functions" for recommendations and details.
- `format.pval()` gains a `...` argument to pass arguments such as `nsmall` to `format()`. (Wish of [PR#9574](#))
- `legend()` supports `title.adj`. (Wish of [PR#13415](#))
- Added support for subsetting "raster" objects, plus assigning to a subset, conversion to a matrix (of colour strings), and comparisons (`==` and `!=`).
- Added a new `parseLatex()` function (and related functions `deparseLatex()` and `latexToUtf8()`) to support conversion of bibliographic entries for display in R.
- Text rendering of `\itemize` in help uses a Unicode bullet in UTF-8 and most single-byte Windows locales.
- Added support for polygons with holes to the graphics engine. This is implemented for the `pdf()`, `postscript()`, `x11(type="cairo")`, `windows()`, and `quartz()` devices (and associated raster formats), but not for `x11(type="Xlib")` or `xfig()` or `pictex()`. The user-level interface is the `polypath()` function in **graphics** and `grid.path()` in **grid**.
- File `'NEWS'` is now generated at installation with a slightly different format: it will be in UTF-8 on platforms using UTF-8, and otherwise in ASCII. There is also a PDF version, `'NEWS.pdf'`, installed at the top-level of the R distribution.
- `kmeans(x, 1)` now works. Further, `kmeans` now returns between and total sum of squares.

- `arrayInd()` and `which()` gain an argument `useNames`. For `arrayInd`, the default is now `false`, for speed reasons.
- As is done for closures, the default `print` method for the formula class now displays the associated environment if it is not the global environment.
- A new facility has been added for inserting code into a package without re-installing it, to facilitate testing changes which can be selectively added and backed out. See `?insertSource`.
- New function `readRenviron` to (re-)read files in the format of `'~/.Renviron'` and `'Renviron.site'`.
- `require()` will now return `FALSE` (and not fail) if loading the package or one of its dependencies fails.
- `aperm()` now allows argument `perm` to be a character vector when the array has named dimnames (as the results of `table()` calls do). Similarly, `array()` allows `MARGIN` to be a character vector. (Based on suggestions of Michael Lachmann.)
- Package `utils` now exports and documents functions `aspell_package_Rd_files()` and `aspell_package_vignettes()` for spell checking package Rd files and vignettes using Aspell, Ispell or Hunspell.
- Package `news` can now be given in Rd format, and `news()` prefers these `'inst/NEWS.Rd'` files to old-style plain text `'NEWS'` or `'inst/NEWS'` files.
- New simple function `packageVersion()`.
- The PCRE library has been updated to version 8.10.
- The standard Unix-alike terminal interface declares its name to readline as `'R'`, so that can be used for conditional sections in `'~/.inputrc'` files.
- 'Writing R Extensions' now stresses that the standard sections in `'Rd'` files (other than `'\alias'`, `'\keyword'` and `'\note'`) are intended to be unique, and the conversion tools now drop duplicates with a warning.
The `'Rd'` conversion tools also warn about an unrecognized type in a `'\docType'` section.
- `ecdf()` objects now have a `quantile()` method.
- `format()` methods for date-time objects now attempt to make use of a `"tzone"` attribute with `"%Z"` and `"%z"` formats, but it is not always possible. (Wish of [PR#14358](#).)
- `tools::texi2dvi(file, clean = TRUE)` now works in more cases (e.g. where emulation is used and when `'file'` is not in the current directory).
- New function `droplevels()` to remove unused factor levels.
- `system(command, intern = TRUE)` now gives an error on a Unix-alike (as well as on Windows) if `command` cannot be run. It reports a non-success exit status from running `command` as a warning.
On a Unix-alike an attempt is made to return the actual exit status of the command in `system(intern = FALSE)`: previously this had been system-dependent but on POSIX-compliant systems the value return was 256 times the status.
- `system()` has a new argument `ignore.stdout` which can be used to (portably) ignore standard output.
- `system(intern = TRUE)` and `pipe()` connections are guaranteed to be available on all builds of R.
- `Sys.which()` has been altered to return `""` if the command is not found (even on Solaris).
- A facility for defining reference-based S4 classes (in the OOP style of Java, C++, etc.) has been added experimentally to package `methods`; see `?ReferenceClasses`.

- The `predict` method for "loess" fits gains an `na.action` argument which defaults to `na.pass` rather than the previous default of `na.omit`. Predictions from "loess" fits are now named from the row names of `newdata`.
- Parsing errors detected during `Sweave()` processing will now be reported referencing their original location in the source file.
- New `adjustcolor()` utility, e.g., for simple translucent color schemes.
- `qr()` now has a trivial `lm` method with a simple (fast) validity check.
- An experimental new programming model has been added to package `methods` for reference (OOP-style) classes and methods. See `?ReferenceClasses`.
- `bzip2` has been updated to version 1.0.6 (bug-fix release). `'--with-system-bzlib'` now requires at least version 1.0.6.
- R now provides `'jss.cls'` and `'jss.bst'` (the class and bib style file for the Journal of Statistical Software) as well as `'RJournal.bib'` and `'Rnews.bib'`, and R CMD ensures that the `'bst'` and `'bib'` files are found by BibTeX.
- Functions using the `TAR` environment variable no longer quote the value when making `system` calls. This allows values such as `'tar --force-local'`, but does require additional quotes in, e.g., `TAR = "'/path with spaces/mytar'"`.

DEPRECATED & DEFUNCT:

- Supplying the parser with a character string containing both octal/hex and Unicode escapes is now an error.
- File extension `'C'` for C++ code files in packages is now defunct.
- R CMD `check` no longer supports configuration files containing Perl configuration variables: use the environment variables documented in 'R Internals' instead.
- The `save` argument of `require()` now defaults to `FALSE` and `save = TRUE` is now deprecated. (This facility is very rarely actually used, and was superseded by the 'Depends' field of the 'DESCRIPTION' file long ago.)
- R CMD `check --no-latex` is deprecated in favour of `'--no-manual'`.
- R CMD `Sd2Rd` is formally deprecated and will be removed in R 2.13.0.

PACKAGE INSTALLATION:

- `install.packages()` has a new argument `libs_only` to optionally pass `'--libs-only'` to R CMD `INSTALL` and works analogously for Windows binary installs (to add support for 64- or 32-bit Windows).
- When sub-architectures are in use, the installed architectures are recorded in the `Archs` field of the 'DESCRIPTION' file. There is a new default filter, `"subarch"`, in `available.packages()` to make use of this. Code is compiled in a copy of the `'src'` directory when a package is installed for more than one sub-architecture: this avoid problems with cleaning the sources between building sub-architectures.
- R CMD `INSTALL --libs-only` no longer overrides the setting of locking, so a previous version of the package will be restored unless `'--no-lock'` is specified.

UTILITIES:

- R CMD `Rprof|build|check` are now based on R rather than Perl scripts. The only remaining Perl scripts are the deprecated R CMD `Sd2Rd` and `install-info.pl` (used only if `install-info` is not found) as well as some maintainer-mode-only scripts. **NB:** because these have been completely rewritten, users should not expect undocumented details of previous implementations to have been duplicated. R CMD no longer manipulates the environment variables `PERL5LIB` and `PERLLIB`.

- R CMD **check** has a new argument '`--extra-arch`' to confine tests to those needed to check an additional sub-architecture.
Its check for "Subdirectory 'inst' contains no files" is more thorough: it looks for files, and warns if there are only empty directories.
Environment variables such as `R_LIBS` and those used for customization can be set for the duration of checking *via* a file '`~/R/check.Renviron`' (in the format used by '`.Renviron`', and with sub-architecture specific versions such as '`~/R/check.Renviron.i386`' taking precedence).
There are new options '`--multiarch`' to check the package under all of the installed sub-architectures and '`--no-multiarch`' to confine checking to the sub-architecture under which **check** is invoked. If neither option is supplied, a test is done of installed sub-architectures and all those which can be run on the current OS are used.
Unless multiple sub-architectures are selected, the install done by **check** for testing purposes is only of the current sub-architecture (*via* R CMD `INSTALL --no-multiarch`).
It will skip the check for non-ascii characters in code or data if the environment variables `_R_CHECK_ASCII_CODE_` or `_R_CHECK_ASCII_DATA_` are respectively set to `FALSE`. (Suggestion of Vince Carey.)
- R CMD **build** no longer creates an 'INDEX' file (R CMD `INSTALL` does so), and `--force` removes (rather than overwrites) an existing 'INDEX' file.
It supports a file '`~/R/build.Renviron`' analogously to **check**.
It now runs build-time `\Sexpr` expressions in help files.
- R CMD **Rd2dvi** makes use of `tools::texi2dvi()` to process the package manual. It is now implemented entirely in R (rather than partially as a shell script).
- R CMD **Rprof** now uses `utils::summaryRprof()` rather than Perl. It has new arguments to select one of the tables and to limit the number of entries printed.
- R CMD **Sweave** now runs R with `--vanilla` so the environment setting of `R_LIBS` will always be used.

C-LEVEL FACILITIES:

- `lang5()` and `lang6()` (in addition to pre-existing `lang[1-4]()`) convenience functions for easier construction of `eval()` calls. If you have your own definition, do wrap it inside `#ifndef lang5 #endif` to keep it working with old and new R.
- Header '`R.h`' now includes only the C headers it itself needs, hence no longer includes `errno.h`. (This helps avoid problems when it is included from C++ source files.)
- Headers '`Rinternals.h`' and '`R_ext/Print.h`' include the C++ versions of '`stdio.h`' and '`stdarg.h`' respectively if included from a C++ source file.

INSTALLATION:

- A C99 compiler is now required, and more C99 language features will be used in the R sources.
- Tcl/Tk ≥ 8.4 is now required (increased from 8.3).
- System functions `access`, `chdir` and `getcwd` are now essential to configure R. (In practice they have been required for some time.)
- **make check** compares the output of the examples from several of the base packages to reference output rather than the previous output (if any). Expect some differences due to differences in floating-point computations between platforms.
- File '`NEWS`' is no longer in the sources, but generated as part of the installation. The primary source for changes is now '`doc/NEWS.Rd`'.
- The `popen` system call is now required to build R. This ensures the availability of `system(intern = TRUE)`, `pipe()` connections and printing from `postscript()`.

- The `pkg-config` file `'libR.pc'` now also works when R is installed using a sub-architecture.
- R has always required a BLAS that conforms to IE60559 arithmetic, but after discovery of more real-world problems caused by a BLAS that did not, this is tested more thoroughly in this version.

BUG FIXES:

- Calls to `selectMethod()` by default no longer cache inherited methods. This could previously corrupt methods used by `as()`.
- The densities of non-central chi-squared are now more accurate in some cases in the extreme tails, e.g. `dchisq(2000, 2, 1000)`, as a series expansion was truncated too early. (PR#14105)
- `pt()` is more accurate in the left tail for `ncp` large, e.g. `pt(-1000, 3, 200)`. (PR#14069)
- The default C function (`R_binary`) for binary ops now sets the S4 bit in the result if either argument is an S4 object. (PR#13209)
- `source(echo=TRUE)` failed to echo comments that followed the last statement in a file.
- S4 classes that contained one of `"matrix"`, `"array"` or `"ts"` and also another class now accept superclass objects in `new()`. Also fixes failure to call `validObject()` for these classes.
- Conditional inheritance defined by argument `test` in `methods::setIs()` will no longer be used in S4 method selection (caching these methods could give incorrect results). See `?setIs`.
- The signature of an implicit generic is now used by `setGeneric()` when that does not use a definition nor explicitly set a signature.
- A bug in `callNextMethod()` for some examples with `"..."` in the arguments has been fixed. See file `'src/library/methods/tests/nextWithDots.R'` in the sources.
- `match(x, table)` (and hence `%in%`) now treat `"POSIXlt"` consistently with, e.g., `"POSIXct"`.
- Built-in code dealing with environments (`get()`, `assign()`, `parent.env()`, `is.environment()` and others) now behave consistently to recognize S4 subclasses; `is.name()` also recognizes subclasses.
- The `abs.tol` control parameter to `nlminb()` now defaults to 0.0 to avoid false declarations of convergence in objective functions that may go negative.
- The standard Unix-alike termination dialog to ask whether to save the workspace takes a EOF response as `n` to avoid problems with a damaged terminal connection. (PR#14332)
- Added `warn.unused` argument to `hist.default()` to allow suppression of spurious warnings about graphical parameters used with `plot=FALSE`. (PR#14341)
- `predict.lm()`, `summary.lm()`, and indeed `lm()` itself had issues with residual DF in zero-weighted cases (the latter two only in connection with empty models). (Thanks to Bill Dunlap for spotting the `predict()` case.)
- `aperm()` treated `resize = NA` as `resize = TRUE`.
- `constrOptim()` now has an improved convergence criterion, notably for cases where the minimum was (very close to) zero; further, other tweaks inspired from code proposals by Ravi Varadhan.
- Rendering of S3 and S4 methods in man pages has been corrected and made consistent across output formats.
- Simple markup is now allowed in `'\title'` sections in `' .Rd'` files.

- The behaviour of `as.logical()` on factors (to use the levels) was lost in R 2.6.0 and has been restored.
- `prompt()` did not backquote some default arguments in the ‘`\usage`’ section. (Reported by Claudia Beleites.)
- `writeBin()` disallows attempts to write 2GB or more in a single call. (PR#14362)
- `new()` and `getClass()` will now work if `Class` is a subclass of `"classRepresentation"` and should also be faster in typical calls.
- The `summary()` method for data frames makes a better job of names containing characters invalid in the current locale.
- `[[` sub-assignment for factors could create an invalid factor (reported by Bill Dunlap).
- `Negate(f)` would not evaluate argument `f` until first use of returned function (reported by Olaf Mersmann).
- `quietly=FALSE` is now also an optional argument of `library()`, and consequently, `quietly` is now propagated also for loading dependent packages, e.g., in `require(*, quietly=TRUE)`.
- If the loop variable in a `for` loop was deleted, it would be recreated as a global variable. (Reported by Radford Neal; the fix includes his optimizations as well.)
- Task callbacks could report the wrong expression when the task involved parsing new code. (PR#14368)
- `getNamespaceVersion()` failed; this was an accidental change in 2.11.0. (PR#14374)
- `identical()` returned `FALSE` for external pointer objects even when the pointer addresses were the same.
- `L$a@x[] <- val` did not duplicate in a case it should have.
- `tempfile()` now always gives a random file name (even if the directory is specified) when called directly after startup and before the R RNG had been used. (PR#14381)
- `quantile(type=6)` behaved inconsistently. (PR#14383)
- `backSpline(.)` behaved incorrectly when the knot sequence was decreasing. (PR#14386)
- The reference BLAS included in R was assuming that `0*x` and `x*0` were always zero (whereas they could be `NA` or `NaN` in IEC 60559 arithmetic). This was seen in results from `tcrossprod`, and for example that `log(0) %*% 0` gave 0.
- The calculation of whether text was completely outside the device region (in which case, you draw nothing) was wrong for screen devices (which have `[0, 0]` at top-left). The symptom was (long) text disappearing when resizing a screen window (to make it smaller). (PR#14391)
- `model.frame(drop.unused.levels = TRUE)` did not take into account `NA` values of factors when deciding to drop levels. (PR#14393)
- `library.dynam.unload` required an absolute path for `libpath`. (PR#14385)
Both `library()` and `loadNamespace()` now record absolute paths for use by `searchpaths()` and `getNamespaceInfo(ns, "path")`.
- The self-starting model `NLSstClosestX` failed if some deviation was exactly zero. (PR#14384)
- `X11(type = "cairo")` (and other devices such as `png` using `cairographics`) and which use Pango font selection now work around a bug in Pango when very small fonts (those with sizes between 0 and 1 in Pango’s internal units) are requested. (PR#14369)
- Added workaround for the font problem with `X11(type = "cairo")` and similar on Mac OS X whereby italic and bold styles were interchanged. (PR#13463 amongst many other reports.)
- `source(chdir = TRUE)` failed to reset the working directory if it could not be determined – that is now an error.

- Fix for crash of `example(rasterImage)` on `x11(type="Xlib")`.
- Force Quartz to bring the on-screen display up-to-date immediately before the snapshot is taken by `grid.cap()` in the Cocoa implementation. (PR#14260)
- `model.frame` had an unstated 500 byte limit on variable names. (Example reported by Terry Therneau.)
- The 256-byte limit on names is now documented.
- Subassignment by `[`, `[[` or `$` on an expression object with value `NULL` coerced the object to a list.

CHANGES IN R VERSION 2.11.1 patched

NEW FEATURES:

- `install.packages()` has a new optional argument `INSTALL_opts` which can be used to pass options to R CMD `INSTALL` for source-package installs.
- R CMD `check` now runs the package-specific tests with `LANGUAGE=en` to facilitate comparison to `‘.Rout.save’` files.
- `sessionInfo()` gives more detailed platform information, including 32/64-bit and the sub-architecture if one is used.

DEPRECATED & DEFUNCT:

- The use of Perl configuration variables for R CMD `check` (as previously documented in ‘Writing R Extensions’) is deprecated and will be removed in R 2.12.0. Use the environment variables documented in ‘R Internals’ instead.

BUG FIXES:

- R CMD `Rd2dvi` failed if run from a path containing space(s). This also affected R CMD `check`, which calls `Rd2dvi`.
- `stripchart()` could fail with an empty factor level. (PR#14317)
- Text help rendering of `\tabular{}` has been improved: under some circumstances leading blank columns were not rendered.
- `strsplit(x, fixed=TRUE)` marked UTF-8 strings with the local encoding when no splits were found.
- `weighted.mean(NA, na.rm=TRUE)` and similar now returns `NaN` again, as it did prior to R 2.10.0.
- R CMD had a typo in its detection of whether the environment variable `TEXINPUTS` was set (reported by Martin Morgan).
- The command-line parser could mistake `‘--file=size...’` for one of the options for setting limits for `Ncells` or `Vcells`.
- The internal `strptime()` could corrupt its copy of the timezone which would then lead to spurious warnings. (PR#14338)
- `dir.create(recursive = TRUE)` could fail if one of the components existed but was a directory on a read-only file system. (Seen on Solaris, where the error code returned is not even listed as possible on the man page.)
- The `postscript()` and `pdf()` devices will now allow `lwd` values less than 1 (they used to force such values to be 1).
- Fixed font face for CID fonts in `pdf()` graphics output. (PR#14326)
- `GERaster()` now checks for width or height of zero and does nothing in those cases; previously the behaviour was undefined, probably device-specific, and possibly dangerous.

- `wilcox.test(x, y, conf.int = TRUE)` failed with an unhelpful message if `x` and `y` were constant vectors, and similarly in the one-sample case. (PR#14329)
- Improperly calling `Recall()` from outside a function could cause a segfault. (Reported by Robert McGehee.)
- `\Sexpr[result=rd]` in an Rd file added a spurious newline, which was displayed as extra whitespace when rendered.
- `require(save = TRUE)` recorded the names of packages it failed to load.
- `packageStatus()` could return a data frame with duplicate row names which could then not be printed.
- `txtProgressBar(style = 2)` did not work correctly.
`txtProgressBar(style = 3)` did not display until a non-minimum value was set.
- `contour()` did not display dashed line types properly when contour lines were labelled. (Reported by David B. Thompson.)
- `tools::undoc()` again detects undocumented data objects. Of course, this also affects R CMD check.
- `ksmooth(x, NULL)` no longer segfaults.
- `approxfun()`, `approx()`, `splinefun()` and `spline()` could be confused by `x` values that were different but so close as to print identically. (PR#14377)

CHANGES IN R VERSION 2.11.1

NEW FEATURES:

- R CMD INSTALL checks if dependent packages are available early on in the installation of source packages, thereby giving clearer error messages.
- R CMD INSTALL --build now names the file in the format used for Mac OS X binary files on that platform.
- `BIC()` in package **stats4** now also works with multiple fitted models, analogously to `AIC()`.

DEPRECATED & DEFUNCT:

- Use of file extension `‘.C’` for C++ code in packages is now deprecated: it has caused problems for some `makes` on case-insensitive file systems (although it currently works with the recommended toolkits).

INSTALLATION:

- Command `gnutar` is preferred to `tar` when `configure` sets `TAR`. This is needed on Mac OS 10.6, where the default `tar`, `bsdtar 2.6.2`, has been reported to produce archives with illegal extensions to `tar` (according to the POSIX standard).

BUG FIXES:

- The C function `mkCharLenCE` now no longer reads past `len` bytes (unlikely to be a problem except in user code). (PR#14246)
- On systems without any default `LD_LIBRARY_PATH` (not even `‘/usr/local/lib’`), `[DY]LIB_LIBRARY_PATH` is now set without a trailing colon. (PR#13637)
- More efficient implementation of `utf8ToInt()` on long multi-byte strings with many multi-byte characters. (PR#14262)
- `aggregate.ts()` gave platform-dependent results due to rounding error for `ndeltat != 1`.
- `package.skeleton()` sometimes failed to fix filenames for `‘.R’` or `‘.Rd’` files to start with an alphanumeric. (PR#14253)
 It also failed when only an S4 class without any methods was defined. (PR#14280)

- `splinefun(method = "monoH.FC")` was not quite monotone in rare cases. (PR#14215)
- `Rhttpd` no longer crashes due to SIGPIPE when the client closes the connection prematurely. (PR#14266)
- `format.POSIXlt()` could cause a stack overflow and crash when used on very long vectors. (PR#14267)
- `Rd2latex()` incorrectly escaped special characters in `\usage` sections.
- `mcnemar.test()` could alter the levels (dropping unused levels) if passed `x` and `y` as factors (reported by Greg Snow).
- `Rd2pdf` sometimes needed a further `pdflatex` pass to get hyperlinked pages correct.
- `interaction()` produced malformed results when levels were duplicated, causing segfaults in `split()`.
- `cut(d, breaks = <n>)` now also works for "Date" or "POSIXt" argument `d`. (PR#14288)
- `memDecompress()` could decompress incompletely rare `xz`-compressed input due to incorrect documentation of `xz` utils. (Report and patch from Olaf Mersmann.)
- The `S4 initialize()` methods for "matrix", "array", and "ts" have been fixed to call `validObject()`. (PR#14284)
- R CMD INSTALL now behaves the same way with or without `'--no-multiarch'` on platforms with only one installed architecture. (It used to clean the `'src'` directory without `'--no-multiarch'`.)
- `[<-.data.frame` was not quite careful enough in assigning (and potentially deleting) columns right-to-left. (PR#14263)
- `rbeta(n, a, b)` no longer occasionally returns NaN for `a >> 1 > b`. (PR#14291)
- `pnorm(x, log.p = TRUE)` could return NaN not `-Inf` for `x` near (minus for `lower.tail=TRUE`) the largest representable number.
- Compressed data files `*.(txt|tab|csv).(gz|bz2|xz)` were not recognized for the list of data topics and hence for packages using `LazyData`. (PR#14273)
- `textConnection()` did an unnecessary translation on strings in a foreign encoding (e.g. UTF-8 strings on Windows) and so was slower than it could have been on very long input strings. (PR#14286)
- `tools::Rd2txt()` did not render poorly written Rd files consistently with other renderers.
It computed widths of strings that would be `print()`ed with escapes incorrectly, for example in the computation of column width for `\tabular`.
- `na.action()` did not extract the `na.action` component as documented.
- `do.call()`ing `NextMethod` in erroneous ways no longer segfaults. (PR#13487)

CHANGES IN R VERSION 2.11.0

SIGNIFICANT USER-VISIBLE CHANGES:

- Packages must have been installed under R \geq 2.10.0, as the current help system is the only one now supported.
- A port to 64-bit Windows is now available as well as binary package repositories: see the 'R Administration and Installation Manual'.
- Argument matching for primitive functions is now done in the same way as for interpreted functions except for the deliberate exceptions
`call switch .C .Fortran .Call .External`

all of which use positional matching for their first argument, and also some internal-use-only primitives.

- The default device for command-line R at the console on Mac OS X is now `quartz()` and not `X11()`.

NEW FEATURES:

- The `open` modes for connections are now interpreted more consistently. `open = "r"` is now equivalent to `open = "rt"` for all connections. The default `open = ""` now means `"rt"` for all connections except the compressed-file connections `gzfile()`, `bzfile()` and `xzfile()` for which it means `"rb"`.
- R CMD INSTALL now uses the internal `untar()` function in package `utils`: this ensures that all platforms can install `bzip2`- and `xz`-compressed tarballs. In case this causes problems (as it has on some Windows file systems when run from Cygwin tools) it can be overridden by the environment variable `R_INSTALL_TAR`: setting this to a modern external tar program will speed up unpacking of large (tens of Mb or more) tarballs.
- `help(try.all.packages = TRUE)` is much faster (although the time taken by the OS to find all the packages the first time it is used can dominate the time).
- R CMD check has a new option `--timings` to record per-example timings in file `'<pkg>.Rcheck/<pkg>-Ex.timings'`.
- The TRE library has been updated to version 0.8.0 (minor bugfixes).
- `grep[1]`, `[g]sub` and `[g]regexpr` now work in bytes in an 8-bit locales if there is no marked UTF-8 input string: this will be somewhat faster, and for `[g]sub()` give the result in the native encoding rather than in UTF-8 (which returns to the behaviour prior to R 2.10.0).
- A new argument `skipCalls` has been added to `browser()` so that it can report the original context when called by other debugging functions.
- More validity checking of UTF-8 and MBCS strings is done by `agrep()` and the regular-expression matching functions.
- The undocumented restriction on `gregexpr()` to `length(text) > 0` has been removed.
- Package `tcltk` now sends strings to Tcl in UTF-8: this means that strings with a marked UTF-8 encoding are supported in non-UTF-8 locales.
- The graphics engine now supports rendering of raster (bitmap) images, though not all graphics devices can provide (full) support. Packages providing graphics devices (e.g., `Cairo`, `RSvgDevice`, `cairoDevice`) will need to be reinstalled. There is also support in the graphics engine for capturing raster images from graphics devices (again not supported on all graphics devices).
- R CMD check now also checks if the package and namespace can be unloaded: this provides a check of the `.Last.lib()` and `.onUnload()` hook functions (unless `'--install=fake'`).
- `prop.table(x)` now accepts a one-dimensional table for `x`.
- A new function `vapply()` has been added, based on a suggestion from Bill Dunlap. It requires that a template for the function value be specified, and uses it to determine the output type and to check for consistency in the function values.
- The main HTML help page now links to a reformatted copy of this 'NEWS' file. (Suggested by Henrik Bengtsson.) Package index files link to the package 'DESCRIPTION' and 'NEWS' files and a list of demos when using dynamic help.
- The `[` method for class `"AsIs"` allows the next method to change the underlying class. (Wish of Jens Oehlschlägel.)

- `write.csv[2]` no longer allow argument `append` to be changed: as ever, direct calls to `write.table()` give more flexibility as well as more room for error.
- The index page for HTML help for a package now collapses multiple signatures for S4 methods into a single entry.
- The use of `.required` by `require()` and `detach()` has been replaced by `.Depends` which is set from the `Depends` field of a package (even in packages with name spaces). By default `detach()` prevents such dependencies from being detached: this can be overridden by the argument `force`.
- `bquote()` has been extended to work on function definitions. (Wish of [PR#14031](#)).
- `detach()` when applied to an object other than a package returns the environment that has been detached, to parallel `attach()`.
- `readline()` in non-interactive use returns "" and does not attempt to read from the 'terminal'.
- New function `file_ext()` in package **tools**.
- `xtfrm()` is now primitive and internally generic, as this allows S4 methods to be set on it without name-space scoping issues.
There are now "AsIs" and "difftime" methods, and the default method uses `unclass(x)` if `is.numeric(x)` is true (which will be faster but relies on `is.numeric()` having been set correctly for the class).
- `is.numeric(x)` is now false for a "difftime" object (multiplication and division make no sense for such objects).
- The default method of `weighted.mean(x, w)` coerces `w` to be numeric (aka double); previously only integer weights were coerced. Zero weights are handled specially so an infinite value with zero weight does not force an NaN result.
There is now a "difftime" method.
- `bug.report()` now has arguments `package` and `lib.loc` to generate bug reports about packages. When this is used, it looks for a `BugReports` field in the package 'DESCRIPTION' file, which will be assumed to be a URL at which to submit the report, and otherwise generates an email to the package maintainer. (Suggested by Barry Rowlingson.)
- `quantile()` now has a method for the date-time class "POSIXt", and types 1 and 3 (which never interpolate) work for Dates and ordered factors.
- `length(<POSIXlt>)` now returns the length of the corresponding abstract timelength-vector rather than always 9 (the length of the underlying list structure). (Wish of [PR#14073](#) and [PR#10507](#).)
- The `readline` completion backend no longer sorts possible completions alphabetically (e.g., function argument names) if R was built with `readline >= 6`.
- `select.list()` gains a `graphics` argument to allow Windows/Mac users to choose the text interface. This changes the behaviour of `new.packages(ask=TRUE)` to be like `update.packages(ask=TRUE)` on those platforms in using a text menu: use `ask="graphics"` for a graphical menu.
- New function `chooseBioCmirror()` to set the "BioC_mirror" option.
- The R grammar now prevents using the argument `name` in signatures of S4 methods for `$` and `$<-`, since they will always be called with a character string value for `name`. The implicit S4 generic functions have been changed to reflect this: packages which included `name` in the signature of their methods need to be updated and re-installed.
- The handling of the `method` argument of `glm()` has been refined following suggestions by Ioannis Kosmidis and Heather Turner.
- `str()` gains a new argument `list.len` with default 99, limiting the number of `list()` items (per level), thanks to suggestions from David Winsenius.

- Having formal arguments of an S4 method in a different order from the generic is now an error (the warning having been ignored by some package maintainers for a long time).
- New functions `enc2native()` and `enc2utf8()` convert character vectors with possibly marked encodings to the current locale and UTF-8 respectively.
- Unrecognized escapes and embedded nuls in character strings are now an error, not just a warning. Thus option `"warnEscapes"` is no longer needed. `rawToChar()` now removes trailing nuls silently, but other embedded nuls become errors.
- Informational messages about masked objects displayed when a package is attached are now more compact, using `strwrap()` instead of one object per line.
- `print.rle()` gains argument `prefix`.
- `download.file()` gains a `"curl"` method, mainly for use on platforms which have `curl` but not `wget`, but also for some hard-to-access URLs.
- In Rd, `\eqn` and `\deqn` will render in HTML (and convert to text) upper- and lower-case Greek letters (entered as `\alpha ...`), ..., `\ge` and `\le`.
- `utf8ToInt()` and `intToUtf8()` now map NA inputs to NA outputs.
- `file()` has a new argument `raw` which may help if it is used with something other than a regular file, e.g. a character device.
- New function `strtoi()`, a wrapper for the C function `strtol`.
- `as.octmode()` and `as.hexmode()` now allow inputs of length other than one. The `format()` and `print()` methods for `"octmode"` now preserve names and dimensions (as those for `"hexmode"` did). The `format()` methods for classes `"octmode"` and `"hexmode"` gain a `width` argument.
- `seq.int()` returns an integer result in some further cases where `seq()` does, e.g. `seq.int(1L, 9L, by = 2L)`.
- Added `\subsection{ }{ }` macro to Rd syntax, for subsections within sections.
- n-dimensional arrays with dimension names can now be indexed by an n-column character matrix. The indices are matched against the dimension names. NA indices are propagated to the result. Unmatched values and `" "` are not allowed and result in an error.
- `interaction(drop=TRUE)` uses less memory (related to [PR#14121](#)).
- `summary()` methods have been added to the `"srcref"` and `"srcfile"` classes, and various encoding issues have been cleaned up.
- If option `"checkPackageLicense"` is set to TRUE (not currently the default), users will be asked to agree to non-known-to-be-FOSS package licences at first use.
- Checking `setAs(a, b)` methods only gives a message instead of a warning, when one of `a` or `b` is unknown.
- New function `norm()` to compute a matrix norm. `norm()` and also `backsolve()` and `sample()` have implicit S4 generics.
- Files `'Renviron.site'` and `'Rprofile.site'` can have architecture-specific versions on systems with sub-architectures.
- R CMD `check` now (by default) also checks Rd files for auto-generated content in need of editing, and missing argument descriptions.
- `aggregate()` gains a formula method thanks to a contribution by Arni Magnusson. The data frame method now allows summary functions to return arbitrarily many values.
- `path.expand()` now propagates NA values rather than converting them to `"NA"`.
- `file.show()` now disallows NA values for file names, headers, and pager.

- The ‘fuzz’ used by `seq()` and `seq.int()` has been reduced from `1e-7` to `1e-10`, which should be ample for the double-precision calculations used in R. It ensures that the fuzz never comes into play with sequences of integers (wish of [PR#14169](#)).
- The default value of `RSiteSearch(restrict=)` has been changed to include vignettes but to exclude R-help. The R-help archives available have been split, with a new option of `"Rhelp10"` for those from 2010.
- New function `rasterImage()` in the `graphics` package for drawing raster images.
- `stats::extractAIC.coxph()` now omits aliased terms when computing the degrees of freedom (suggestion of Terry Therneau).
- `cor()` and `cov()` now test for misuse with non-numeric arguments, such as the non-bug report [PR#14207](#).
- `pchisq(ncp =, log.p = TRUE)` is more accurate for probabilities near one. E.g. `pchisq(80, 4, ncp=1, log.p=TRUE)`. (Maybe what was meant in [PR#14216](#).)
- `maintainer()` has been added, to give convenient access to the name of the maintainer of a package (contributed by David Scott).
- `sample()` and `sample.int()` allow zero items to be sampled from a zero-length input. `sample.int()` gains a default value `size=n` to be more similar to `sample()`.
- `switch()` returned `NULL` on error (not previously documented on the help page): it now does so invisibly, analogously to `if-without-else`.
It is now primitive: this means that argument `EXPR` is always matched to the first argument and there is no danger of partial matching to later named arguments.
- Primitive functions `UseMethod()`, `attr()`, `attr<-()`, `on.exit()`, `retracemem()` and `substitute()` now use standard argument matching (rather than positional matching). This means that all multi-argument primitives which are not internal now use standard argument matching except where positional matching is desirable (as for `switch()`, `call()`, `.C()` ...).
- All the one-argument primitives now check that any name supplied for their first argument is a partial match to the argument name as documented on the help page: this also applies to replacement functions of two arguments.
- `base::which()` uses a new `.Internal` function when `arr.ind` is `FALSE` resulting in a 10x speedup. Thanks to Patrick Aboyoun for implementation suggestions.
- Help conversion to text now uses the first part of `\enc{ }{ }` markup if it is representable in the current output encoding. On the other hand, conversion to LaTeX with the default `outputEncoding = "ASCII"` uses the second part.
- A new class `"listOfMethods"` has been introduced to represent the methods in a methods table, to replace the deprecated class `"MethodsList"`.
- `any()` and `all()` return early if possible. This may speed up operations on long vectors.
- `strptime()` now accepts `"%z"` (for the offset from UTC in the RFC822 format of `+/-hhmm`).
- The PCRE library has been updated to version 8.02, a bug-fix release which also updates tables to Unicode 5.02.
- Functions which may use a graphical `select.list()` (including `menu()` and `install.packages()`) now check on a Unix-alike that Tk can be started (and not just `capabilities("tcltk") && capabilities("X11")`).
- The parser no longer marks strings containing octal or hex escapes as being in UTF-8 when entered in a UTF-8 locale.
- On platforms with cairo but not Pango (notably Mac OS X) the initial default `X11()` type is set to `"Xlib"`: this avoids several problems with font selection when done by cairo rather than Pango (at least on Mac OS X).

- New function `arrayInd()` such that `which(x, arr.ind = TRUE)` for an array 'x' is now equivalent to `arrayInd(which(x), dim(x), dimnames(x))`.

DEPRECATED & DEFUNCT:

- Bundles of packages are defunct.
- `stats::clearNames()` is defunct: use `unname()`.
- Basic regular expressions are defunct, and `strsplit()`, `grep()`, `grepl()`, `sub()`, `gsub()`, `regexpr()` and `gregexpr()` no longer have an `extended` argument.
- `methods::trySilent()` is defunct.
- `index.search()` (which was deprecated in 2.10.0) is no longer exported and has a different argument list.
- Use of multiple arguments to `return()` is now defunct.
- The use of `UseMethod()` with more than two arguments is now defunct.
- In the **methods** package, the "MethodsList" metadata objects which had been superseded by hash tables (environments) since R 2.8.0 are being phased out. Objects of this class are no longer assigned or used as metadata by the package. `getMethods()` is now deprecated, with its internal use replaced by `findMethods()` and other changes. Creating objects from the "MethodsList" class is also deprecated.
- Parsing strings containing both octal/hex and Unicode escapes now gives a warning and will become an error in R 2.12.0.

INSTALLATION:

- UTF-8 is now used for the reference manual and package manuals. This requires LaTeX '2005/12/01' or later.
- `configure` looks for a POSIX compliant `tr`, Solaris's `/usr/ucb/tr` having been found to cause `Rdiff` to malfunction.
- `configure` is now generated with `autoconf 2.65`, which works better on recent systems and on Mac OS X.

PACKAGE INSTALLATION:

- Characters in R source which are not translatable to the current locale are now handled more tolerantly: these will be converted to hex codes with a warning. Such characters are only really portable if they appear in comments.
- R CMD INSTALL now tests that the installed package can be loaded (and backs out the installation if it cannot): this can be suppressed by '`--no-test-load`'. This avoids installing/updating a package that cannot be used: common causes of failures to load are missing/incompatible external software and missing/broken dependent packages.
- Package installation on Windows for a package with a '`src`' directory now checks if a DLL is created unless there is a '`src/Makefile.win`' file: this helps catch broken installations where the toolchain has not reported problems in building the DLL. (Note: this can be any DLL, not just one named '`<pkg-name>.dll`'.)

BUG FIXES:

- Using `with()`, `eval()` etc with a list with some unnamed elements now works. ([PR#14035](#))
- The "quick" dispatch of S4 methods for primitive functions was not happening, forcing a search each time. (Dispatch for closures was not affected.) A side effect is that default values for arguments in a method that do not have defaults in the generic will now be ignored.

- Trying to dispatch S4 methods for primitives during the search for inherited methods slows that search down and potentially could cause an infinite recursion. An internal switch was added to turn off all such methods from `findInheritedMethods()`.
- R framework installation (on Mac OS X) would not work properly if a rogue `Resources` directory was present at the top level. Such a non-symlink will now be renamed to `Resources.old` (and anything previously named `Resources.old` removed) as part of the framework installation process.
- The checks for conforming S4 method arguments could fail when the signature of the generic function omitted some of the formal arguments (in addition to `...`). Arguments omitted from the method definition but conforming (per the documentation) should now be ignored (treated as "ANY") in dispatching.
- The computations for S4 method evaluation when `...` was in the signature could fail, treating `...` as an ordinary symbol. This has been fixed, for the known cases.
- Various `ar()` fitting methods have more protection for singular fits.
- `callNextMethod` now works again with the `drop=` argument in `'[`
- `parse()` and `parse_Rd()` miscounted columns when multibyte UTF-8 characters were present.
- Formatting of help pages has had minor improvements: extra blank lines have been removed from the text format, and empty package labels removed from HTML.
- `cor(A, B)` where `A` is $n \times 1$ and `B` a 1-dimensional array segfaulted or gave an internal error. (The case `cor(B, A)` was [PR#7116](#).)
- `cut.POSIXt()` applied to a start value after the DST transition on a DST-change day could give the wrong time for argument `breaks` in units of days or longer. ([PR#14208](#))
- `do_par()` UNPROTECTED too early ([PR#14214](#))
- Subassignment `x[[...]] <- y` didn't check for a zero-length right hand side, and inserted a rubbish value. ([PR#14217](#))
- `fisher.test()` no longer gives a P-value **very** slightly > 1 , in some borderline cases.
- Internal function `matchArgs()` no longer modifies the general purpose bits of the SEXPs that make up the formals list of R functions. This fixes an invalid error message that would occur when a garbage collection triggered a second call to `matchArgs` for the same function *via* a finalizer.
- `gsub()` in 2.10.x could fail from stack overflow for extremely long strings due to temporary data being allocated on the stack. Also, `gsub()` with `fixed=TRUE` is in some circumstances considerably faster.
- Several primitives, including `attributes()`, `attr<-()`, `interactive()`, `nargs()` and `proc.time()`, did not check that they were called with the correct number of arguments.
- A potential race condition in `list.files()` when other processes are operating on the directory has been fixed; the code now dynamically allocates memory for file listings in a single pass instead of making an initial count pass.
- `mean(x, trim=, na.rm = FALSE)` failed to return NA if `x` contained missing values. (Reported by Bill Dunlap.)
- Extreme tail behavior of, `pbeta()` {and hence `pf()`}, e.g., `pbeta(x, 3, 2200, lower.tail=FALSE, log.p=TRUE)` now returns finite values instead of jumping to $-\text{Inf}$ too early. ([PR#14230](#)).
- `parse(text=x)` misbehaved for objects `x` that were not coerced internally to character, notably symbols. (Reported to R-devel by Bill Dunlap.)
- The internal C function `coerceSymbol` now handles coercion to character, and warns if coercion fails (rather than silently returning NULL). This allows a name to be given where a character vector is required in functions which coerce internally.

- The interpretation by `strptime()` of `"%c"` was non-standard (not that it is ever advisable to use locale- and system-specific input formats).
- `capabilities("X11")` now works the same way on Mac OS X as on other platforms (and as documented: it was always true for R built with `'--with-aqua'`, as the CRAN builds are).
- The `X11()` device with `cairo` but not `Pango` (notably Mac OS X) now checks validity of text strings in UTF-8 locales (since `Pango` does but `cairo` it seems does not).
- `read.fwf()` misread multi-line records when `n` was specified. (PR#14241)
- `all.equal(tolerance = e)` passes the numeric tolerance also to the comparison of the attributes.
- `pgamma(0, 0)`, a boundary case, now returns 0, its limit from the left, rather than the limit from the right.
- Issuing POST requests to the internal web server could stall the request under certain circumstances.
- `gzcon(<textConnection>)`, an error, no longer damages the connection (in a way to have it segfault). (PR#14237)
- All the results from `hist()` now use the nominal `breaks` not those adjusted by the numeric 'fuzz': in recent versions the nominal `breaks` were reported but the 'density' referred to the intervals used in the calculation – which mattered very slightly for one of the extreme bins. (Based on a report by Martin Becker.)
- If `xy[z].coords` (used internally by many graphics functions) are given a list as `x`, they now check that the list has suitable names and give a more informative error message. (PR#13936)

CHANGES IN R VERSION 2.10.1 patched

NEW FEATURES:

- The handling of line textures in the `postscript()` and `pdf()` devices was set up for round end caps (the only type which existed at the time): it has now been adjusted for butt endcaps.
- `lchoose(a, k)` is now defined as `log(abs(choose(a,k)))`, analogously to `lfactorial()`.
- Although `\eqn{}` in Rd files is defined as a 'verbatim' macro, many packages expected `...` and `...` to be interpreted there (as was the case in R < 2.10.0), so this is now done (using an ellipsis in HTML rendering).
- Escaping of braces in quoted strings in R-code sections of Rd files is allowed again. This had been changed for the new Rd format in R 2.10.0 but was only documented on the developer site and was handled inconsistently by the converters: text and example conversion removed the escapes but HTML conversion did not.
- The PCRE library has been updated to version 8.01, a bug-fix release.
- `tools::readNEWS()` now accepts a digit as the first character of a news section.

BUG FIXES:

- Using `read.table(header=TRUE)` on a header with an embedded new line would copy part of the header into the data. (PR#14103)
- `qpois(p = 1, lambda = 0)` now gives 0 as for all other `p`. (PR#14135)
- Functions related to string comparison (e.g. `unique()`, `match()`) could cause crashes when used with strings not in the native encoding, e.g. UTF-8 strings on Windows. (PR#14114 and PR#14125)
- `x[, drop=TRUE]` dropped an NA level even if it was in use.

- The dynamic HTML help system reported the wrong MIME type for the style sheet.
- `tools::codoc()` (used by `R CMD check`) was missing cases where the function had no arguments but was documented to have some.
- Help links containing special characters (e.g. "?") were not generated correctly when rendered in HTML. ([PR#14155](#))
- `lchoose(a, k)` no longer wrongly gives NaN for negative `a`.
- `ks.test()` could give a p-value that was off by one observation due to rounding error. ([PR#14145](#))
- `readBin()/readChar()` when reading millions of character strings in a single call used excessive amounts of memory (which also slowed them down).
- R CMD SHLIB could fail if used with paths that were not alphanumeric, e.g. contained `+`. ([PR#14168](#))
- `sprintf()` was not re-entrant, which potentially caused problems if an `as.character()` method called it.
- The `quartz()` device did not restore the clipping region when filling the background for a new page. This could be observed in multi-page bitmap output as stale outer regions of the plot.
- `p.adjust(method, n)` now works correctly for the rare case `n > length(p)`, also when method differs from `"bonferroni"` or `"none"`, thanks to a patch from Gordon Smyth.
- `tools::showNonASCII()` failed to detect non-ASCII characters if `iconv()` (incorrectly) converted them to different ASCII characters. (Seen on Windows only.)
- `tcrossprod()` wrongly failed in some cases when one of the arguments was a vector and the other a matrix.
- `[cr]bind(..., deparse.level=2)` was not always giving names when documented to do so. (Discovered whilst investigating [PR#14189](#).)
- `match(incomparables=<non-NULL>)` could in rare cases infinite-loop.
- `poisson.test()` needed to pass argument `conf.level` to `binom.test()`. ([PR#14195](#))
- The `"nls"` method for `df.residual()` gave incorrect results for models fitted with `na.action = na.exclude`. ([PR#14194](#))
- A change to `options(scipen=)` was only implemented when printing next occurred, even though it should have affected intervening calls to `axis()`, `contour()` and `filledcontour()`.
- `prettyNum(drop0trailing=TRUE)` did not handle signs of imaginary parts of complex numbers correctly (and this was used by `str()`: [PR#14201](#)).
- `system.time()` had the `sys.child` component wrong (copied `user.child` instead) on systems with `HAVE_GETRUSAGE`. ([PR#14210](#))
- Changing both line texture and line cap (end) resulted in the latter to be omitted from the PDF code. In addition, line cap (end) and join are now set explicitly in PDF output to ensure correct defaults.
- The suppression of auto-rotation in `bitmap()` and `dev2bitmap()` with the `"pdfwrite"` device was not working correctly.
- `plot(ecdf(), log="x")` no longer gives an incorrect warning.
- `read.fwf()` works again when argument `file` is a connection.
- Startup files will now be found if their paths exceed 255 bytes. ([PR#14228](#))
- `contrasts<-` (in the `stats` package) no longer has an undeclared dependence on `methods` (introduced in 2.10.0).

CHANGES IN R VERSION 2.10.1

NEW FEATURES:

- The PCRE library has been updated to version 8.00.
- R CMD INSTALL has new options ‘--no-R’, ‘--no-libs’, ‘--no-data’, ‘--no-help’, ‘--no-demo’, ‘--no-exec’, and ‘--no-inst’ to suppress installation of the specified part of the package. These are intended for special purposes (e.g. building a database of help pages without fully installing all packages).
- The documented line-length limit of 4095 bytes when reading from the console now also applies also to `parse(file="")` (which previously had a limit of around 1024 bytes).
- A Bioconductor mirror can be set for use by `setRepositories()` *via* the option "BioC_mirror", e.g. the European mirror can be selected by `options(BioC_mirror="http://bioconductor.statistik.tu-dortmund.de")`.
- Double-clicking in a `tk_select.list()` list box now selects the item and closes the list box (as happens on the Windows `select.list()` widget).

INSTALLATION:

- `configure` will be able to find a usable `libtiff` in some rare circumstances where it did not previously (where `libtiff` needed to be linked explicitly against `-ljpeg`).
- Making `refman.pdf` works around a problem with the indexing with `hyperref 6.79d` and later.

DEPRECATED & DEFUNCT:

- The `extended` argument is deprecated in `strsplit()`, `grep()`, `grepl()`, `sub()`, `gsub()`, `regexpr()` and `gregexpr()` (not just the value `extended = FALSE`) and will be removed in R 2.11.0.

BUG FIXES:

- `trigamma(x)` and other `psigamma(x, n)` calls are now accurate for very large `abs(x)`. (PR#14020)
- `[g]sub(perl=FALSE, fixed=FALSE)` could use excessive stack space when used with a very long vector containing some non-ASCII strings.
- The default method of `weighted.mean(na.rm = TRUE)` did not omit weights for NA observations in 2.10.0. (PR#14032)
- `[g]regexpr(pattern, fixed = TRUE)` returned match positions in bytes (not characters) in an MBCS locale if `pattern` was a single byte.
`[g]sub(fixed = TRUE)` with a single-byte pattern could conceivably have matched part of a multibyte character in a non-UTF-8 MBCS.
- `findLineNum()` and `setBreakpoint()` would sometimes fail if the specified file was not in the current directory.
- Package `tcltk`'s `demo(tkdensity)` was broken in 2.9.0 when `demo()` was changed to `set par(ask = TRUE)`.
- `gsub()` with backrefs could fail on extremely long strings (hundreds of thousands of characters) due to integer overflow in a length calculation.
- `abline(utf=TRUE)` now uses a better x-grid in log-scale, e.g., for `plot(c(1,300), c(1,300), log="xy"); abline(4,1, utf=TRUE)`.
- `detach()/unloadNamespace()` arrange to flush the package's lazyload cache of R objects once the package/namespace is no longer needed.

- There have been small fixes to the rendering of help, e.g. `\command` is now rendered verbatim (so e.g. `--` is not interpreted, [PR#14045](#)).
Also, there are many small changes to help files where the new converters were not rendering them in the same way as before.
- `available.packages()` would fail when run on a repository with no packages meeting the filtering conditions. ([PR#14042](#))
- `rep(x, times, each = 2)` gave invalid results when the `times` argument was a vector longer than `x`. Reported by Bill Dunlap.
- An error when `unloadNamespace()` attempted to run the `.onUnload()` function gave an error in the reporting function and so was not reported properly.
- Text help rendering did not handle very long input lines properly.
- `promptMethods()` generated signature documentation improperly.
- `pgamma(x, a, lower.tail=FALSE)` and `qgamma(...)` are now considerably more accurate in some regions for very small `a`. `qgamma()` now correctly returns 0 instead of NaN in similar extreme cases, and `qgamma()` no longer warns in the case of small `a`, see ([PR#12324](#)).
- `unname()` now also removes names from a zero length vector.
- Printing results from `ls.str()` no longer evaluates unevaluated calls.
- `complete.cases()` failed on a 0-column data frame argument. (Underlies [PR#14066](#).)
It could return nonsensical results if no input determined the number of cases (seen in the no-segfault tests).
- An error in `nls()` with a long formula could cause a segfault. ([PR#14059](#))
- `qchisq(p, df, ncp, lower.tail = FALSE)` with `ncp >= 80` was inaccurate for small `p` (as the help page said): it is now less inaccurate. (In part, [PR#13999](#).)
For `ncp` less than but close to 80, `pchisq()` and `qchisq()` are more accurate for probabilities very close to 1 (a series expansion was truncated slightly too early).
`pchisq(x, df, ncp)` can no longer return values just larger than one for large values of `ncp`.
- `intToUtf8()` could fail when asked to produce 10Mb or more strings, something it was never intended to do: unfortunately Windows crashed R (other OSes reported a lack of resources). ([PR#14068](#))
- `chisq.test()` could fail when given argument `x` or `y` which deparsed to more than one line. (Reported by Laurent Gauthier.)
- S4 methods are uncached whenever the name space containing them is unloaded (by `unloadNamespace()` as well as by `detach(unload = TRUE)`).
- The internal record-keeping by `dyn.load/dyn.unload` was incomplete, which could crash R if a DLL that registered `.External` routines had earlier been unloaded.
- `bessel[JY](x, nu)` with `nu` a negative integer (a singular case) is now correct, analogously to `besselI()`, see [PR#13556](#).
- `tools::file_path_as_absolute()` doubled the file separator when applied to a file such as `"/vmunix"` or (on Windows) `"d:/afile"` in a directory for which `getwd()` would return a path with a trailing separator (largely cosmetic, as reasonable file systems handle such a path correctly). (Perhaps what was meant by [PR#14078](#).)
- `unsplit(drop = TRUE)` applied to a data frame failed to pass `drop` to the computation of row names. ([PR#14084](#))
- The `"difftime"` method of `mean()` ignored its `na.rm` argument.
- `tcltk::tk_select.list()` is now more likely to remove the widget immediately after selection is complete.

- Adding/subtracting a "difftime" object to/from a "POSIXt" or "Date" object works again (it was broken by the addition of `Ops.difftime`).
- Conversion to latex of an Rd file with no aliases failed.
- `wilcox.test(conf.int=TRUE)` has `achieved.level` corrected and, for `exact=FALSE`, now returns a `estimate` component which does not depend on the alternative used.
- `help.search()` failed when the package argument was specified. ([PR#14113](#))
- `switch(EXPR = "A")` now returns `NULL`, as does `switch(1)` (which used to signal an error).

CHANGES IN R VERSION 2.10.0

SIGNIFICANT USER-VISIBLE CHANGES:

- Package help is now converted from Rd by the R-based converters that were first introduced in 2.9.0. This means
 - Packages that were installed by R-devel after 2009-08-09 should not be used with earlier versions of R, and most aspects of package help (including the runnable examples) will be missing if they are so used.
 - Text, HTML and latex help and examples for packages installed under the new system are converted on-demand from stored parsed Rd files. (Conversions stored in packages installed under `R < 2.10.0` are used if no parsed Rd files are found. It is recommended that such packages be re-installed.)
- HTML help is now generated dynamically using an HTTP server running in the R process and listening on the loopback interface.
 - Those worried about security implications of such a server can disable it by setting the environment variable `R_DISABLE_HTTPD` to a non-empty value. This disables `help.start()` and HTML help (so text help is shown instead).
 - The Java/Javascript search engine has been replaced by an HTML interface to `help.search()`. `help.start()` no longer has an argument `searchEngine` as it is no longer needed.
 - The HTML help can now locate cross-references of the form `\link[pkg]{foo}` and `\link[pkg:foo]{bar}` where `foo` is an alias in the package, rather than the documented (basename of a) filename (since the documentation has been much ignored).

NEW FEATURES:

- `polygon()`, `pdf()` and `postscript()` now have an argument `fillOddEven` (default `FALSE`), which controls the mode used for polygon fills of self-intersecting shapes.
- New `debugonce()` function; further, `getOption("deparse.max.lines")` is now observed when debugging, from a code suggestion by John Brzustowski. ([PR#13647/8](#))
- `plot()` methods for "stepfun" and hence "ecdf" no longer plot points by default for `n >= 1000`.
- `[g]sub(perl=TRUE)` now also supports "\E" in order to end "\U" and "\L" case changes, thanks to a patch from Bill Dunlap.
- `factor()`, `'levels<-'()`, etc, now ensure that the resulting factor levels are unique (as was always the implied intention). Factors with duplicated levels are still constructible by low-level means, but are now declared illegal.
- New `print()` (S3) method for class "function", also used for auto-printing. Further, `.Primitive` functions now print and auto-print identically. The new method is based on code suggestions by Romain François.

- The `print()` and `toLatex()` methods for class `"sessionInfo"` now show the locale in a nicer format and have arguments to suppress locale information.
- In addition to previously only `round()`, there are other `Math` group (S3) methods for `"difftime"`, such as `floor()`, `signif()`, `abs()`, etc.
- For completeness, `old.packages()` and `available.packages()` allow arguments `type` to be specified (you could always specify arguments `available` or `contriburl`).
- `available.packages()` by default only returns information on the latest versions of packages whose version requirements are satisfied by the currently running R.
- `tools::write_PACKAGES()` has a new argument `latestOnly`, which defaults to `TRUE` when only the latest versions in the repository will be listed in the index.
- `getOption()` has a new argument `default` that is returned if the specified option is not set. This simplifies querying a value and checking whether it is `NULL` or not.
- `parse()` now warns if the requested encoding is not supported.
- The `"table"` method of `as.data.frame()` gains a `stringsAsFactors` argument to allow the classifying factors to be returned as character vectors rather than the default factor type.
- If `model.frame.default()` encounters a character variable where `xlev` indicates a factor, it now converts the variable to a factor (with a warning).
- `curve()` now returns a list containing the points that were drawn.
- `spineplot()` now accepts `axes = FALSE`, for consistency with other functions called by `plot.factor()`.
- The Kendall and Spearman methods of `cor.test()` can optionally use continuity correction when not computing exact p-values. (The Kendall case is the wish of [PR#13691](#).)
- R now keeps track of line numbers during execution for code sourced with `options(keep.source = TRUE)`. The source reference is displayed by debugging functions such as `traceback()`, `browser()`, `recover()`, and `dump.frames()`, and is stored as an attribute on each element returned by `sys.calls()`.
- More functions now have an implicit (S4) generic definition.
- `quantile.default()` now disallows factors (wish of [PR#13631](#)) and its help documents what numeric-like properties its input need to have to work correctly.
- `weighted.mean()` is now generic and has `"Date"`, `"POSIXct"` and `"POSIXlt"` methods.
- Naming subscripts (e.g. `x[i=1, j=2]`) in `data.frame` methods for `[` and `[[` now gives a warning. (Names are ignored in the default method, but could have odd semantics for other methods, and do for the `data.frame` ones.)
- `as.data.frame()` has an `"aovproj"` method. (Wish of [PR#13505](#))
- `as.character(x)` for numeric `x` no longer produces strings such as `"0.30"`, i.e., with trailing zeros. This change also renders levels construction in `factor()` more consistent.
- `codocClasses()`, which checks consistency of the documentation of S4 class slots, now does so in considerably more cases. The documentation of inherited slots (from superclasses) is now optional. This affects R CMD `check <pkg>` when the package defines S4 classes.
- `codoc()` now also checks S4 methods for code/documentation mismatches.
- `for()`, `while()`, and `repeat()` loops now always return `NULL` as their (invisible) value. This change was needed to address a reference counting bug without creating performance penalties for some common use cases.
- The `print()` method for `ls.str()` results now obeys an optional `digits` argument.

- The `method` argument of `glm()` now allows user-contributed methods.
- More general `reorder.default()` replaces functionality of `reorder.factor()` and `reorder.character()`.
- The function `aspell()` has been added to provide an interface to the `Aspell` spell-checker.
- Filters `RdTextFilter()` and `SweaveTeXFilter()` have been added to the `tools` package to provide support for `aspell()` or other spell checkers.
- `xtabs()` with the new argument `sparse = TRUE` now returns a sparse Matrix, using package **Matrix**.
- `contr.sum()` etc gain an argument `sparse` which allows sparse matrices to be returned.
`contrasts()` also gains a `sparse` argument which it passes to the actual contrast function if that has a formal argument `sparse`.
`contrasts(f, .) <- val` now also works when `val` is a sparse Matrix. It is planned that `model.matrix()` will work with such factors `f` in the future.
- `readNEWS()` will recognize a UTF-8 byte-order mark (BOM) in the 'NEWS' file. However, it is safer to use only ASCII code there because not all editors recognize BOMs.
- New utility function `inheritedSlotNames()` for S4 class programming.
- `tabulate()` now allows NAs to pass through (and be ignored).
- If `debug()` is called on an S3 generic function then all methods are debugged as well.
- Outlier symbols drawn by `boxplot()` now obey the `outlwd` argument. Reported by Jurgen Kluge.
- `svd(x)` and `eigen(x)` now behave analogously to `qr(x)` in accepting logical matrices `x`.
- File 'NEWS' is now in UTF-8, and has a BOM (often invisible) on the first line, and Emacs local variables set for UTF-8 at the end. `RShowDoc("NEWS")` should display this correctly, given suitable fonts.
- `terms.formula(simplify = TRUE)` now does not deparse the LHS and so preserves non-standard responses such as 'a: b' (requested by Sundar Dorai-Raj).
- New function `news()` for building and querying R or package news information.
- z^n for integer `n` and complex `z` is more accurate now if `|n| <= 65536`.
- `factor(NULL)` now returns the same as `factor(character(0))` instead of an error, and `table(NULL)` consequently does analogously.
- `as.data.frame.vector()` (and its copies) is slightly faster by avoiding a copy if there are no names (following a suggestion of Tim Hesterberg).
- `writeLines()`, `writeBin()` and `writeChar()` have a new argument `useBytes`. If false, character strings with marked encodings are translated to the current locale (as before) but if true they are written byte-by-byte.
- `iconv()` has a new argument `mark` which can be used (by experts) to suppress the declaration of encodings.
- 'DESCRIPTION' file's `LinkingTo` specs are now recognized as installation dependencies, and included in package management computations.
- Standardized 'DESCRIPTION' file `License` specs are now available for package management computations.
- `"\uxxxx"` and `"\Uxxxxxxxx"` escapes can now be parsed to a UTF-8 encoded string even in non-UTF-8 locales (this has been implemented on Windows since R 2.7.0). The semantics have been changed slightly: a string containing such escapes is always stored in UTF-8 (and hence is suitable for portably including Unicode text in packages).
- New `as.raw()` method for `"tc10bj"` objects (wish of [PR#13758](#)).

- `Rd.sty` now makes a better job of setting email addresses, including using a monospaced font.
- `textConnection()` gains an `encoding` argument to determine how input strings with marked encodings will be handled.
- R CMD `Rd2pdf` is available as a shortcut for R CMD `Rd2dvi --pdf`.
- R CMD `check` now checks links where a package is specified (`\link[pkg]{file}` or `\link[pkg:file]{topic}`), if the package is available. It notes if the package is not available, as in many cases this is an error in the link.
- `identical()` gains three logical arguments, which allow for even more differentiation, notably `-0` and `0`.
- `legend()` now can specify the `border` color of filled boxes, thanks to a patch from Frederic Schutz.
- Indexing with a vector index to `[[]]` has now been extended to all recursive types.
- Pairlists may now be assigned as elements of lists. (Lists could always be created with pairlist elements, but `[[<-` didn't support assigning them.)
- The parser now supports C-preprocessor-like `#line` directives, so error messages and source references may refer to the original file rather than an intermediate one.
- New functions `findLineNum()` and `setBreakpoint()` work with the source references to find the location of source lines and set breakpoints (using `trace()`) at those lines.
- Namespace importing is more careful about warning on masked generics, thanks to a patch by Yohan Chalabi.
- `detach()` now has an argument `character.only` with the same meaning as for `library()` or `require()`.
- `available.packages()` gains a `filters` argument for specifying the filtering operations performed on the packages found in the repositories. A new built-in "license/FOSS" filter only retains packages for which installation can proceed solely based on packages which can be verified as Free or Open Source Software (FOSS) employing the available license specifications.
- In registering an S3 class by a call to `setOldClass()`, the data part (e.g., the object type) required for the class can be included as one of the superclasses in the `Classes` argument.
- The argument `f` to `showMethods()` can be an expression evaluating to a generic function, allowing methods to be shown for non-exported generics and other nonstandard cases.
- `sprintf()` now supports `%o` for octal conversions.
- New function `Sys.readlink()` for information about symbolic links, including if a file is a symbolic link.
- Package **tools** has new functions `checkRdaFiles()` and `resaveRdaFiles()` to report on the format of `‘.rda’/‘.RData’` data files, and to re-save them in a different compressed format, including choosing the most compact format available. A new `INSTALL` option, `‘--resave-data’`, makes use of this.
- File `‘~/R/config’` is used in preference to `‘~/R.Rconfig’`, and these are now documented in ‘R Installation and Administration’.
- Logic operations with complex numbers now work, as they were always documented to, and as in S.
- `arrows()` and `segments()` allow one of `x1` or `y1` to be omitted to simplify the specification of vertical or horizontal lines (suggestion of Tim Hesterberg).
- `approxfun()` is faster by avoiding repeated `NA` checks (diagnosis and patch by Karline Soetaert & Thomas Petzoldt).

- There are the beginnings of a Nynorsk translation by Karl Ove Hufthammer.
- `stripchart()` allows `par bg` to be passed in for the background colour for `pch = 21` (wish of [PR#13984](#)).
- New generic function `.DollarNames()` to enable class authors to customize completion after the `$` extractor.
- `load()`, `save()`, `dput()` and `dump()` now open a not-yet-open connection in the appropriate mode (as other functions using connections directly already did).

REGULAR EXPRESSIONS:

- A different regular expression engine is used for basic and extended regexps and is also for approximate matching. This is based on the TRE library of Ville Laurikari, a modified copy of which is included in the R sources.
This is often faster, especially in a MBCS locale.
Some known differences are that it is less tolerant of invalid inputs in MBCS locales, and in its interpretation of undefined (extended) regexps such as `"^*"`. Also, the interpretation of ranges such as `[W-z]` in caseless matching is no longer to map the range to lower case.
This engine may in future be used in 'literal' mode for `fixed = TRUE`, and there is a compile-time option in `src/main/grep.c` to do so.
- The use of repeated boundary regexps in `gsub()` and `gregexpr()` as warned about in the help page does not work in this engine (it did in the previous one since 2005).
- Extended (and basic) regexps now support same set of options as for `fixed = TRUE` and `perl = TRUE`, including `useBytes` and support for UTF-8-encoded strings in non-UTF-8 locales.
- `agrep()` now has full support for MBCS locales with a modest speed penalty. This enables `help.search()` to use approximate matching character-wise rather than byte-wise.
- `[g]sub` use a single-pass algorithm instead of matching twice and so is usually faster.
- The `perl = TRUE` versions now work correctly in a non-UTF-8 MBCS locale, by translating the inputs to UTF-8.
- `useBytes = TRUE` now inhibits the translation of inputs with marked encodings.
- `strsplit()` gains a `useBytes` argument.
- The algorithm used by `strsplit()` has been reordered to batch by elements of `split`: this can be much faster for `fixed = FALSE` (as multiple compilation of regexps is avoided).
- The help pages, including `?regexp`, have been updated and should be consulted for details of the new implementations.

HELP & Rd FILE CHANGES:

- A new dynamic HTML help system is used by default, and may be controlled using `tools::startDynamicHelp()`. With this enabled, HTML help pages will be generated on request, resolving links by searching through the current `.libPaths()`. The user may set `options("help.ports")` to control which IP port is used by the server.
- `help.start()` no longer sets `options(htmlhelp = TRUE)` (it used to on Unix but not on Windows). Nor does it on Unix reset the `"browser"` option if given an argument of that name.
Arguments `update` and `remote` are now available on all platforms: the default is `update = FALSE` since the http server will update the package index at first use.
- `help()` has a new argument `help_type` (with default set by the option of that name) to supersede arguments `offline`, `htmlhelp` and `chmhelp` (although for now they still

work if `help_type` is unset). There is a new type, "PDF" to allow offline PDF (rather than PostScript).

A function `offline_help_helper()` will be used if this exists in the workspace or further down the search path, otherwise the function of that name in the `utils` name space is used.

- Plain text help is now used as the fallback for HTML help (as it always was for Compiled HTML help on Windows).
- It is possible to ask for static HTML pages to be prebuilt *via* the configure option '`--enable-prebuilt-html`'. This may be useful for those who wish to make HTML help available outside R, e.g. on a local web site.
- An experimental tag `\Sexpr` has been added to Rd files, to evaluate expressions at build, install, or render time. Currently install time and render time evaluation are supported.
- Tags `\if`, `\ifelse` and `\out` have been added to allow format-specific (or more general, using `\Sexpr`) conditional text in man pages.
- The `parse_Rd()` parser has been made more tolerant of coding errors in Rd files: now all syntax errors are reported as warnings, and an attempt is made to continue parsing.
- `parse_Rd()` now has an argument `fragment` (default `FALSE`) to accept small fragments of Rd files (so that `\Sexpr` can output Rd code which is then parsed).
- `parse_Rd()` now always converts its input to UTF-8. The `Rd2*` rendering functions have a new argument, `outputEncoding`, which controls how their output is encoded.
- `parse_Rd()` no longer includes the newline as part of a `"%"`-style comment.
- There have been various bug fixes and code reorganization in the Rd renderers `Rd2HTML`, `Rd2latex`, `Rd2txt`, and `Rd2ex`.
All example files are now created with either ASCII or UTF-8 encoding, and the encoding is only marked in the file if there is any non-UTF-8 code (previously it was marked if the help file had non-ASCII contents, possibly in other sections).
- `print.Rd()` now adds necessary escape characters so that printing and re-parsing an Rd object should produce an equivalent object.
- `parse_Rd()` was incorrectly handling multiple backslashes in R code strings, converting $4n+3$ backslashes to $2n+1$ instead of $2n+2$.
- `parse_Rd()` now recognizes the `\var` tag within a quoted string in R-like text.
- `parse_Rd()` now treats the argument of `\command` as LaTeX-like, rather than verbatim.

COMPRESSION:

- New function `untar()` to list or unpack tar archives, possibly compressed. This uses either an external `tar` command or an internal implementation.
- New function `tar()` to create (possibly compressed) tar archives.
- New functions `memCompress()` and `memDecompress()` for in-memory compression and decompression.
- `bzfile()` has a `compress` argument to select the amount of effort put into compression when writing.
- New function `xzfile()` for use with `xz`-compressed files. (This can also read files compressed by some versions of `lzma`.)
- `gzfile()` looks at the file header and so can now also read `bzip2`-ed files and `xz`-compressed files.

- There are the new options of `save(compress = "bzip2")` and `"xz"` to use `bzip2` or `xz` compression (which will be slower, but can give substantially smaller files). Argument `compression_level` gives finer control over the space/time tradeoffs. `load()` can read such saves (but only as from this version of R).
- R CMD `INSTALL/check` and `tools::writePACKAGES` accept a wider range of compressed tar archives. Precisely how wide depends on the capabilities of the host system's `tar` command: they almost always include `.tar.bz2` archives, and with modern versions of `tar` other forms of compression such as `lzma` and `xz`, and arbitrary extensions.
- R CMD `INSTALL` has a new option `'--data-compress'` to control the compression used when lazy-loading data. New possibilities are `'--data-compress=bzip2'` which will give ca 15% better compression at the expense of slower installation times, and `'--data-compress=xz'`, often giving even better compression on large datasets at the expense of much longer installation times. (The latter is used for the recommended packages: it is particularly effective for `survival`.)
- `file()` for `open = "", "r" or "rt"` will automatically detect compressed files (from `gzip`, `bzip2` or `xz`). This means that compressed files can be specified by file name (rather than *via* a `gzfile()` connection) to `read.table()`, `readlines()`, `scan()` and so on.
- `data()` can handle compressed text files with extensions `'{txt,tab,csv}.{gz,bz2,xz}'`.

DEPRECATED & DEFUNCT:

- `png(type="cairo1")` is defunct: the value is no longer recognized.
- `tools::Rd_parse()` is defunct (as this version of R uses only Rd version 2).
- Use of file `'~/Rconf'` (which was deprecated in favour of `'~/Rconfig'` in 2004) has finally been removed.
- Bundles of packages are deprecated. See 'Writing R Extensions' for the steps needed to unbundle a bundle.
- `help()` arguments `offline`, `htmlhelp` and `chmhelp` are deprecated in favour of `help_type`.
- `clearNames()` (in package `stats`) is deprecated for `unname()`.
- Basic regular expressions (`extended = FALSE`) are deprecated in `strsplit`, `grep` and friends. There is a precise POSIX standard for them, but it is not what recent RE engines implement, and it seems that in almost all cases package authors intended `fixed = TRUE` when using `extended = FALSE`.
- `methods::trySilent()` is deprecated in favour of `try(silent=TRUE)` or – more efficiently and flexibly – something like `tryCatch(error = function(e) e)`.
- `index.search()` is deprecated: there are no longer directories of types other than `help`.

INSTALLATION:

- `cairo >= 1.2` is now required (1.2.0 was released in July 2006) for cairo-based graphics devices (which remain optional).
- A suitable `iconv()` is now required: support for configure option `'--without-iconv'` has been withdrawn (it was deprecated in R 2.5.0).
- Perl is no longer 'essential'. R can be built without it, but scripts R CMD `build`, `check`, `Rprof` and `Sd2d` currently require it.
- A system `glob` function is now essential (a working `Sys.glob()` has been assumed since R 2.9.0 at least).

- C99 support for MBCS is now required, and configure option ‘--disable-mbcs’ has been withdrawn.
- Having a version of **tar** capable of automagically detecting compressed archives is useful for `utils::untar()`, and so **gtar** (a common name for GNU **tar**) is preferred to **tar**: set environment variable **TAR** to specify a particular **tar** command.

INTERNATIONALIZATION:

- There is some makefile support for adding/updating translations in packages: see `po/README` and ‘Writing R Extensions’.
- There is support for the use of `dngettext` for C-level translations in packages: see ‘Writing R Extensions’.

BUG FIXES:

- Assigning an extra 0-length column to a data frame by `DF[, "foo"] <- value` now works in most cases (by filling with NAs) or fails. (It used to give a corrupt data frame.)
- `validObject()` avoids an error during evaluation in the case of various incorrect slot definitions.
- `n:m` now returns a result of type `"integer"` in a few more boundary cases.
- The `zap.ind` argument to `printCoefmat()` did not usually work as other code attempted to ensure that non-zero values had a non-zero representation.
- `printCoefmat()` formatted groups of columns together, not just the `cs.ind` group but also the `zap.ind` group and a residual group. It now formats all columns except the `cs.ind` group separately (and zaps the `zap.ind` group column-by-column). The main effect will be seen in the output from `print.anova()`, as this grouped SS-like columns in the `zap.ind` group.
- `R_RepIDLLinit()` initializes the top-level jump so that some embedded applications on Windows no longer crash on error.
- `identical()` failed to take the encoding of character strings into account, so identical byte patterns are not necessarily identical strings, and similarly Latin-1 and UTF-8 versions of the same string differ in byte pattern.
- `methods(f)` used to warn unnecessarily for an S4 generic `f` which had been created based on an existing S3 generic.
- The check for consistent ordering of superclasses was not ignoring all conditional relations (the symptom was usually spurious warnings for classes extending `"array"`).
- Trying to assign into a raw vector with an index vector containing NAs could cause a segfault. Reported by Hervé Pagès.
- `Rscript` could segfault if (by user error) its filename argument was missing. Reported by Martin Morgan.
- `getAnywhere()` (and functions that use it, including argument completion in the console) did not handle special built-in functions. Reported by Romain Francois.
- `order()` was missing a `PROTECT()` call and so could segfault when called on character data under certain (rare) circumstances involving marked non-native encodings.
- `prettyNum(z, drop0trailing=TRUE)` did not work correctly when `z` was a complex vector. Consequently, `str(z, ...)` also did not. ([PR#13985](#))
- `make distclean` removed too many files in `etc/` if `builddir = srcdir`.
- R CMD replaced `TEXINPUTS` rather than appending to it (as documented and intended).
- `help.start()` no longer fails on unix when `"browser"` is a function.
- `pbeta(x, ..., log.p = TRUE)` is sometimes more accurate, e.g., for very small `x`.
- Unserializing a pre-2.8 workspace containing pure ASCII character objects with a Latin-1 or UTF-8 encoding would corrupt the `CHARSXP` cache.