

Changes in R 2.6.1

by the R Core Team

New features

- The `data.frame` and `factor` methods for `[[` now support the `exact` argument introduced in 2.6.0.
- `plot.lm()` gains a new argument `cex.caption` to allow the size of the captions to be controlled.
- A series of changes make the `CHARSXP` cache introduced in 2.6.0 faster (and in some cases many times faster) in sessions with a large number (e.g., a million) of unique character strings, and also if there are many empty strings.
- `embedFonts()`, `bitmap()` and `dev2bitmap()` explicitly turn off auto-rotation in Ghostscript when generating PDF.
- The canonical architecture is no longer checked when loading packages using a non-empty sub-architecture, since it is possible to (e.g.) build packages for `i386-pc-linux-gnu` on both that architecture and on `x86_64-unknown-linux-gnu`.
- Deparsing will (if option `warnIncomplete` is set) warn on strings longer than the parser limit (8192 bytes).
- `url()` now uses the `UserAgent` header in `http` transactions in the same way as `download.file()` (making use of option `"HTTPUserAgent"`).

Bug fixes

- `iconv()` is again able to translate character strings with embedded nuls (such as those in UCS-2).
- `new.packages()` and `update.packages()` failed when called on an empty library, since `old.packages()` threw an error. `old.packages()` now returns `NULL` (as documented) in that case.
- Builds on Mac OS X 10.4 or higher now allocate enough space in the binary headers to relocate dependent libraries into the framework.
- R CMD build now computes the exclusion list on the copy it makes: this avoids problems if the original sources contain symbolic links

(which are resolved in the copy). Thanks to Michael Lawrence for diagnosis and patch.

- `object.size()` had slightly too low a size for objects of type `"S4"`.
- `symbol()` in `plotmath` expressions was only accepting valid character strings, which made it impossible to specify symbols such as `aleph` (obtained by `symbol("300")`) in a UTF-8 locale.
- An event handling issue caused `autorepeat` functions to misbehave with `tcltk` (notably scrollbars).
- `plot(sin, -5, 5)` gives `ylab "sin(x)"` again, where it resulted in `"x(x)"` in 2.6.0. Further, `plot(sin)` again plots from `[0,1]` also in cases where a previously used coordinate system differs.
- `curve()` with unspecified `from`, `to` and `xlim` now reuses the previous `x` limits, and not slightly larger ones.
- It was intended that R code filenames in packages should start with an ASCII letter or digits (and R CMD INSTALL uses that), but the test used in R CMD build (`[A-Za-z0-9]`) was locale-specific (and excluded `t` to `y` in Estonian, for example). (PR#10351)
- R CMD build could misbehave when faced with files with CRLF line endings *and* no line ending on the final line of the file, removing the last byte of the file.
- `DF[i, j]` failed in 2.6.0 if `j` was a logical vector selecting a single column.
- `Unix x11()` would fail if a valid `display` was specified but `DISPLAY` was unset. (PR#10379)
- `postscript()` was not always ignoring `.Postscript.Options` in the workspace (where it should not have occurred).
- `help.search()` would give an error if it found a badly installed package, even if package was not specified.
- `tclServiceMode()` (package `tcltk`) now works under Unix-alikes. (Although documented, it used only to work under Windows.)
- As Mac OS X 10.5.0 comes with incompatible `/bin/sh` shell, we force `SHELL=/bin/bash` (which is ok) in that case. [Only for 2.6.x: another solution is used in 2.7.0.]
- Deliberately using malformed source attributes no longer causes deparsing/printing of functions to crash R. (PR#10437)

- R CMD check and R CMD INSTALL now work with (some) directory names containing spaces.
- choose(n, k) gave incorrect values for negative n and small k.
- plot.ts(x,y) could use wrong default labels; fixed thanks to Antonio Fabio di Narzo.
- reshape() got column names out of sync with contents in some cases; found by Antonio Fabio Di Narzo.
- ar(x) for short x (i.e., length ≤ 10) could fail because the default order.max was \geq length(x) which is non-sensical.
- Keyboard events in getGraphicsEvent() could cause stack imbalance errors (PR#10453)

Changes on CRAN

by Kurt Hornik

New contributed packages

BARD Better Automated ReDistricting and heuristic exploration of redistricter revealed preference. By Micah Altman.

CellularAutomaton An object-oriented implementation of one-dimensional cellular automata. Supports many of the features offered by Mathematica, including elementary rules, user-defined rules, radii, user-defined seeding, and plotting. By John Hughes.

ComPairWise Compare phylogenetic or population genetic data alignments. By Trina E. Roberts.

EDR Estimation of the effective dimension reduction (EDR) space in multi-index regression models. By Joerg Polzehl.

FGN Fractional Gaussian Noise (FGN) model fitting, including MLEs for the H parameter and regression with FGN errors, and simulation of FGN. By A. I. McLeod.

FKBL Fuzzy Knowledge Base Learning, an R/C implementation of a fuzzy inference engine supporting several inference methods. By Alvaro Gonzalez Alvarez.

FieldSim Routines to simulate random fields. By Alexandre Brouste and Sophie Lambert-Lacroix.

GLDEX Fitting single and mixture of Generalized Lambda Distributions (RS and FMKL) using Discretized and Maximum Likelihood methods. The fitting algorithms considered have two major objectives. One is to provide a smoothing device to fit distributions to data using the weight and unweighted discretized approach based on the bin width of the histogram. The other is to provide a definitive fit

to the data set using the maximum likelihood estimation. Diagnostics on goodness of fit can be done via QQ-plots, KS-resample tests and comparing mean, variance, skewness and kurtosis of the data with the fitted distribution. By Steve Su.

GillespieSSA Gillespie's Stochastic Simulation Algorithm (SSA). Provides a simple to use, intuitive, and extensible interface to several stochastic simulation algorithms for generating simulated trajectories of finite population continuous-time models. Currently it implements Gillespie's exact stochastic simulation algorithm (Direct method) and several approximate methods (Explicit tau-leap, Binomial tau-leap, and Optimized tau-leap). Also contains a library of template models that can be run as demo models and can easily be customized and extended. Currently the following models are included: decaying-dimerization reaction set, linear chain system, logistic growth model, Lotka predator-prey model, Rosenzweig-MacArthur predator-prey model, Kermack-McKendrick SIR model, and a meta-population SIRS model. By Mario Pineda-Krch.

HardyWeinberg Exploration of bi-allelic marker data. Focuses on the graphical representation of the results of tests for Hardy-Weinberg equilibrium in a ternary plot. Routines for several tests for Hardy-Weinberg equilibrium are included. By Jan Graffelman.

HiddenMarkov Hidden Markov Models. Contains functions for the analysis of Discrete Time Hidden Markov Models, Markov Modulated GLMs and the Markov Modulated Poisson Process. Includes functions for simulation, parameter estimation, and the Viterbi algorithm. The algorithms are based of those of Walter Zucchini. By David Harte.

JADE JADE and ICA performance criteria. The