Bibliography

K. Hornik (2002). *The R FAQ*. ISBN 3-901167-51-X, http://www.ci.tuwien.ac.at/~hornik/R/. 32

P. Murrell and R. Ihaka (2000). An Approach to Providing Mathematical Annotation in Plots, *Journal of Computational and Graphical Statistics*, 9(3): 582–599. 32

W. N. Venables and B. D. Ripley (2000). *S Programming*. Springer-Verlag, New York. 32

W. N. Venables (2002). Programmer's Niche, *R News*, 2(2): 24–26, ISSN 1609-3631, http://CRAN.R-project.org/doc/Rnews/. 32, 33

Uwe Ligges
Fachbereich Statistik, Universität Dortmund, Germany
ligges@statistik.uni-dortmund.de

Changes in R

by the R Core Team

User-visible changes

- The default colour palette now has "grey" instead of "white" in location 8. See palette().

- grid(nx) behaves differently (but the same as in R versions <= 0.64).

New features

- barplot() has a new argument 'axis.lty', which if set to 1 allows the pre-1.6.0 behaviour of plotting the axis and tick marks for the categorical axis. (This was apparently not intentional, but axis() used to ignore lty=0.) The argument 'border' is no longer "not yet used".

- New operator :: in the grammar, for name spaces.

- New faster rowsum(), also works on data frames.

- grep(), sub(), gsub() and regexpr() have a new argument 'perl' which if TRUE uses Perl-style regexps from PCRE (if installed). New capabilities option "PCRE" to say if PCRE is available.

- Preparations for name space support:

  – Functions in the **base** package are now defined in a name space. As a temporary measure, you can disable this by defining the environment variable R_NO_BASE_NAMESPACE.

  – UseMethod dispatching now searches for methods in the environment of the caller of the generic function rather than the environment where the generic is defined.

- The objects created in the **methods** package to represent classes, generic functions, method definitions, and inheritance relations now themselves belong to true classes. In particular, the "classRepresentation" objects follow the description in "Programming with Data" (section 7.6).

- Other additions and changes to the **methods** package:

  – The function setOldClass() has been added, following the description on page 450 of "Programming with Data". Use it if old-style classes are to be supplied in signatures for setMethod, particularly if the old-style classes have inheritance. Many of the old-style classes in the base package should be pre-specified; try getClass("mlm"), e.g.

  – The setGeneric() function applies some heuristics to warn about possibly erroneous generic function definitions. (Before, obscure bugs could result.)

  – The function promptMethods() has been revised to work better and to provide aliases for individual methods.

  – The behavior of the as() function has been generalized, in particular with a 'strict=' argument, the general goal being to let simple extensions of classes pass through in method dispatch and related computations without altering the objects. More to make method behavior more "natural" than for direct use.

  – Some inconsistencies following detach("package:methods") have been removed, so it *should* be possible to detach/re-attach the methods package.

- New methods ([[, print, str) and extended plot() method (including logical 'horiz') for "dendrogram" class.

- `sprintf()` now checks the agreement between formats and object types, and handles special values (`NA`, `Inf`, ...) correctly.

- `chol()` now uses a tolerance for non-positive-definiteness and so should give more consistent results across platforms.

- New function `agrep()` for approximate (fuzzy) string matching.

- `help.search()` can now use both approximate (fuzzy) and regular expression matching. By default, if the pattern to be matched consists of only alphanumeric characters, whitespace or a dash, approximate matching is used.

- `axis()` has three new optional arguments 'col', 'lty', and 'lwd' all for drawing the axis line and tick marks.

- Function `vcov()` (formerly in **MASS**), a generic function to return the variance-covariance matrix of the parameter estimates of a fitted model.

- `duplicated()` and `unique()` have methods for matrices and arrays (based on ideas from Jens Oehlschlägel).

- Internally memory sizes and counts of cons cells are now stored in unsigned longs. This allows memory limits to be set and objects created in the range 2-4Gb on 32-bit platforms, and allows 64-bit platforms to use much larger amounts of memory.

- Command-line flags to set memory can now use the suffix 'G' for gigabytes. The setting of maximum vsize is now only limited by the platform's address space.

- All warning and error messages are truncated to a length set by `options(warning.length=)`, defaulting to 1000. (Previously most (but not quite all) were truncated at 8192 characters.)

- `[dpqr]gamma()` check for shape parameter > 0.

- `as.POSIX[cl]t()` can now convert logical `NA`s.

- All installed packages (even those shipped with R) are given a 'Built' field in the 'DESCRIPTION' file.

- `as.data.frame()` now coerces logical matrices into logical columns (rather than factors).

- `[[<-.data.frame` no longer coerces character replacement values to factor. This is consistent with using '$' to replace and with S4.

- `library()` attempts to detect improperly installed packages, so as from this version an installed package must have a 'DESCRIPTION' file and that file must have been stamped with a 'Built:' line (which was introduced in 1.2.0). Under Unix-alikes, the platform is checked against that used for installation.

- `print.factor()` has new arguments 'max.levels' (with a smart default) and 'width'. `print.ordered()` is no longer needed.

- `RNGkind()` has an additional option for normal random generators: `"Inversion"`.

- `data.frame()` recycles factors and `"AsIs"` objects as well as atomic vectors.

- `predict.lm()` warns if 'newdata' is supplied and the fit was rank-deficient, as this can be misleading.

- `rect()` accepts additional graphics parameters through a '...' argument (in the same way as polygon).

- `strwidth()` and `strheight()` now coerce their first argument in exactly the same way `text()` does, so a wider range of inputs is allowed.

- `prompt()`'s default and data.frame methods have a new 3rd argument 'name' allowing them to used more easily in scripts and loops.

- `rgb()` has a new 'maxColorValue' argument, allowing r,g,b in $[0, M]$, particularly in $0 : 255$, efficiently and non-error-prone.

- `summaryRprof()` provides the functionality of R CMD Rprof in R code, though more slowly.

- `stop()` accepts multiple arguments (which are concatenated) just as `warning()` does.

- `scan()` now throws an error with incorrect logical input (which was previously taken as `FALSE`).

- `pdf()` now uses PDF not R code for clipping, which ensures that partially visible text strings are (partially) shown.

- Each R session uses a per-session temporary directory which is removed at normal termination. The directory name is given by the `tempdir()` function, and filenames returned by `tempfile()` will be within that directory.

- `help.start()` on Unix now uses a '.R' subdirectory of the per-session temporary directory and not '~/.R'. A side effect is that '~/.R' is now never deleted by R.

  This now uses the remote control mechanism only if the X display is local to the R process (as otherwise it might use a browser running on an arbitrary machine).

- *Very* experimental `browseEnv()` for browsing objects in an environment.

- `cbind()` and `rbind()` used to ignore all zero-length vectors, an undocumented quirk for S-compatibility. This caused problems when combining zero-extent matrices and zero-length vectors, and now zero-length vectors are ignored unless the result would have zero rows/columns.

- `read.table(stdin())` will now work.

- `plot.spec(x)` now also works for other x than AR and Pgram results.

- New functions `La.chol()` and `La.chol2inv()` for Cholesky decomposition and inverse of positive definite matrices using Lapack.

- Changes to the **tcltk** package:

  - Added a few "trivial and obviously missing" functions: `tkchooseDirectory`, `tkpopup`, `tkdialog`, `tkread`.

  - on Unix systems, the Tcl event loop has been integrated with R's own (so that `tkwait.variable()` no longer halts updates of plot windows).

  - also on Unix, stubs have been created to divert R's input and output routines to go via Tcl commands. (Nothing uses this at present, but packages might be developed to take advantage of it.)

  - return value from Tcl commands is no longer invisible. A new print method, `print.tclObj()`, has been introduced.

  - Tcl variables created by `tclVar()` are now explicitly put into Tcl's global namespace, removing potential scoping problems.

  - The tcltk dynamic library now loads with `local=FALSE` since the default had trouble when loading Tcl extensions (e.g., Tix).

  - The `tkpager()` function had not been updated for the return value change from 1.5.0.

- The `bmp()`, `jpeg()` and `png()` devices can produce multiple bitmap files, one for each page. The default filenames have been changed to include a sequence number.

- New function `axTicks()` returning tick mark locations like `axis()`.

- `grid()` has a more sensible default behavior. Tick axis alignment only happens when no numbers of grid cells are specified. New arguments 'lwd' and 'equilogs'; nx/ny = NA for not drawing, see ?grid.

- `installed.packages()` has a new argument 'priority'.

- `termplot()` uses factor levels rather than $1, 2, 3, \ldots$ for x-axis.

- Workaround for optimization bugs on gcc 3.1/2 on 32-bit Solaris.

- The `trace()` function has been robustified and a new function `tracingState()` added to turn tracing temporarily on and off.

- New `cophenetic()` in **mva** as utility for hierarchical clustering.

- `p.adjust()` has two new methods, 'Hommel' and 'FDR', contributed by Gordon Smyth <smyth@wehi.edu.au>.

- `stars()` now has add and plot arguments.

- Enhancements to mathematical annotation of plots:

  - expressions involving `dot`(*something*) now produce a dot accent above the *something* (initial patch from Ben Bolker).

  - within an expression, the symbol 'partialdiff' is now converted to a partial differential symbol (greek delta).

- `smooth.spline()` has a new argument 'nknots' allowing to set the default number of knots (when 'all.knots = FALSE' as per default).

Build Issues

- Toplevel 'Makefile' was missing dependency of 'docs' on 'R' (causing parallel makes to go wrong).

- When building with recommended packages those were installed into the first path in R_LIBS, if the environment variable was present.

Deprecated & defunct

- The assignment operator '_' is deprecated: a warning is given once per R session.

- `machine()`, `Machine()` and `Platform()` are deprecated in favour of `.Platform$OS.type`, `.Machine` and `.Platform`.

- `arima0.diag()` (package **ts**) is defunct.

- `piechart()` is defunct.

- `print.ordered()` has been removed, so `print.factor()` is used.

- The global internal variables `.Dyn.libs` and `.lib.loc` are removed in favor of the internal functions `.dynLibs()` and `.libPaths()`.

- `restart()` is deprecated in preparation for proper exception handling. Use `try()`, as has long been recommended.

Documentation changes

- New `demo(persp)` containing some of the former `example(persp)` ones and more.

C-level facilities

- 'Rversion.h' is no longer automatically included by 'R.h'. Include it explicitly if you need it.

- New entry point `R_tmpnam` in 'R_ext/Utils.h'.

- The Unix event loop interface has been changed to facilitate integration with other loops. `R_checkActivity` and `R_runHandlers` should eventually replace `getSelectedHandler`.

Installation changes

- Perl 5.005 or newer is now required.

- `R CMD INSTALL` is now guaranteed to sort the R source files in ASCII order.

Utilities

- `R CMD check` now tests for mis-use on an installed or binary package, and sets `T` and `F` to `NULL` when running the examples.

- New function `SweaveSyntConv()` converts between Sweave file syntaxes. `RweaveLatex()` now gets its prompt from `options()` and uses the text width as linebreak cutoff for deparsing input statements.

See the file 'NEWS' in the R distribution for additional information on bug fixes.

Changes on CRAN

by Kurt Hornik

CRAN packages

CGIwithR Facilities for the use of R to write CGI scripts. By David Firth.

ISwR Data sets and scripts for text examples and exercises in P. Dalgaard (2002), "Introductory Statistics with R", Springer Verlag. By Peter Dalgaard.

KMsurv Data sets and functions for Klein and Moeschberger (1997), "Survival Analysis, Techniques for Censored and Truncated Data", Springer. Original by Klein and Moeschberger, modifications by Jun Yan.

MPV Data sets from the book "Introduction to Linear Regression Analysis" by D. C. Montgomery, E. A. Peck, and C. G. Vining, 2001, John Wiley and Sons. By W. J. Braun.

RColorBrewer The package provides palettes for drawing nice maps shaded according to a variable. By Erich Neuwirth.

SparseM Basic linear algebra for sparse matrices. By Roger Koenker and Pin Ng.

StatDataML Read and write StatDataML files, alpha implementation of the StatDataML proposal. By Torsten Hothorn, Friedrich Leisch, and David Meyer.

ape Ape provides functions for reading, writing, and plotting phylogenetic trees in parenthetic format (standard Newick format), analyses of comparative data in a phylogenetic framework, analyses of diversification and macroevolution, computing distances from allelic and nucleotide data, reading nucleotide sequences from GenBank via internet, and several tools such as Mantel's test, computation of minimum spanning tree, or the population parameter theta based on various approaches. By Emmanuel Paradis, Korbibian Strimmer, Julien Claude, Yvonnick Noel, and Ben Bolker.

deal Bayesian networks with continuous and/or discrete variables can be learned and compared from data. By Susanne Gammelgaard Bøttcher and Claus Dethlefsen.

geepack Generalized estimating equations solver for parameters in mean, scale, and correlation structures, through mean link, scale link, and correlation link. Can also handle clustered categorical responses. By Jun Yan.