R Installation and Administration

Version 2.13.1 (2011-07-08)

R Development Core Team

ISBN 3-900051-09-7

Table of Contents

1 Obtaining R

Sources, binaries and documentation for R can be obtained via CRAN, the "Comprehensive R Archive Network" whose current members are listed at http://cran.r-project.org/mirrors.html.

1.1 Getting and unpacking the sources

The simplest way is to download the most recent '`R-x.y.z.tar.gz`' file, and unpack it with

```
tar -xf R-x.y.z.tar.gz
```

on systems that have a suitable[1] `tar` installed. On other systems you need to have the `gzip` program installed, when you can use

```
gzip -dc R-x.y.z.tar.gz | tar -xf -
```

The pathname of the directory into which the sources are unpacked should not contain spaces, as most `make` programs (and specifically GNU `make`) do not expect spaces.

If you want the build to be usable by a group of users, set `umask` before unpacking so that the files will be readable by the target group (e.g., `umask 022` to be usable by all users). Keep this setting of `umask` whilst building and installing.

If you use a recent GNU version of `tar` and do this as a root account (which on Windows includes accounts with administrator privileges) you may see many warnings about changing ownership. In which case you can use

```
tar --no-same-owner -xf R-x.y.z.tar.gz
```

and perhaps also include the option '`--no-same-permissions`'. (These options can also be set in the `TAR_OPTIONS` environment variable: if more than one option is included they should be separated by spaces.)

1.2 Getting patched and development versions

A patched version of the current release, '`r-patched`', and the current development version, '`r-devel`', are available as daily tarballs and via access to the R Subversion repository. (For the two weeks prior to the release of a minor (2.x.0) version, '`r-patched`' tarballs may refer to beta/release candidates of the upcoming release, the patched version of the current release being available via Subversion.)

The tarballs are available from ftp://ftp.stat.math.ethz.ch/pub/Software/R/. Download '`R-patched.tar.gz`' or '`R-devel.tar.gz`' (or the '`.tar.bz2`' versions) and unpack as described in the previous section. They are built in exactly the same way as distributions of R releases.

1.2.1 Using Subversion and rsync

Sources are also available via https://svn.R-project.org/R/, the R Subversion repository. If you have a Subversion client (see http://subversion.apache.org/), you can check out and update the current '`r-devel`' from https://svn.r-project.org/R/trunk/ and the current '`r-patched`' from '`https://svn.r-project.org/R/branches/R-x-y-branch/`' (where *x* and *y* are the major and minor number of the current released version of R). E.g., use

```
svn checkout https://svn.r-project.org/R/trunk/ path
```

to check out '`r-devel`' into directory *path* (which will be created if necessary). The alpha, beta and RC versions of an upcoming *x.y.0* release are available from

[1] e.g. GNU `tar` version 1.15 or later, or that from the '`libarchive`' (as used on OS 10.6) or 'Heirloom Toolchest' distributions.

'`https://svn.r-project.org/R/branches/R-x-y-branch/`' in the four-week period prior to the release.

Note that '`https:`' is required, and that the SSL certificate for the Subversion server of the R project should be recognized as from a trusted source.

Note that retrieving the sources by e.g. `wget -r` or `svn export` from that URL will not work: the Subversion information is needed to build R.

The Subversion repository does not contain the current sources for the recommended packages, which can be obtained by `rsync` or downloaded from CRAN. To use `rsync` to install the appropriate sources for the recommended packages, run `./tools/rsync-recommended` from the top-level of the R sources.

If downloading manually from CRAN, do ensure that you have the correct versions of the recommended packages: if the number in the file '`VERSION`' is '*x.y.z*' you need to download the contents of '`http://CRAN.R-project.org/src/contrib/dir`', where *dir* is '*x.y.z/Recommended*' for r-devel or '*x.y-patched/Recommended*' for r-patched, respectively, to directory '`src/library/Recommended`' in the sources you have unpacked. After downloading manually you need to execute `tools/link-recommended` from the top level of the sources to make the requisite links in '`src/library/Recommended`'. A suitable incantation from the top level of the R sources using `wget` might be (for the correct value of '*dir*')

```
wget -r -l1 --no-parent -A\*.gz -nd -P src/library/Recommended \
  http://CRAN.R-project.org/src/contrib/dir
./tools/link-recommended
```

2 Installing R under Unix-alikes

R will configure and build under most common Unix and Unix-alike platforms including '*cpu-*-linux-gnu*' for the 'alpha', 'arm', 'hppa', 'ix86', 'ia64', 'm68k', 'mips', 'mipsel', 'powerpc', 's390', 'sparc', and 'x86_64' CPUs, 'i386-apple-darwin', 'x86_64-apple-darwin', 'i386-sun-solaris', 'sparc-sun-solaris', 'x86_64-*-freebsd', and 'powerpc-ibm-aix6*' as well as perhaps (it is tested less frequently on these platforms) 'powerpc-apple-darwin', 'i386-*-freebsd', 'i386-*-netbsd' and 'i386-*-openbsd'.

In addition, binary distributions are available for some common Linux distributions and for OS X (formerly Mac OS). See the FAQ for current details. These are installed in platform-specific ways, so for the rest of this chapter we consider only building from the sources.

2.1 Simple compilation

First review the essential and useful tools and libraries in Appendix A [Essential and useful other programs under a Unix-alike], page 31, and install those you want or need. Ensure that the environment variable `TMPDIR` is either unset (and '`/tmp`' exists and can be written in and scripts can be executed from) or points to a valid temporary directory (one from which execution of scripts is allowed).

Choose a place to install the R tree (R is not just a binary, but has additional data sets, help files, font metrics etc). Let us call this place *R_HOME*. Untar the source code. This should create directories '`src`', '`doc`', and several more under a top-level directory: change to that top-level directory (At this point North American readers should consult Section B.3.1 [Setting paper size], page 40.) Issue the following commands:

```
./configure
make
```

(See Section B.5 [Using make], page 41 if your make is not called '`make`'.)

Then check the built system works correctly by

```
make check
```

Failures are not necessarily problems as they might be caused by missing functionality,[1] but you should look carefully at any reported discrepancies. (Some non-fatal errors are expected in locales that do not support Latin-1, in particular in true `C` locales and non-UTF-8 non-Western-European locales.) A failure in '`tests/ok-errors.R`' may indicate inadequate resource limits (see Chapter 5 [Running R], page 17).

More comprehensive testing can be done by

```
make check-devel
```

or

```
make check-all
```

see file '`tests/README`'.

If the command `configure` and `make` commands execute successfully, a shell-script front-end called '`R`' will be created and copied to '*R_HOME/bin*'. You can copy this script to a place where users can invoke it, for example to '`/usr/local/bin/R`'. You could also copy the man page '`R.1`' to a place where your `man` reader finds it, such as '`/usr/local/man/man1`'. If you want to install the complete R tree to, e.g., '`/usr/local/lib/R`', see Section 2.4 [Installation], page 5. Note: you do not *need* to install R: you can run it from where it was built.

You do not necessarily have to build R in the top-level source directory (say, '*TOP_SRCDIR*'). To build in '*BUILDDIR*', run

[1] for example, if you configured R with '`--without-recommended`'.

```
cd BUILDDIR
TOP_SRCDIR/configure
make
```

and so on, as described further below. This has the advantage of always keeping your source tree clean and is particularly recommended when you work with a version of R from Subversion. (You may need GNU `make` to allow this, and the pathname of the build directory should not contain spaces.)

Now `rehash` if necessary, type *`R`*, and read the R manuals and the R FAQ (files '`FAQ`' or '`doc/manual/R-FAQ.html`', or http://CRAN.R-project.org/doc/FAQ/R-FAQ.html which always has the version for the latest release of R).

2.2 Help options

By default HTML help pages are created when needed rather than being built at install time.

If you need to disable the server and want HTML help, there is the option to build HTML pages when packages are installed (including those installed with R). This is enabled by the `configure` option '`--enable-prebuilt-html`'. Whether `R CMD INSTALL` (and hence `install.packages`) pre-builds HTML pages is determined by looking at the R installation and is reported by `R CMD INSTALL --help`: it can be overridden by specifying one of the `INSTALL` options '`--html`' or '`--no-html`'.

The server is disabled by setting the environment variable `R_DISABLE_HTTPD` to a non-empty value, either before R is started or within the R session before HTML help (including `help.start`) is used. It is also possible that system security measures will prevent the server from being started, for example if the loopback interface has been disabled. See `?tools::startDynamicHelp` for more details.

2.3 Making the manuals

There is a set of manuals that can be built from the sources,

'`refman`' Printed versions of the help pages for the base packages.

'`fullrefman`'
Printed versions of all the help pages for base and recommended packages (over 3200 pages).

'`R-FAQ`' R FAQ

'`R-intro`' "An Introduction to R".

'`R-data`' "R Data Import/Export".

'`R-admin`' "R Installation and Administration", this manual.

'`R-exts`' "Writing R Extensions".

'`R-lang`' "The R Language Definition".

To make these (except '`fullrefman`'), use

```
make dvi        to create DVI versions
make pdf        to create PDF versions
make info       to create info files (not 'refman').
```

You will not be able to build any of these unless you have `makeinfo` version 4.7 or later installed, and for DVI or PDF you must have `texi2dvi` and '`texinfo.tex`' installed (which are part of the GNU **texinfo** distribution but are, especially '`texinfo.tex`', often made part of the TEX package in re-distributions).

The DVI versions can be previewed and printed using standard programs such as `xdvi` and `dvips`. The PDF versions can be viewed using any recent PDF viewer: they have hyperlinks that can be followed. The info files are suitable for reading online with Emacs or the standalone GNU `info` program. The DVI and PDF versions will be created using the paper size selected at configuration (default ISO a4): this can be overridden by setting `R_PAPERSIZE` on the `make` command line, or setting `R_PAPERSIZE` in the environment and using `make -e`. (If re-making the manuals for a different paper size, you should first delete the file '`doc/manual/version.texi`'.)

There are some issues with making the reference manual, and in particular with the PDF version '`refman.pdf`'. The help files contain both ISO Latin1 characters (e.g. in '`text.Rd`') and upright quotes, neither of which are contained in the standard LaTeX Computer Modern fonts. We have provided four alternatives:

`times` (The default for PDF.) Using standard PostScript fonts. This works well both for on-screen viewing and for printing. One disadvantage is that the Usage and Examples sections may come out rather wide (and R 2.14.0 will have workarounds for this).

Note that in most LaTeX installations this will not actually use the standard fonts for PDF, but rather embed the URW clones NimbusRom, NimbusSans and (for Courier, if used) NimbusMon.

`lm` Using the *Latin Modern* fonts. These are not often installed as part of a TeX distribution, but can obtained from http://www.ctan.org/tex-archive/fonts/ps-type1/lm/ and mirrors. This uses fonts rather similar to Computer Modern, but is not so good on-screen as `times`.

`cm-super` Using type-1 versions of the Computer Modern fonts by Vladimir Volovich. This is a large installation, obtainable from http://www.ctan.org/tex-archive/fonts/ps-type1/cm-super/ and its mirrors. These type-1 fonts have poor hinting and so are nowhere near so readable on-screen as the other three options.

`ae` (The default for DVI.) A package to use composites of Computer Modern fonts. This works well most of the time, and its PDF is more readable on-screen than the previous two options. There are three fonts for which it will need to use bitmapped fonts, '`tctt0900.600pk`', '`tctt1000.600pk`' and '`tcrm1000.600pk`'. Unfortunately, if those files are not available, Acrobat Reader will substitute completely incorrect glyphs so you need to examine the logs carefully.

The default can be overridden by setting the environment variables `R_RD4PDF` and `R_RD4DVI`. (On Unix-alikes, these will be picked up at install time and stored in '`etc/Renviron`', but can still be overridden when the manuals are built.) The default value for `R_RD4PDF` is '`times,hyper`': omit '`hyper`' if you do not want hyperlinks, e.g. for printing. The default for `R_RD4DVI` is '`ae`'.

2.4 Installation

To ensure that the installed tree is usable by the right group of users, set `umask` appropriately (perhaps to '`022`') before unpacking the sources and throughout the build process.

After

```
./configure
make
make check
```

(or, when building outside the source, *`TOP_SRCDIR/configure`*, etc) have been completed successfully, you can install the complete R tree to your system by typing

```
make install
```

A parallel make can be used (but run `make all` first).

This will install to the following directories:

'*prefix/bin*' or '*bindir*'
: the front-end shell script and other scripts and executables

'*prefix/man/man1*' or '*mandir/man1*'
: the man page

'*prefix/LIBnn/R*' or '*libdir/R*'
: all the rest (libraries, on-line help system, ...). Here *LIBnn* is usually '`lib`', but may be '`lib64`' on some 64-bit Linux systems. This is known as the R home directory.

where *prefix* is determined during configuration (typically '`/usr/local`') and can be set by running `configure` with the option '`--prefix`', as in

```
./configure --prefix=/where/you/want/R/to/go
```

This causes `make install` to install the R script to '`/where/you/want/R/to/go/bin`', and so on. The prefix of the installation directories can be seen in the status message that is displayed at the end of `configure`. You can install into another directory tree by using

```
make prefix=/path/to/here install
```

at least with GNU `make` (and current Solaris and FreeBSD `make`, but not some older Unix makes).

More precise control is available at configure time via options: see `configure --help` for details. (However, most of the 'Fine tuning of the installation directories' options are not used by R.)

Configure options '`--bindir`' and '`--mandir`' are supported and govern where a copy of the `R` script and the `man` page are installed.

The configure option '`--libdir`' controls where the main R files are installed: the default is '*eprefix/LIBnn*', where *eprefix* is the prefix used for installing architecture-dependent files, defaults to *prefix*, and can be set via the configure option '`--exec-prefix`'.

Each of `bindir`, `mandir` and `libdir` can also be specified on the `make install` command line (at least for GNU `make`).

The `configure` or `make` variables `rdocdir` and `rsharedir` can be used to install the system-independent '`doc`' and '`share`' directories to somewhere other than `libdir`. The C header files can be installed to the value of `rincludedir`: note that as the headers are not installed into a subdirectory you probably want something like `rincludedir=/usr/local/include/R-2.13.1`.

If you want the R home to be something other than '*libdir/R*', use '`rhome`': for example

```
make install rhome=/usr/local/lib64/R-2.13.0
```

will use a version-specific R home on a Linux 64-bit system.

If you have made R as a shared/dynamic library you can install it in your system's library directory by

```
make prefix=/path/to/here install-libR
```

where `prefix` is optional, and `libdir` will give more precise control.

```
make install-strip
```

will install stripped executables, and on platforms where this is supported, stripped libraries in directories '`lib`' and '`modules`' and in the standard packages.

Note that installing R into a directory which contains spaces is not supported, and at least some aspects (such as installing source packages) will not work.

To install DVI, info and PDF versions of the manuals, use one or more of

```
make install-dvi
make install-info
make install-pdf
```

Once again, it is optional to specify `prefix`, `libdir` or `rhome` (the DVI and PDF manuals are installed under the R home directory). (`make install-info` needs Perl installed if there is no command `install-info` on the system.)

More precise control is possible. For info, the setting used is that of `infodir` (default '*prefix/info*', set by configure option '`--infodir`'). The DVI and PDF files are installed into the R '`doc`' tree, set by the `make` variable `rdocdir`.

A staged installation is possible, that it is installing R into a temporary directory in order to move the installed tree to its final destination. In this case `prefix` (and so on) should reflect the final destination, and `DESTDIR` should be used: see http://www.gnu.org/prep/standards/html_node/DESTDIR.html.

You can optionally install the run-time tests that are part of `make check-all` by

```
make install-tests
```

which populates a '`tests`' directory in the installation.

2.5 Uninstallation

You can uninstall R by

```
make uninstall
```

optionally specifying `prefix` etc in the same way as specified for installation.

This will also uninstall any installed manuals. There are specific targets to uninstall DVI, info and PDF manuals in file '`doc/manual/Makefile`'.

Target `uninstall-tests` will uninstall any installed tests, as well as removing the directory '`tests`' containing the test results.

2.6 Sub-architectures

Some platforms can support closely related builds of R which can share all but the executables and dynamic objects. Examples include builds under Solaris for different chips (in particular, 32- and 64-bit builds), 64- and 32- bit builds on '`x86_64`' Linux and different CPUs (e.g. '`ppc`', '`i386`' and '`x86_64`') under (Mac) OS X >= 10.4.

R supports the idea of architecture-specific builds, specified by adding '`r_arch=`*name*' to the `configure` line. Here *name* can be anything non-empty, and is used to name subdirectories of '`lib`', '`etc`', '`include`' and the package '`libs`' subdirectories. Example names from other systems are the use of '`sparcv9`' on Sparc Solaris and '`32`' by `gcc` on '`x86_64`' Linux.

If you have two or more such builds you can install them over each other (and for 32/64-bit builds on one architecture, one build can be done without '`r_arch`'). The space savings can be considerable: on '`x86_64`' Linux a basic install (without debugging symbols) took 63Mb, and adding a 32-bit build added 6Mb. If you have installed multiple builds you can select which build to run by

```
R --arch=name
```

and just running '`R`' will run the last build that was installed.

`R CMD INSTALL` will detect if more than one build is installed and try to install packages with the appropriate library objects for each. This will not be done if the package has an executable `configure` script or a '`src/Makefile`' file. In such cases you can install for extra builds by

```
R --arch=name CMD INSTALL --libs-only pkg(s)
```

If you want to mix sub-architectures compiled on different platforms (for example 'x86_64' Linux and 'i686' Linux), it is wise to use explicit names for each, and you may also need to set 'libdir' to ensure that they install into the same place.

When sub-architectures are used the version of Rscript in e.g. '/usr/bin' will be the last installed, but architecture-specific versions will be available in e.g. '/usr/lib64/R/bin/exec${*R_ARCH*}'. Normally all installed architectures will run on the platform so the architecture of Rscript itself does not matter. The executable Rscript will run the R script, and at that time the setting of the R_ARCH environment variable determines the architecture which is run.

When running post-install tests with sub-architectures, use

```
R --arch=name CMD make check[-devel|all]
```

to select a sub-architecture to check.

Sub-architectures are also used on Windows, but by selecting executables within the appropriate 'bin' directory, '*R_HOME/bin/i386*' or '*R_HOME/bin/x64*'. For backwards compatibility with R < 2.12.0, there are executables '*R_HOME/bin/R.exe*' or '*R_HOME/bin/Rscript.exe*': these will run an executable from one of the subdirectories, which one being taken first from the R_ARCH environment variable, then from the '--arch' command-line option[2] and finally from the installation default (which is 32-bit for a combined 32/64 bit R installation).

2.6.1 Multilib

On Linux, there is an alternative mechanism for mixing 32-bit and 64-bit libraries known as *multilib*. If a Linux distribution supports multilib, then parallel builds of R may be installed in the sub-directories 'lib' (32-bit) and 'lib64' (64-bit). The build to be run may then be chosen using the setarch command. For example, a 32-bit build may be chosen by

```
setarch i686 R
```

The setarch command is only operational if both 32-bit and 64-bit builds are installed. If there is only one installation of R, then this will always be run regardless of the architecture specified by the setarch command.

There can be problems with installing packages on the non-native architecture. It is a good idea to run e.g. setarch i686 R for sessions in which packages are to be installed, even if that is the only version of R installed (since this tells the package installation code the architecture needed).

At present there is a potential problem with packages using Java, as the post-install for a 'i386' RPM on 'x86_64' Linux reconfigures Java and will find the 'x86_64' Java. If you know where a 32-bit Java is installed you may be able to run (as root)

```
export JAVA_HOME=<path to jre directory of 32-bit Java>
setarch i686 R CMD javareconf
```

to get a suitable setting.

When this mechanism is used, the version of Rscript in e.g. '/usr/bin' will be the last installed, but an architecture-specific version will be available in e.g. '/usr/lib64/R/bin'. Normally all installed architectures will run on the platform so the architecture of Rscript does not matter.

2.7 Testing an Installation

Full testing is possible only if the test files have been installed with

[2] with possible values 'i386', 'x64', '32' and '64'.

```
make install-tests
```

which populates a 'tests' directory in the installation.

If this has been done, two testing routes are available. The first is to move to the home directory of the R installation (as given by `R.home()`) and run

```
cd tests
## followed by one of
../bin/R CMD make check
../bin/R CMD make check-devel
../bin/R CMD make check-all
```

and other useful targets are `test-BasePackages` and `test-Recommended` to the run tests of the standard and recommended packages (if installed) respectively.

This re-runs all the tests relevant to the installed R (including for example code in the package vignettes), but not for example the ones checking the example code in the manuals nor making the standalone Rmath library. This can occasionally be useful when the operating environment has been changed, for example by OS updates or by substituting the BLAS (see Section A.3.1.5 [Shared BLAS], page 36).

Alternatively, the installed R can be run, preferably with '`--vanilla`'. Then

```
library("tools")
testInstalledBasic("both")
testInstalledPackages(scope = "base")
testInstalledPackages(scope = "recommended")
```

runs the basic tests and then all the tests on the standard and recommended packages. These tests can be run from anywhere: the basic tests write their results in the '`tests`' folder of the R home directory and run slightly fewer tests than the first approach: in particular they do not test Internet access.

These tests work best if `diff` (in '`Rtools*.exe`' for Windows users) is in the path, and on some systems need the collation locale set manually (the R code tries to do so but it may not be possible to reset it): so if needed try setting the environment variable `LC_COLLATE` to '`C`' before starting R.

It is possible to test the installed packages (but not the package-specific tests) by `testInstalledPackages` even if `make install-tests` was not run.

Note that the results may depend on the language set for times and messages: for maximal similarity to reference results you may want to try setting

```
LANGUAGE=en LC_TIME=C
```

2.8 Byte-compiler

R 2.13.0 introduced a new 'byte' compiler which compiles R code to a 'byte code' representation. This is by default not used, but is available for experimentation. To compile all the base and recommended packages, run

```
make bytecode
```

in the top-level build directory. This will build (or re-build) R in the normal way, then compile the package code (which for the recommended packages entails a complete re-install of the package from the tarball). To update R, set the environment variables

```
R_COMPILE_PKGS=1
R_COMPILER_SUPPRESS_ALL=1
```

before re-running `make`.

It is however recommended that experimentation is done in the R-devel version of R, which has an updated version of the byte-compiler and more facilities to use it.

3 Installing R under Windows

The '`bin/windows`' directory of a CRAN site contains binaries for a base distribution and a large number of add-on packages from CRAN to run on Windows XP or later on ix86 CPUs (including AMD64/Intel64[1] chips and Windows x64).

Your file system must allow long file names (as is likely except perhaps for some network-mounted systems).

Installation is *via* the installer '`R-2.13.1-win.exe`'. Just double-click on the icon and follow the instructions. When installing on a 64-bit version of Windows the options will include 32- or 64-bit versions of R (and the default is to install both). You can uninstall R from the Control Panel or from the (optional) R program group on the Start Menu.

Note that you will be asked to choose a language for installation, and that choice applies to both installation and un-installation but not to running R itself.

See the R Windows FAQ for more details on the binary installer.

3.1 Building from source

R can be built as either a 32-bit or 64-bit application on Windows: to build the 64-bit application you need a 64-bit edition of Windows: such an OS can also be used to build 32-bit R.

The standard installer combines 32-bit and 64-bit builds into a single executable which can then be installed into the same location and share all the files except the '`.exe`' and '`.dll`' files and some configuration files in the '`etc`' directory.

3.1.1 Getting the tools

If you want to build R from the sources, you will first need to collect, install and test an extensive set of tools. See Appendix D [The Windows toolset], page 55 (and perhaps updates in `http://www.murdoch-sutherland.com/Rtools/`) for details.

The '`Rtools*.exe`' executable installer described in Appendix D [The Windows toolset], page 55 also includes some source files in addition to the R source as noted below. You should run it first, to obtain a working `tar` and other necessities. Choose a "Full installation", and install the extra files into your intended R source directory, e.g. '`C:/R`'. The directory name *should not contain spaces*. We will call this directory '*`R_HOME`*' below.

For a 64-bit build you will need a 64-bit toolchain and the appropriate 64-bit version of Tcl/Tk: both of these are contained in '`Rtools213.exe`'.

3.1.2 Getting the source files

You need to collect the following sets of files:

- Get the R source code tarball '`R-2.13.1.tar.gz`' from CRAN. Open a command window (or another shell) at directory *R_HOME*, and run

  ```
  tar -xf R-2.13.1.tar.gz
  ```

  to create the source tree in *R_HOME*. **Beware**: do use `tar` to extract the sources rather than tools such as WinZip that do not understand symbolic links. If you are using an account with administrative privileges you may get a lot of messages which can be suppressed by

  ```
  tar --no-same-owner -xf R-2.13.1.tar.gz
  ```

  or perhaps better, set the environment variable `TAR_OPTIONS` to the value '`--no-same-owner --no-same-permissions`'.

  It is also possible to obtain the source code using Subversion; see Chapter 1 [Obtaining R], page 1 for details.

[1] formerly known as EM64T.

- If you are not using a tarball you need to obtain copies of the recommended packages from CRAN. Put the '`.tar.gz`' files in '*R_HOME/src/library/Recommended*' and run `make link-recommended`. If you have an Internet connection, you can do this automatically by running in '*R_HOME/src/gnuwin32*'

```
make rsync-recommended
```

The following additional items are normally installed by '`Rtools*.exe`'. If instead you choose to do a completely manual build you will also need

- The Tcl/Tk support files are contained in '`Rtools*.exe`'. Please make sure you install the right version: there is a 32-bit version and a 64-bit version.
- You need `libpng`, `jpeg` and `libtiff` sources (available, e.g., from http://www.libpng.org/, http://www.ijg.org and ftp://ftp.remotesensing.org/pub/libtiff/). The earliest versions that have been tested are '`libpng-1.4.5.tar.gz`', '`jpegsrc.v8b.tar.gz`', '`tiff-3.9.4.tar.gz`' (including betas of '`tiff-4.0.0`'). It is also possible to use '`libjpeg-turbo`' from http://sourceforge.net/projects/libjpeg-turbo/files/.

  Working in the directory '*R_HOME/src/gnuwin32/bitmap*', install the `libpng` and `jpeg` sources in sub-directories. The `jpeg` sub-directory for version 8c is named '`jpeg-8c`'; if you use a different version (e.g. '`jpeg-7`' or '`libjpeg-turbo`'), copy file '`src/gnuwin32/MkRules.dist`' to '`src/gnuwin32/MkRules.local`' and edit the definition of `JPEGDIR`: the names of the '`libpng`' and '`libtiff`' directories can also be set there.

  Example:

```
> tar -zxf libpng-1.4.5.tar.gz
> mv libpng-1.4.5 libpng
> tar -zxf jpegsrc.v8c.tar.gz
> tar -zxf tiff-3.9.4.tar.gz
> mv tiff-3.9.4/libtiff .
> rm -rf tiff-3.9.4
```

  (and see the comment about about '`--no-same-owner`'.

3.1.3 Building the core files

Set the environment variable `TMPDIR` to point to a writable directory, with a path specified with forward slashes and no spaces. (The default is '`/tmp`', which may not be useful on Windows.)

You may need to compile under a case-honouring file system: we found that a `samba`-mounted file system (which maps all file names to lower case) did not work.

Open a command window at '*R_HOME/src/gnuwin32*'. Look at '`MkRules.dist`' and if settings need to be altered, copy it to '`MkRules.local`' and edit the settings there. In particular, this is where a 64-bit build is selected. Then run

```
make all recommended
```

and sit back and wait while the basic compile takes place.

Notes:

- We have had reports that earlier versions of anti-virus software locking up the machine, but not for several years. However, aggressive anti-virus checking such as the on-access scanning of Sophos can slow the build down several-fold.
- By default Doug Lea's `malloc` in the file '*R_HOME/src/gnuwin32/malloc.c*' is used for R's internal memory allocations. You can opt out of this by setting `LEA_MALLOC=NO` in '`MkRules.dist`', in which case the `malloc` in '`msvcrt.dll`' is used. This does impose a considerable performance penalty and has not been tested recently.
- You can run a parallel make by e.g.

```
make -j4 all
make -j4 recommended
```

but this is only likely to be worthwhile on a multi-core machine with ample memory, and is not 100% reliable.

3.1.4 Building the bitmap files

The file '*R_HOME/library/grDevices/libs/{i386,x64}Rbitmap.dll*' is not built automatically.

Running `make` in '*R_HOME/src/gnuwin32/bitmap*' or `make bitmapdll` in '*R_HOME/src/gnuwin32*' should build '`Rbitmap.dll`' and install it under '*R_HOME/library/grDevices/libs*'.

3.1.5 Checking the build

You can test a build by running

```
make check
```

The recommended packages can be checked by

```
make check-recommended
```

Other levels of checking are

```
make check-devel
```

for a more thorough check of the R functionality, and

```
make check-all
```

for `check-devel` and `check-recommended`.

3.1.6 Building the manuals

The PDF manuals can be made by

```
make manuals
```

If you want to make the info versions (not including the Reference Manual), use

```
cd ../../doc/manual
make -f Makefile.win info
```

To make DVI versions of the manuals use

```
cd ../../doc/manual
make -f Makefile.win dvi
```

(all assuming you have `pdftex`/`pdflatex` or `tex`/`latex` installed and in your path).

See the Section 2.3 [Making the manuals], page 4 section in the Unix-alike section for setting options such as the paper size.

3.1.7 Building the Inno Setup installer

You need to have the files for a complete R build, including bitmap and Tcl/Tk support and the manuals, as well as the recommended packages and Inno Setup (see Section D.2 [The Inno Setup installer], page 56).

Once everything is set up

```
make distribution
make check-all
```

will make all the pieces and the installers and put them in the '`gnuwin32/cran`' subdirectory, then check the build. This works by building all the parts in the sequence:

`rbuild` (the executables, the FAQ docs etc.)
`rpackage` (the base packages)
`htmldocs` (the HTML documentation)
`bitmapdll` (the bitmap support files)
`cairodevices` (the cairo-based graphics devices)
`recommended` (the recommended packages)
`vignettes` (the vignettes in package **grid**:
only needed if building from an svn checkout)
`manuals` (the PDF manuals)
`rinstaller` (the install program)
`crandir` (the CRAN distribution directory, only for 64-bit builds)

The parts can be made individually if a full build is not needed, but earlier parts must be built before later ones. (The '`Makefile`' doesn't enforce this dependency—some build targets force a lot of computation even if all files are up to date.) The first four targets are the default build if just `make` (or `make all`) is run.

If you want to customize the installation by adding extra packages, replace `make rinstaller` by something like

```
make rinstaller EXTRA_PKGS='pkg1 pkg2 pkg3'
```

An alternative way to customize the installer starting with a binary distribution is to first make a full installation of R from the standard installer (that is, select '`Full Installation`' from the '`Select Components`' screen), then add packages and make other customizations to that installation. Then (after having customized file '`MkRules`', possibly *via* '`MkRules.local`', and having made R in the source tree) in '`src/gnuwin32/installer`' run

```
make myR IMAGEDIR=rootdir
```

where '`rootdir`' is the path to the root of the customized installation (in double quotes if it contains spaces or backslashes).

Both methods create an executable with a standard name such as '`R-2.13.1-win.exe`', so please rename it to indicate that it is customized. If you intend to *distribute* a customized installer please do check that license requirements are met – note that the installer will state that the contents are distributed under GPL-2 and this has a requirement for *you* to supply the complete sources (including the R sources even if you started with a binary distribution of R, and also the sources of packages (including their external software) which are included).

The defaults for the startup parameters may also be customized. For example

```
make myR IMAGEDIR=rootdir MDISDI=1
```

will create an installer that defaults to installing R to run in SDI mode. See '`src/gnuwin32/installer/Makefile`' for the names and values that can be set.

The standard CRAN distribution of a 32/64-bit installer is made by first building 32-bit R (just

```
make 32-bit
```

is needed), and then building 64-bit R with the macro `HOME32` set in file '`MkRules.local`' to the top-level directory of the 32-bit build. Then the `make rinstaller` step copies the files that differ between architectures from the 32-bit build as it builds the installer image.

3.1.8 Building the MSI installer

It is also possible to build an installer for use with Microsoft Installer. This is intended for use by sysadmins doing automated installs, and is not recommended for casual use.

It makes use of the Windows Installer XML (WiX) toolkit *version 3.0 or 3.5* available from http://wix.sourceforge.net/. Once WiX is installed, set the path to its home directory in '`MkRules.local`'.

You need to have the files for a complete R build, including bitmap and Tcl/Tk support and the manuals, as well as the recommended packages. There is no option in the installer to customize startup options, so edit 'etc/Rconsole' and 'etc/Rprofile.site' to set these as required. Then

```
cd installer
make msi
```

which will result in a file with a name like 'R-2.13.1-win32.msi'. This can be double-clicked to be installed, but those who need it will know what to do with it (usually by running msiexec /i with additional options). Properties that users might want to set from the msiexec command line include 'ALLUSERS', 'INSTALLDIR' (something like 'c:\Program Files\R\R-2.13.1') and 'RMENU' (the path to the 'R' folder on the start menu) and 'STARTDIR' (the starting directory for R shortcuts, defaulting to something like 'c:\Users\name\Documents\R').

The MSI installer can be built both from a 32-bit build of R ('R-2.13.1-win32.msi') and from a 64-bit build of R ('R-2.13.1-win64.msi', optionally including 32-bit files by setting the macro HOME32, when the name is 'R-2.13.1-win.msi'). Unlike the main installer, a 64-bit MSI installer can only be run on 64-bit Windows.

Thanks to David del Campo (Dept of Statistics, University of Oxford) for suggesting WiX and building a prototype installer.

3.1.9 Cross-building on Linux

Support for cross-building was withdrawn at R 2.9.0.

3.1.10 64-bit Windows builds

To build a 64-bit version of R you need a 64-bit toolchain: the only one discussed here (see Section D.4.2 [64-bit toolchain], page 57) is based on the work of the MinGW-w64 project (http://sourceforge.net/projects/mingw-w64/, but commercial compilers such as those from Intel and PGI could be used (and have been by R redistributors).

Support for MinGW-w64 was developed in the R sources over the period 2008–10 and was first released as part of R 2.11.0. The assistance of Yu Gong at a crucial step in porting R to MinGW-w64 is gratefully acknowledged, as well as help from Kai Tietz, the lead developer of the MinGW-w64 project.

3.2 Byte-compiler

R 2.13.0 introduced a new 'byte' compiler which compiles R code to a 'byte code' representation. This is by default not used, but is available for experimentation. To compile all the base and recommended packages, first build R from the sources in the normal way, then run

```
make bytecode
```

in the 'src/gnuwin32' directory. To update R, set the environment variables

```
R_COMPILE_PKGS=1
R_COMPILER_SUPPRESS_ALL=1
```

before re-running make.

3.3 Testing an Installation

The Windows installer contains a set of test files used when building R. These are an optional part of the installation, and if it is desired to run tests on the installation these should be selected as well as the help source files (perhaps most easily by doing a 'Full Installation').

The Rtools are not needed to run these tests. but more comprehensive analysis of errors will be given if diff is in the path.

Once this has been done, launch either Rgui or Rterm, preferably with '--vanilla'. Then

```
library("tools")
testInstalledBasic("both")
testInstalledPackages("base")
testInstalledPackages("recommended")
```

runs the basic tests and then all the tests on the standard and recommended packages. These tests can be run from anywhere: they write their results in the '`tests`' folder of the R home directory (as given by `R.home()`).

4 Installing R under (Mac) OS X

The '`bin/macosx`' directory of a CRAN site contains binaries for OS X for a base distribution and a large number of add-on packages from CRAN to run on OS X 10.5/6/7.

The simplest way is to use '`R-2.13.1.pkg`': just double-click on the icon. Note that Tcl/Tk and the Fortran compiler need to be installed separately if needed.

See the R for Mac OS X FAQ for more details.

4.1 Building from source on (Mac) OS X

If you want to build this port from the sources, you should read the R for Mac OS X FAQ for full details. You will need to collect and install some tools as explained in that document. Then you have to unpack the R sources and configure R appropriately, for example

```
tar -zxvf R-2.13.1.tar.gz
cd R-2.13.1
./configure --with-blas='-framework vecLib' --with-lapack \
  --with-aqua --enable-R-framework
make
```

and then sit back and wait. The first two options are the default and use the BLAS and LAPACK built into OS X. This has pros and cons: it used to be faster (but has not been on some recent multicore hardware) but less accurate.

The second line of options are also default on OS X, but needed only if you want to build R for use with `R.app` Console, and imply '`--enable-R-shlib`' to build R as a shared/dynamic library. These options configure R to be built and installed as a framework called '`R.framework`'. The default installation path for '`R.framework`' is '`/Library/Frameworks`' but this can be changed at configure time by specifying the flag '`--enable-R-framework[=DIR]`' or at install time as

```
make prefix=/where/you/want/R.framework/to/go install
```

(the final '`R.framework`' directory should not be included in the path).

For compatibility with the CRAN distribution you may need to specify '`--with-included-gettext`' to avoid linking against a '`libintl`' dynamic library you may have available, for example in '`/usr/local/lib`'.

Note that building the '`R.app`' GUI console is a separate project: see the OS X FAQ for details.

5 Running R

How to start R and what command-line options are available is discussed in Section "Invoking R" in *An Introduction to R*.

You should ensure that the shell has set adequate resource limits: R expects a stack size of at least 8MB and to be able to open at least 256 file descriptors. (Any modern OS will have default limits at least as large as these, but apparently NetBSD does not. Use the shell command `ulimit` (`sh`/`bash`) or `limit` (`csh`/`tcsh`) to check.)

R makes use of a number of environment variables, the default values of many of which are set in file '*`R_HOME/etc/Renviron`*' (there are none set by default on Windows and hence no such file). These are set at `configure` time, and you would not normally want to change them – a possible exception is `R_PAPERSIZE` (see Section B.3.1 [Setting paper size], page 40). The paper size will be deduced from the '`LC_PAPER`' locale category if it exists and `R_PAPERSIZE` is unset, and this will normally produce the right choice from '`a4`' and '`letter`' on modern Unix-alikes (but can always be overridden by setting `R_PAPERSIZE`).

Various environment variables can be set to determine where R creates its per-session temporary directory. The environment variables `TMPDIR`, `TMP` and `TEMP` are searched in turn and the first one which is set and points to a writable area is used. If none do, the final default is '`/tmp`' on Unix-alikes and the value of `R_USER` on Windows.

Some Unix-alike systems are set up to remove files and directories periodically from '`/tmp`', for example by a `cron` job running `tmpwatch`. Set `TMPDIR` to another directory before running long-running jobs on such a system.

Note that `TMPDIR` will be used to execute `configure` scripts when installing packages, so if '`/tmp`' has been mounted as '`noexec`', `TMPDIR` needs to be set to a directory from which execution is allowed.

6 Add-on packages

It is helpful to use the correct terminology. A *package* is loaded from a *library* by the function `library()`. Thus a library is a directory containing installed packages; the main library is '*R_HOME/library*', but others can be used, for example by setting the environment variable `R_LIBS` or using the R function `.libPaths()`.

6.1 Default packages

The set of packages loaded on startup is by default

```
> getOption("defaultPackages")
[1] "datasets"  "utils"     "grDevices" "graphics"  "stats"     "methods"
```

(plus, of course, **base**) and this can be changed by setting the option in startup code (e.g. in '`~/.Rprofile`'). It is initially set to the value of the environment variable `R_DEFAULT_PACKAGES` if set (as a comma-separated list). Setting `R_DEFAULT_PACKAGES=NULL` ensures that only package **base** is loaded.

Changing the set of default packages is normally used to reduce the set for speed when scripting: in particular not using **methods** will reduce the start-up time by a factor of up to two (and this is done by `Rscript`). But it can also be used to customize R, e.g. for class use.

6.2 Managing libraries

R packages are installed into *libraries*, which are directories in the file system containing a subdirectory for each package installed there.

R comes with a single library, '*R_HOME/library*' which is the value of the R object '`.Library`' containing the standard and recommended[1] packages. Both sites and users can create others and make use of them (or not) in an R session. At the lowest level '`.libPaths()`' can be used to add paths to the collection of libraries or to report the current collection.

R will automatically make use of a site-specific library '*R_HOME/site-library*' if this exists (it does not in a vanilla R installation). This location can be overridden by setting[2] '`.Library.site`' in '*R_HOME/etc/Rprofile.site*', or (not recommended) by setting the environment variable `R_LIBS_SITE`. Like '`.Library`', the site libraries are always included by '`.libPaths()`'.

Users can have one or more libraries, normally specified by the environment variable `R_LIBS_USER`. This has a default value (use '`Sys.getenv("R_LIBS_USER")`' within an R session to see what it is), but only is used if the corresponding directory actually exists (which by default it will not).

Both `R_LIBS_USER` and `R_LIBS_SITE` can specify multiple library paths, separated by colons (semicolons on Windows).

6.3 Installing packages

Packages may be distributed in source form or compiled binary form. Installing source packages which contain C/C++/Fortran code requires that compilers and related tools be installed. Binary packages are platform-specific and generally need no special tools to install, but see the documentation for your platform for details.

Note that you may need to specify implicitly or explicitly the library to which the package is to be installed. This is only an issue if you have more than one library, of course.

[1] unless they were excluded in the build.

[2] its binding is locked once that files has been read, so users cannot easily change it.

If installing packages on a Unix-alike to be used by other users, ensure that the system `umask` is set to give sufficient permissions (see also `Sys.umask` in R). (To a large extent this is unnecessary in recent versions of R, which install packages as if `umask = 022`.)

For most users it suffices to call '`install.packages(pkgname)`' or its GUI equivalent if the intention is to install a CRAN package and internet access is available.[3] On most systems '`install.packages()`' will allow packages to be selected from a list box.

To install packages from source in a Unix-alike use

```
R CMD INSTALL -l /path/to/library pkg1 pkg2 ...
```

The part '`-l /path/to/library`' can be omitted, in which case the first library of a normal R session is used (that shown by `.libPaths()[1]`).

Ensure that the environment variable `TMPDIR` is either unset (and '`/tmp`' exists and can be written in and executed from) or points to a valid temporary directory.

There are a number of options available: use `R CMD INSTALL --help` to see the current list.

Alternatively, packages can be downloaded and installed from within R. First set the option `CRAN` to your nearest CRAN mirror using `chooseCRANmirror()`. Then download and install packages **pkg1** and **pkg2** by

```
> install.packages(c("pkg1", "pkg2"))
```

The essential dependencies of the specified packages will also be fetched. Unless the library is specified (argument `lib`) the first library in the library search path is used: if this is not writable, R will ask the user (in an interactive session) if the default user library should be created, and if allowed to will install the packages there.

If you want to fetch a package and all those it depends on (in any way) that are not already installed, use e.g.

```
> install.packages("Rcmdr", dependencies = TRUE)
```

`install.packages` can install a source package from a local '`.tar.gz`' file by setting argument `repos` to `NULL`: this will be selected automatically if the name given is a single '`.tar.gz`' file.

`install.packages` can look in several repositories, specified as a character vector by the argument `repos`: these can include a CRAN mirror, Bioconductor, Omegahat, R-forge, local archives, local files, ...). Function `setRepositories()` can select amongst those repositories that the R installation is aware of.

Naive users sometimes forget that as well as installing a package, they have to use `library` to make its functionality available.

6.3.1 Windows

What `install.packages` does by default is different on Unix-alikes (except OS X) and Windows. On Unix-alikes it consults the list of available *source* packages on CRAN (or other repository/ies), downloads the latest version of the package sources, and installs them (via `R CMD INSTALL`). On Windows it looks (by default) at the list of *binary* versions of packages available for your version of R and downloads the latest versions (if any), although optionally it will also download and install a source package by setting the `type` argument.

On Windows `install.packages` can also install a binary package from a local '`zip`' file by setting argument `repos` to `NULL`. `Rgui.exe` has a menu `Packages` with a GUI interface to `install.packages`, `update.packages` and `library`.

Windows binary packages for R are nowadays distributed as a single binary containing either or both architectures.

[3] If a proxy needs to be set, see `?download.file`.

A few of the binary packages need other software to be installed on your system: see for example http://cran.r-project.org/bin/windows/contrib/2.13/@ReadMe. For 64-bit builds, packages using Gtk+ (**Cairo**, **RGtk2**, **cairoDevice** and those that depend on them) need the '`bin`' directory of a bundled distribution from http://www.gtk.org/download-windows-64bit.html in the path: note that this conflicts with the directory of DLLs needed for 32-bit Gtk+ so you need the correct one first in your path if using both. One way to achieve this is to set the `PATH` environment variable in files '`R_HOME/etc/i386/Renviron.site`' and '`R_HOME/etc/x64/Renviron.site`'.

`R CMD INSTALL` works in Windows to install source packages if you have set up the tools needed (see Appendix D [The Windows toolset], page 55). No additional tools are needed if the package does not contain compiled code, and `install.packages(type="source")` will work for such packages (and for those with compiled code if the tools are in the path).

We have seen occasional permission problems after unpacking source packages on some Vista/Windows 7/Server 2008 systems: these have been circumvented by setting the environment variable `R_INSTALL_TAR` to '`tar.exe`'.

If you have only a source package that is known to work with current R and just want a binary Windows build of it, you could make use of the building service offered at http://win-builder.r-project.org/.

For almost all packages `R CMD INSTALL` will attempt to install both 32- and 64-bit builds of a package if run from a 32/64-bit install of R on a 64-bit version of Windows. It will report success if the installation of the architecture of the running `R` succeeded, whether or not the other architecture was successfully installed.

The exceptions are packages with a non-empty '`configure.win`' script or which make use of '`src/Makefile.win`'. If '`configure.win`' does something appropriate to both architectures use[4] option '`--force-biarch`': otherwise `R CMD INSTALL --merge-multiarch` can be applied to a source tarball to merge separate 32- and 64-bit installs. (This can only be applied to a tarball, and will only succeed if both installs succeed.)

If you have a package without compiled code and no Windows-specific help, you can zip up an installation on another OS and install from the that zip file on Windows. However, such a package can be installed from the sources on Windows without any additional tools.

6.3.2 OS X

On OS X `install.packages` works as it does on other Unix-alike systems, but there are additional types `mac.binary*` (the default in the CRAN distribution) that can be passed to `install.packages` in order to download and install binary packages from a suitable repository, and type `mac.binary.leopard` is the default for CRAN builds of R. These OS X binary package files have the extension '`tgz`'. The R.app GUI provides for installation of either binary or source packages, from CRAN or local files.

6.3.3 Customizing package compilation

The R system and package-specific compilation flags can be overridden or added to by setting the appropriate Make variables in the personal file '*`HOME/.R/Makevars-R_PLATFORM`*' (but '*`HOME/.R/Makevars.win`*' or '*`HOME/.R/Makevars.win64`*' on Windows), or if that does not exist, '*`HOME/.R/Makevars`*', where '`R_PLATFORM`' is the platform for which R was built, as available in the `platform` component of the R variable `R.version`.

Package developers are encouraged to use this mechanism to enable a reasonable amount of diagnostic messaging ("warnings") when compiling, such as e.g. '`-Wall -pedantic`' for tools from GCC, the Gnu Compiler Collection.

[4] for a small number of CRAN packages where this is known to be safe and is needed by the autobuilder this is the default. Look at the source of '`tools:::.install_packages`' for the list.

Note that this mechanism can also be used when it necessary to change the optimization level for a particular package. For example

```
## for C code
CFLAGS=-g -O
## for C++ code
CXXFLAGS=-g -O
## for Fortran code
FFLAGS=-g -O
## for Fortran 95 code
FCFLAGS=-g -O
```

There is also provision for a site-wide '`Makevars.site`' file under '*`R_HOME/etc`*' (in a sub-architecture-specific directory if appropriate). This is read immediately after '`Makeconf`'.

6.3.4 Multiple sub-architectures

When installing packages from their sources, there are some extra considerations on installations which use sub-architectures. These are commonly used on (Mac) OS X and Windows, but can in principle be used on other platforms.

When a source package is installed by a build of R which supports multiple sub-architectures, the normal installation process installs the packages for all sub-architectures, but only tests it can be loaded under the current sub-architecture. The exceptions are

Unix-alikes
: where there is an '`configure`' script, or a file '`src/Makefile`'.

Windows
: where there is a non-empty '`configure.win`' script, or a file '`src/Makefile.win`' (with some exceptions where the package is known to have an architecture-independent '`configure.win`', or if '`--force-biarch`' is used to assert so).

In those cases only the current architecture is installed. Further sub-architectures can be installed by

```
R CMD INSTALL --libs-only pkg
```

using the path to `R` or `R --arch` to select the additional sub-architecture. On Windows there is also `R CMD INSTALL --merge-multiarch` to build and merge the two architectures, starting with a source tarball.

6.4 Updating packages

The command `update.packages()` is the simplest way to ensure that all the packages on your system are up to date. Set the `repos` argument as in the previous section. The `update.packages()` downloads the list of available packages and their current versions, compares it with those installed and offers to fetch and install any that have later versions on the repositories.

An alternative interface to keeping packages up-to-date is provided by the command `packageStatus()`, which returns an object with information on all installed packages and packages available at multiple repositories. The `print` and `summary` methods give an overview of installed and available packages, the `upgrade` method offers to fetch and install the latest versions of outdated packages.

One sometimes-useful additional piece of information that `packageStatus()` returns is the status of a package, as `"ok"`, `"upgrade"` or `"unavailable"` (in the currently selected repositories). For example

```
> inst <- packageStatus()$inst
> with(inst, inst[Status != "ok", c(1, 3, 14)])
                  Package Version      Status
Biobase           Biobase   2.8.0 unavailable
RCurl               RCurl   1.4-2     upgrade
Rgraphviz       Rgraphviz  1.26.0 unavailable
rgdal               rgdal  0.6-27     upgrade
```

6.5 Removing packages

Packages can be removed in a number of ways. From a command prompt they can be removed by

```
R CMD REMOVE -l /path/to/library pkg1 pkg2 ...
```

From a running R process they can be removed by

```
> remove.packages(c("pkg1", "pkg2"),
                  lib = file.path("path", "to", "library"))
```

Finally, in most installations one can just remove the package directory from the library.

6.6 Setting up a package repository

Utilities such as `install.packages` can be pointed at any CRAN-style repository, and R users may want to set up their own. The 'base' of a repository is a URL such as http://www.omegahat.org/R/: this must be an URL scheme that `download.packages` supports (which also includes '`ftp://`' and '`file://`', but not on most systems '`https://`'). Under that base URL there should be directory trees for one or more of the following types of package distributions:

- `"source"`: located at '`src/contrib`' and containing '`.tar.gz`' files. Other forms of compression can be used, e.g. '`.tar.bz2`' or '`.tar.xz`' files.
- `"win.binary"`: located at '`bin/windows/contrib/`*x.y*' for R versions *x.y.z* and containing '`.zip`' files for Windows.
- `"mac.binary.leopard"`: located at '`bin/macosx/leopard/contrib/`*x.y*' for R versions *x.y.z* and containing '`.tgz`' files.

Each terminal directory must also contain a '`PACKAGES`' file. This can be a concatenation of the '`DESCRIPTION`' files of the packages separated by blank lines, but only a few of the fields are needed. The simplest way to set up such a file is to use function `write_PACKAGES` in the **tools** package, and its help explains which fields are needed. Optionally there can also be a '`PACKAGES.gz`' file, a `gzip`-compressed version of '`PACKAGES`'—as this will be downloaded in preference to '`PACKAGES`' it should be included for large repositories.

To add your repository to the list offered by `setRepositories()`, see the help file for that function.

A repository can contain subdirectories, when the descriptions in the '`PACKAGES`' file of packages in subdirectories must include a line of the form

```
Path: path/to/subdirectory
```

—once again `write_PACKAGES` is the simplest way to set this up.

6.7 Checking installed source packages

It can be convenient to run `R CMD check` on an installed package, particularly on a platform which uses sub-architectures. The outline of how to do this is, with the source package in directory '*pkgname*' (or a tarball filename):

```
R CMD INSTALL -l libdir pkgname > pkgname.log 2>&1
R CMD check -l libdir --install=check:pkgname.log pkgname
```

Where sub-architectures are in use the `R CMD check` line can be repeated with additional architectures by

```
R --arch arch CMD check -l libdir --extra-arch --install=check:pkgname.log pkgname
```

where '`--extra-arch`' selects only those checks which depend on the installed code and not those which analyse the sources. (If multiple sub-architectures fail only because they need different settings, e.g. environment variables, '`--no-multiarch`' may need to be added to the `INSTALL` lines.) On (Mac) OS X and other Unix-alikes the architecture to run is selected by '`--arch`': this can also be used on Windows with '*R_HOME/bin/R.exe*', but it is more usual to select the path to the `Rcmd.exe` of the desired architecture.

So on Windows to install, check and package for distribution a source package from a tarball which has been tested on another platform one might use

```
.../bin/i386/Rcmd INSTALL -l libdir tarball --build > pkgname.log 2>&1
.../bin/i386/Rcmd check -l libdir --extra-arch --install=check:pkgname.log pkgname
.../bin/x64/Rcmd check -l libdir --extra-arch --install=check:pkgname.log pkgname
```

where one might want to run the second and third lines in a different shell with different settings for environment variables and the path (to find external software, notably for Gtk+).

`R CMD INSTALL` can do a `i386` install and then add the `x64` DLL by

```
R CMD INSTALL --merge-multiarch -l libdir tarball
```

and '`--build`' can be added to zip up the installation.

7 Internationalization and Localization

Internationalization refers to the process of enabling support for many human languages, and *localization* to adapting to a specific country and language.

Historically R worked in the ISO Latin-1 8-bit character set and so covered English and most Western European languages (if not necessarily their currency symbols). Since R 2.1.0 it has supported multi-byte character sets such as UTF-8 and others used for Chinese, Japanese and Korean.

Current builds of R support all the character sets that the underlying OS can handle. These are interpreted according to the current `locale`, a sufficiently complicated topic to merit a separate section. Note though that R has no built-in support for right-to-left languages and bidirectional output, relying on the OS services. For example, how character vectors in UTF-8 containing both English digits and Hebrew characters are printed is OS-dependent (and perhaps locale-dependent).

The other aspect of the internationalization is support for the translation of messages. This is enabled in almost all builds of R.

7.1 Locales

A *locale* is a description of the local environment of the user, including the preferred language, the encoding of characters, the currency used and its conventions, and so on. Aspects of the locale are accessed by the R functions `Sys.getlocale` and `Sys.localeconv`.

The system of naming locales is OS-specific. There is quite wide agreement on schemes, but not on the details of their implementation. A locale needs to specify

- A human language. These are generally specified by a lower-case two-character abbreviation following ISO 639 (see e.g. http://en.wikipedia.org/wiki/ISO_639-1).
- A 'territory', used mainly to specify the currency. These are generally specified by an upper-case two-character abbreviation following ISO 3166 (see e.g. http://en.wikipedia.org/wiki/ISO_3166). Sometimes the combination of language and territory is used to specify the encoding, for example to distinguish between traditional and simplified Chinese.
- A charset encoding, which determines both how a byte stream should be divided into characters, and which characters the subsequences of bytes represent.
- Optionally, a modifier, for example to indicate that Austria is to be considered pre- or post-Euro. The modifier is also used to indicate the script (`@latin`, `@cyrillic`, `@iqtelif`) or language dialect (e.g. `saaho`, a dialect of Afar, and `bokmal` and `nynorsk`, dialects of Norwegian regarded by some OSes are separate languages, `no` and `nn`).

R is principally concerned with the first (for translations) and third. Note that the charset may be deducible from the language, as some OSes offer only one charset per language, and most OSes have only one charset each for many languages. Note too the remark above about Chinese.

7.1.1 Locales under Linux

Modern Linux uses the XPG[1] locale specifications which have the form '`en_GB`', '`en_GB.UTF-8`', '`aa_ER.UTF-8@saaho`', '`de_AT.iso885915@euro`', the components being in the order listed above. (See `man locale` and `locale -a` for more details.) Similar schemes are used by most Unix-alikes: some (including some distributions of Linux) use '`.utf8`' rather than '`.UTF-8`'.

[1] 'X/Open Portability Guide', which has had several versions.

7.1.2 Locales under Windows

Windows also uses locales, but specified in a rather less concise way. Most users will encounter locales only via drop-down menus, but more information and lists can be found at http://msdn.microsoft.com/library/default.asp?url=/library/en-us/vccore98/html/_crt_language_and_country_strings.asp.

It offers only one encoding per language.

Some care is needed with Windows' locale names. For example, `chinese` is Traditional Chinese and not Simplified Chinese as used in most of the Chinese-speaking world.

7.1.3 Locales under OS X

OS X supports locales in its own particular way, but the R GUI tries to make this easier for users. See http://developer.apple.com/documentation/MacOSX/Conceptual/BPInternational/ for how users can set their locales. As with Windows, end users will generally only see lists of languages/territories. Users of R in a terminal may need to set the locale to something like '`en_GB.UTF-8`' if it defaults to '`C`' (as it often does when logging it remotely and in batch jobs).

Internally OS X uses a form similar to Linux. It is based on ICU locales IDs (http://userguide.icu-project.org/locale) and not XPG ones, but utilities such as `Sys.setlocale()` do now normally accept XPG forms. So there are locales like `de_AT.ISO8859-15` (German in Austria in Latin-9, which covers the Euro): the main difference from other Unix-alikes is that where a character set is not specified it is assumed to be `UTF-8`.

7.2 Localization of messages

The preferred language for messages is by default taken from the locale. This can be overridden first by the setting of the environment variable `LANGUAGE` and then[2] by the environment variables `LC_ALL`, `LC_MESSAGES` and `LANG`. (The last three are normally used to set the locale and so should not be needed, but the first is only used to select the language for messages.) The code tries hard to map locales to languages, but on some systems (notably Windows) the locale names needed for the environment variable `LC_ALL` do not all correspond to XPG language names and so `LANGUAGE` may need to be set. (One example is '`LC_ALL=es`' on Windows which sets the locale to Estonian and the language to Spanish.)

It is usually possible to change the language once R is running *via* (not Windows) `Sys.setlocale("LC_MESSAGES", "new_locale")`, or by setting an environment variable such as `LANGUAGE`, *provided*[3] the language you are changing to can be output in the current character set. But this is OS-specific, and has been known to stop working on an OS upgrade.

Messages are divided into *domains*, and translations may be available for some or all messages in a domain. R makes use of the following domains.

- Domain `R` for basic C-level error messages.
- Domain `R-`*`pkg`* for the R `stop`, `warning` and `message` messages in each package, including `R-base` for the **base** package.
- Domain *`pkg`* for the C-level messages in each package.
- Domain `RGui` for the menus etc of the R for Windows GUI front-end.

Dividing up the messages in this way allows R to be extensible: as packages are loaded, their message translation catalogues can be loaded too.

[2] On some systems setting `LC_ALL` or `LC_MESSAGES` to '`C`' disables `LANGUAGE`.

[3] If you try changing from French to Russian except in a UTF-8 locale, you will most likely find messages change to English.

Translations are looked for by domain according to the currently specified language, as specifically as possible, so for example an Austrian ('`de_AT`') translation catalogue will be used in preference to a generic German one ('`de`') for an Austrian user. However, if a specific translation catalogue exists but does not contain a translation, the less specific catalogues are consulted. For example, R has catalogues for '`en_GB`' that translate the Americanisms (e.g., '`gray`') in the standard messages into English.[4] Two other examples: there are catalogues for '`es`', which is Spanish as written in Spain and these will by default also be used in Spanish-speaking Latin American countries, and also for '`pt_BR`', which are used for Brazilian locales but not for locales specifying Portugal.

Translations in the right language but the wrong charset be made use of by on-the-fly re-encoding. The `LANGUAGE` variable (only) can be a colon-separated list, for example '`se:de`', giving a set of languages in decreasing order of preference. One special value is '`en@quot`', which can be used in a UTF-8 locale to have American error messages with pairs of quotes translated to Unicode directional quotes.

If no suitable translation catalogue is found or a particular message is not translated in any suitable catalogue, 'English'[5] is used.

See `http://developer.r-project.org/Translations.html` for how to prepare and install translation catalogues.

[4] the language written in England: some people living in the USA appropriate this name for their language.

[5] with Americanisms.

8 Choosing between 32- and 64-bit builds

Many current CPUs have both 32- and 64-bit sets of instructions: this has long been true for UltraSparc and more recently for MIPS, PPC and '`x86_64`' (sometimes known as '`amd64`' and '`Intel64`' and earlier as '`EM64T`': practically all current '`ix86`' CPUs support this set of instructions). Most OSes running on such CPUs offer the choice of building a 32-bit or a 64-bit version of R (and details are given below under specific OSes). For most a 32-bit version is the default, but for some (e.g., '`x86_64`' Linux and (Mac) OS X >= 10.6) 64-bit is.

All current versions of R use 32-bit integers and IEC 60559[1] double-precision reals, and so compute to the same precision[2] and with the same limits on the sizes of numerical quantities. The principal difference is in the size of the pointers.

64-bit builds have both advantages and disadvantages:

- The total virtual memory space made available to a 32-bit process is limited by the pointer size to 4GB, and on most OSes to 3GB (or even 2GB). The limits for 64-bit processes are much larger (e.g. 8–128TB).

  R allocates memory for large objects as needed, and removes any unused ones at garbage collection. When the sizes of objects become an appreciable fraction of the address limit, fragmentation of the address space becomes an issue and there may be no hole available that is the size requested. This can cause more frequent garbage collection or the inability to allocate large objects. As a guide, this will become an issue with objects more than 10% of the size of the address space (around 300Mb) or when the total size of objects in use is around one third (around 1Gb).
- Most 32-bit OSes by default limit file sizes to 2GB (and this may also apply to 32-bit builds on 64-bit OSes). This can often be worked around: and `configure` selects suitable defines if this is possible. (We have also largely worked around that limit on 32-bit Windows.) 64-bit builds have much larger limits.
- Because the pointers are larger, R's basic structures are larger. This means that R objects take more space and (usually) more time to manipulate. So 64-bit builds of R will, all other things being equal, run slower than 32-bit builds. (On Sparc Solaris the difference was 15-20%.)
- However, 'other things' may not be equal. In the specific case of '`x86_64`' *vs* '`ix86`', the 64-bit CPU has features (such as SSE2 instructions) which are guaranteed to be present but are optional on the 32-bit CPU, and also has more general-purpose registers. This means that on chips like a desktop Intel Core 2 Duo the vanilla 64-bit version of R is around 10% faster on both Linux and OS X. (Laptop CPUs are relatively slower in 64-bit mode.)

So, for speed you may want to use a 32-bit build, but to handle large datasets (and perhaps large files) a 64-bit build. You can often build both and install them in the same place: See Section 2.6 [Sub-architectures], page 7. (This is done in the OS X and Windows binary distributions.)

Even on 64-bit builds of R there are limits on the size of R objects (see `help("Memory-limits")`, some of which stem from the use of 32-bit integers (especially in FORTRAN code). On all builds of R, the maximum length (number of elements) of a vector is $2^{31} - 1$, about 2 billion, and on 64-bit builds the size of a block of memory allocated is limited to $2^{34} - 1$ bytes (8GB). It is anticipated these will be raised eventually[3] but the need for 8GB objects is (when this was written in 2011) exceptional.

[1] also known as IEEE 754

[2] at least when storing quantities: the on-FPU precision is allowed to vary

[3] this comment has been in the manual since 2005.

9 The standalone Rmath library

The routines supporting the distribution and special[1] functions in R and a few others are declared in C header file '`Rmath.h`'. These can be compiled into a standalone library for linking to other applications. (Note that they are not a separate library when R is built, and the standalone version differs in several ways.)

The makefiles and other sources needed are in directory '`src/nmath/standalone`', so the following instructions assume that is the current working directory (in the build directory tree on a Unix-alike if that is separate from the sources).

'`Rmath.h`' contains '`R_VERSION_STRING`', which is a character string containing the current R version, for example `"2.11.0"`.

There is full access to R's handling of `NaN`, `Inf` and `-Inf` via special versions of the macros and functions

```
ISNAN, R_FINITE, R_log, R_pow and R_pow_di
```

and (extern) constants `R_PosInf`, `R_NegInf` and `NA_REAL`.

There is no support for R's notion of missing values, in particular not for `NA_INTEGER` nor the distinction between `NA` and `NaN` for doubles.

A little care is needed to use the random-number routines. You will need to supply the uniform random number generator

```
double unif_rand(void)
```

or use the one supplied (and with a shared library or DLL you will have to use the one supplied, which is the Marsaglia-multicarry with an entry point

```
set_seed(unsigned int, unsigned int)
```

to set its seeds).

The facilities to change the normal random number generator are available through the constant N01_kind. This takes values from the enumeration type

```
typedef enum {
    BUGGY_KINDERMAN_RAMAGE,
    AHRENS_DIETER,
    BOX_MULLER,
    USER_NORM,
    INVERSION,
    KINDERMAN_RAMAGE
} N01type;
```

(and '`USER_NORM`' is not available).

9.1 Unix-alikes

If R has not already be made in the directory tree, `configure` must be run as described in the main build instructions.

Then (in '`src/nmath/standalone`')

```
make
```

will make standalone libraries '`libRmath.a`' and '`libRmath.so`': '`make static`' and `make shared` will create just one of them.

NB: certain compilers are unable to do compile-time IEEE-754 arithmetic and so cannot compile '`mlutils.c`' and several other files. The known example is old versions of Sun's `cc` (e.g. Forte 6 and 7).

[1] e.g. Bessel, beta and gamma functions

To use the routines in your own C or C++ programs, include

```
#define MATHLIB_STANDALONE
#include <Rmath.h>
```

and link against '`-lRmath`' (and '`-lm`' if needed on your OS). The example file '`test.c`' does nothing useful, but is provided to test the process (via `make test`). Note that you will probably not be able to run it unless you add the directory containing '`libRmath.so`' to the `LD_LIBRARY_PATH` environment variable.

The targets

```
make install
make uninstall
```

will (un)install the header '`Rmath.h`' and shared and static libraries (if built). Both `prefix=` and `DESTDIR` are supported, together with more precise control as described for the main build.

'`make install`' installs a file for `pkg-config` to use by e.g.

```
$(CC) `pkg-config --cflags libRmath` -c test.c
$(CC) `pkg-config --libs libRmath` test.o -o test
```

On some systems '`make install-strip`' will install a stripped shared library.

9.2 Windows

You need to set up almost all the tools to make R and then run (in a Unix-like shell)

```
(cd ../../include; make -f Makefile.win config.h Rconfig.h Rmath.h)
make -f Makefile.win
```

For '`cmd.exe`' use

```
cd ../../include
make -f Makefile.win config.h Rconfig.h Rmath.h
cd ../nmath/standalone
make -f Makefile.win
```

This creates a static library '`libRmath.a`' and a DLL '`Rmath.dll`'. If you want an import library '`libRmath.dll.a`' (you don't need one), use

```
make -f Makefile.win shared implib
```

To use the routines in your own C or C++ programs using MinGW, include

```
#define MATHLIB_STANDALONE
#include <Rmath.h>
```

and link against '`-lRmath`'. This will use the first found of '`libRmath.dll.a`', '`libRmath.a`' and '`Rmath.dll`' in that order, so the result depends on which files are present. You should be able to force static or dynamic linking *via*

```
-Wl,-Bstatic -lRmath -Wl,dynamic
-Wl,-Bdynamic -lRmath
```

or by linking to explicit files (as in the '`test`' target in '`Makefile.win`': this makes two executables, '`test.exe`' which is dynamically linked, and '`test-static.exe`', which is statically linked).

It is possible to link to '`Rmath.dll`' using other compilers, either directly or via an import library: if you make a MinGW import library as above, you will create a file '`Rmath.def`' which can be used (possibly after editing) to create an import library for other systems such as Visual C++.

If you make use of dynamic linking you should use

```
#define MATHLIB_STANDALONE
#define RMATH_DLL
#include <Rmath.h>
```

to ensure that the constants like `NA_REAL` are linked correctly. (Auto-import will probably work with MinGW, but it is better to be sure. This is likely to also work with VC++, Borland and similar compilers.)

Appendix A Essential and useful other programs under a Unix-alike

This appendix gives details of programs you will need to build R on Unix-like platforms, or which will be used by R if found by `configure`.

Remember that some package management systems (such as RPM and deb) make a distinction between the user version of a package and the development version. The latter usually has the same name but with the extension '`-devel`' or '`-dev`': you need both versions installed.

A.1 Essential programs and libraries

You need a means of compiling C and FORTRAN 77 (see Section B.6 [Using FORTRAN], page 41). Some add-on packages also need a C++ compiler. Your C compiler should be IEC 60059[1], POSIX 1003.1 and C99-compliant. R tries to choose suitable flags for the C compilers it knows about, but you may have to set `CC` or `CFLAGS` suitably. For recent versions of `gcc` with `glibc` this means including '`-std=gnu99`'[2]. If the compiler is detected as `gcc`, '`-std=gnu99`' will be appended to `CC` unless it conflicts with a setting of `CFLAGS`. (Note that options essential to run the compiler even for linking, such as those to set the architecture, should be specified as part of `CC` rather than of `CFLAGS`.)

Unless you do not want to view graphs on-screen you need '`X11`' installed, including its headers and client libraries. (You also need '`X11`' installed to use the `jpeg()`, `png()`, `tiff()` and `bmp()` devices.) For recent Fedora distributions it means (at least) RPMs '`libX11`', '`libX11-devel`', '`libXt`' and '`libXt-devel`'. On Debian we recommend the meta-package '`xorg-dev`'. If you really do not want these you will need to explicitly configure R without X11, using '`--with-x=no`'.

The command-line editing (and command completion) depends on the GNU `readline` library: version 4.2 or later is needed for all the features to be enabled. Otherwise you will need to configure with '`--with-readline=no`' (or equivalent).

A suitably comprehensive `iconv` function is essential. The R usage requires `iconv` to be able to translate between `"latin1"` and `"UTF-8"`, to recognize `""` (as the current encoding) and `"ASCII"`, and to translate to and from the Unicode wide-character formats `"UCS-[24][BL]E"` — this is true for `glibc` but not of most commercial Unixes. However, you can make use of GNU `libiconv` (possibly as a plug-in replacement: see http://www.gnu.org/software/libiconv/).

The OS needs to have enough support[3] for wide-character types: this is checked at configuration.

A `tar` program is needed to unpack the sources and packages (including the recommended packages). A version[4] that can automagically detect compressed archives is preferred for use with `untar()`: the configure script looks for `gtar` and `gnutar` before `tar`: use environment variable `TAR` to override this.

There need to be suitable versions of the tools `grep`, `sed` and `tr`: the problems are usually with old AT&T and BSD variants. `configure` will try to find suitable versions (including looking in '`/usr/xpg4/bin`' which is used on some commercial Unixes).

You will not be able to build most of the manuals unless you have `makeinfo` version 4.7 or later installed, and if not some of the HTML manuals will be linked to CRAN. To make DVI or PDF versions of the manuals you will also need file '`texinfo.tex`' installed (which is part of the

[1] also known as IEEE 754

[2] '`-std=c99`' excludes POSIX functionality, but '`config.h`' will turn on all GNU extensions to include the POSIX functionality.

[3] specifically, the C99 functionality of headers '`wchar.h`' and '`wctype.h`', types `wctans_t` and `mbstate_t` and functions `mbrtowc`, `mbstowcs`, `wcrtomb`, `wcscoll`, `wcstombs`, `wctrans`, `wctype`, and `iswctype`.

[4] Such as GNU `tar` 1.15 or later, `bsdtar` (from http://code.google.com/p/libarchive/, as used by FreeBSD and OS 10.6) or `tar` from the Heirloom Toolchest (http://heirloom.sourceforge.net/tools.html).

GNU **texinfo** distribution but is often made part of the TEX package in re-distributions) as well as `texi2dvi`.[5] Further, the versions of `texi2dvi` and '`texinfo.tex`' need to be compatible: we have seen problems with older TEX distributions (TeXLive 2007 and MiKTeX 2.8) used with **texinfo** 4.13.

The DVI and PDF documentation (including '`doc/NEWS.pdf`' and building vignettes needs `tex` and `latex`, or `pdftex` and `pdflatex`. We require LATEX version `2005/12/01` or later (for UTF-8 support). Building PDF package manuals (including the R reference manual) and vignettes is sensitive to the version of the LATEX package **hyperref** and we recommend that the TEX distribution used is keep up-to-date. A number of LATEX packages are required (including **url.sty**, and **listings.sty**) and others such as **hyperref** are desirable.

If you want to build from the R Subversion repository you need both `makeinfo` and `pdflatex`.

The essential programs should be in your `PATH` at the time `configure` is run: this will capture the full paths.

A.2 Useful libraries and programs

The ability to use translated messages makes use of `gettext` and most likely needs GNU `gettext`: you do need this to work with new translations, but otherwise the version contained in the R sources will be used if no suitable external `gettext` is found.

The 'modern' version of the `X11()`, `jpeg()`, `png()` and `tiff()` graphics devices uses the `cairo` and (optionally) `Pango` libraries. Cairo version 1.2.0 or later is required. Pango needs to be at least version 1.10, and 1.12 is the earliest version we have tested. (For Fedora users we believe the `pango-devel` RPM and its dependencies suffice.) R checks for `pkg-config`, and uses that to check first that the '`pangocairo`' package is installed (and if not, '`cairo`') and if additional flags are needed for the '`cairo-xlib`' package, then if suitable code can be compiled. These tests will fail if `pkg-config` is not installed, and are likely to fail if `cairo` was built statically (unusual). Most systems with `Gtk+` 2.8 or later installed will have suitable libraries. OS X comes with none of these libraries, but `cairo` support (without `Pango`) has been added to the binary distribution: `pkg-config` is still needed and can be installed from the sources.

For the best font experience with these devices you need suitable fonts installed: Linux users will want the `urw-fonts` package. On platforms which have it available, the `msttcorefonts` package[6] provides TrueType versions of Monotype fonts such as Arial and Times New Roman. Another useful set of fonts is the 'liberation' truetype fonts available at https://www.redhat.com/promo/fonts/,[7] which cover the Latin, Greek and Cyrillic alphabets plus a fair range of signs. These share metrics with Arial, Times New Roman and Courier New, and contain fonts rather similar to the first two (http://en.wikipedia.org/wiki/Liberation_fonts). Then there is the 'Free UCS Outline Fonts' project (http://www.gnu.org/software/freefont/) which are OpenType/TrueType fonts based on the URW fonts but with extended Unicode coverage. See the R help on `X11` on selecting such fonts.

The bitmapped graphics devices `jpeg()`, `png()` and `tiff()` need the appropriate headers and libraries installed: `jpeg` (version 6b or later, or `libjpeg-turbo`) or `libpng` (version 1.2.7 or later, including 1.4.x and 1.5.x) and `zlib` or `libtiff` (any recent version of 3.x.y – 3.8.2 and 3.9.[124] have been tested) respectively.

[5] `texi2dvi` is normally a shell script. Some versions, e.g. that from **texinfo** 4.13a, need to be run under `bash` rather than a Bourne shell as on, say, Solaris.

[6] also known as `ttf-mscorefonts-installer` in the Debian/Ubuntu world: see also http://en.wikipedia.org/wiki/Core_fonts_for_the_Web.

[7] `ttf-liberation` in Debian/Ubuntu.

If you have them installed (including the appropriate headers and of suitable versions), `zlib`, `libbz2` and PCRE will be used if specified by '`--with-system-zlib`' (version 1.2.3 or later, but beware that bugs have been found with 1.2.4 and 1.2.5, and these should not be used on 32-bit Linux systems with LFS), '`--with-system-bzlib`' or '`--with-system-pcre`': otherwise versions in the R sources will be compiled in. As the latter suffice and are tested with R you should not need to change this.

`liblzma` from `xz-utils` version 4.999 or later (preferably 5.0.0 or later) will be used if installed: the version in the R sources can be selected instead by configuring with '`--with-system-xz=no`'.

Use of the X11 clipboard selection requires the `Xmu` headers and libraries. These are normally part of an X11 installation (e.g. the Debian meta-package '`xorg-dev`'), but some distributions have split this into smaller parts, so for example recent versions of Fedora require the '`libXmu`' and '`libXmu-devel`' RPMs.

Some systems (notably OS X and at least some FreeBSD systems) have inadequate support for collation in multibyte locales. It is possible to replace the OS's collation support by that from ICU (International Components for Unicode, http://site.icu-project.org/), and this provides much more precise control over collation on all systems. ICU is available as sources and as binary distributions for (at least) most Linux distributions, Solaris, FreeBSD and AIX, usually as `libicu` or `icu4c`. It will be used by default where available (including on OS X >= 10.4): should a very old or broken version of ICU be found this can be suppressed by '`--without-ICU`'.

The `bitmap` and `dev2bitmap` devices and also `embedFonts()` use ghostscript (http://www.cs.wisc.edu/~ghost). This should either be in your path when the command is run, or its full path specified by the environment variable `R_GSCMD` at that time.

A.2.1 Tcl/Tk

The **tcltk** package needs Tcl/Tk >= 8.4 installed: the sources are available at http://www.tcl.tk/. To specify the locations of the Tcl/Tk files you may need the configuration options

'`--with-tcltk`'
: use Tcl/Tk, or specify its library directory

'`--with-tcl-config=`*`TCL_CONFIG`*'
: specify location of '`tclConfig.sh`'

'`--with-tk-config=`*`TK_CONFIG`*'
: specify location of '`tkConfig.sh`'

or use the configure variables `TCLTK_LIBS` and `TCLTK_CPPFLAGS` to specify the flags needed for linking against the Tcl and Tk libraries and for finding the '`tcl.h`' and '`tk.h`' headers, respectively. If you have both 32- and 64-bit versions of Tcl/Tk installed, specifying the paths to the correct config files may be necessary to avoid confusion between them.

Versions of Tcl/Tk up to 8.5.9 have been tested (including most versions of 8.4.x, but not recently).

A.2.2 Java support

`configure` looks for Java support on the host system, and if it finds it sets some settings which are useful for Java-using packages. `JAVA_HOME` can be set during the `configure` run to point to a specific JRE/JDK.

Principal amongst these are setting some library paths to the Java libraries and JVM, which are stored in environment variable `R_JAVA_LD_LIBRARY_PATH` in file '*`R_HOME/etc/ldpaths`*' (or a sub-architecture-specific version). A typical setting for '`x86_64`' Linux is

```
JAVA_HOME=/usr/lib/jvm/java-1.6.0-openjdk-1.6.0.0.x86_64/jre
R_JAVA_LD_LIBRARY_PATH=${JAVA_HOME}/lib/amd64/server:${JAVA_HOME}/lib/amd64
```

Note that this unfortunately depends on the exact version of the JRE/JDK installed, and so may need updating if the Java installation is updated. This can be done by running `R CMD javareconf`. The script re-runs Java detection in a manner similar to that of the `configure` script and updates settings in both '`Makeconf`' and '*`R_HOME/etc/ldpaths`*'. See `R CMD javareconf --help` for details.

Another alternative of overriding those setting is to set `R_JAVA_LD_LIBRARY_PATH` (e.g. in '`~/.Renviron`'), or use '`/etc/ld.so.conf`' to specify the Java runtime library paths to the system. Other settings are recorded in '`etc/Makeconf`' (or a sub-architecture-specific version), e.g.

```
JAVA = /usr/bin/java
JAVAC = /usr/bin/javac
JAVA_HOME = /usr/java/jdk1.5.0_06/jre
JAVA_LD_LIBRARY_PATH = $(JAVA_HOME)/lib/amd64/server:$(JAVA_HOME)/lib/amd64:\
$(JAVA_HOME)/../lib/amd64:/usr/local/lib64
JAVA_LIBS = -L$(JAVA_HOME)/lib/amd64/server -L$(JAVA_HOME)/lib/amd64
            -L$(JAVA_HOME)/../lib/amd64 -L/usr/local/lib64 -ljvm
```

where '`JAVA_LIBS`' contains flags necessary to link JNI programs. Some of the above variables can be queried using `R CMD config`.

A.3 Linear algebra

A.3.1 BLAS

The linear algebra routines in R can make use of enhanced BLAS (Basic Linear Algebra Subprograms, http://www.netlib.org/blas/faq.html) routines. However, these have to be explicitly requested at configure time: R provides an internal BLAS which is well-tested and will be adequate for most uses of R.

You can specify a particular BLAS library *via* a value for the configuration option '`--with-blas`' and not to use an external BLAS library by '`--without-blas`' (the default). If '`--with-blas`' is given with no `=`, its value is taken from the environment variable `BLAS_LIBS`, set for example in '`config.site`'. If neither the option nor the environment variable supply a value, a search is made for a suitable BLAS. If the value is not obviously a linker command (starting with a dash or giving the path to a library), it is prefixed by '`-l`', so

```
--with-blas="foo"
```

is an instruction to link against '`-lfoo`' to find an external BLAS (which needs to be found both at link time and run time).

The configure code checks that the external BLAS is complete (it must include all double precision and double complex routines[8], as well as `LSAME`), and appears to be usable. However, an external BLAS has to be usable from a shared object (so must contain position-independent code), and that is not checked.

Some enhanced BLASes are compiler-system-specific (`sunperf` on Solaris[9], `libessl` on IBM, `vecLib` on OS X). The correct incantation for these is usually found *via* '`--with-blas`' with no value on the appropriate platforms.

Some of the external BLASes are multi-threaded. One issue is that R profiling (which uses the `SIGPROF` signal) may cause problems, and you may want to disable profiling if you use a multi-threaded BLAS. Note that using a multi-threaded BLAS can result in taking more CPU time

[8] unless FORTRAN double complex is not supported on the platform

[9] Using the Oracle Solaris Studio `cc` and `f95` compilers

and even more elapsed time (occasionally dramatically so) than using a similar single-threaded BLAS.

Note that under Unix (but not under Windows) if R is compiled against a non-default BLAS and '`--enable-BLAS-shlib`' is **not** used, then all BLAS-using packages must also be. So if R is re-built to use an enhanced BLAS then packages such as **quantreg** will need to be re-installed.

R relies on IEC 60559 compliance of an external BLAS. This can be broken if for example the code assumes that terms with a zero factor are always zero and do not need to be computed—whereas `x*0` can be `NaN`. (The version of the reference BLAS used prior to R 2.12.0 did so.)

A.3.1.1 ATLAS

ATLAS (`http://math-atlas.sourceforge.net/`) is a "tuned" BLAS that runs on a wide range of Unix-alike platforms. Unfortunately it is usually built as a static library that on some platforms cannot be used with shared objects such as are used in R packages. Be careful when using pre-built versions of ATLAS (they seem to work on '`ix86`' platforms, but not on '`x86_64`' ones).

The usual way to specify ATLAS will be via

```
--with-blas="-lf77blas -latlas"
```

if the libraries are in the library path, otherwise by

```
--with-blas="-L/path/to/ATLAS/libs -lf77blas -latlas"
```

For systems with multiple processors it is possible to use a multi-threaded version of ATLAS, by specifying

```
--with-blas="-lptf77blas -lpthread -latlas"
```

Consult its file '`INSTALL.txt`' for how to build ATLAS with position-independent code (at least on version 3.8.0 and later): that file also describes how to build ATLAS as a shared library.

ATLAS can also be used on Windows: see see Section 3.1.2 [Getting the source files], page 10 when building from source, and R Windows FAQ for adding pre-compiled support to binary versions.

A.3.1.2 ACML

For '`x86_64`' and '`i686`' processors under Linux there is the AMD Core Math Library (ACML) `http://www.amd.com/acml`. For the `gcc` version we could use

```
--with-blas="-lacml"
```

if the appropriate library directory (such as '`/opt/acml4.4.0/gfortran64/lib`') is in the `LD_LIBRARY_PATH`. For other compilers, see the ACML documentation. There is a multithreaded Linux version of ACML available for recent versions of `gfortran`. To make use of this you will need something like

```
--with-blas="-L/opt/acml4.4.0/gfortran64_mp/lib -lacml_mp"
```

See see Section A.3.1.5 [Shared BLAS], page 36 for an alternative (and in many ways preferable) way to use ACML.

A.3.1.3 Goto BLAS

Dr Kazushige Goto has written another tuned BLAS which is available for several processors and OSes. The current version is known as GotoBLAS2, and has (in November 2010) been re-released under a much less restrictive licence. Source code can be obtained from `http://www.tacc.utexas.edu/tacc-projects/gotoblas2/`

Once it is built and installed, it can be used by configuring R with

```
--with-blas="-lgoto2"
```

See see Section A.3.1.5 [Shared BLAS], page 36 for an alternative (and in many ways preferable) way to use it.

Our understanding is that this project is now frozen and so will not be updated for CPUs released since mid-2010.

A.3.1.4 Intel MKL

svn up For Intel processors under Linux, there is Intel's Math Kernel Library (http://www.intel.com/software/products/mkl/). You are strongly encouraged to read the MKL User's Guide, which is installed with the library, before attempting to link to MKL.

Version 10 of MKL supports two linking models: the default model, which is backward compatible with version 9 (see below), and the pure layered model. The layered model gives the user fine-grained control over four different library layers: interface, threading, computation, and run-time library support. Some examples of linking to MKL using this layered model are given below. (These examples are for GCC compilers on '`x86_64`'.) The choice of interface layer is important on '`x86_64`' since the Intel Fortran compiler returns complex values in different registers from the GNU Fortran compiler. You must therefore use the interface layer that matches your compiler (`mkl_intel*` or `mkl_gf*`).

R can be linked to a sequential version of MKL by something like

```
MKL_LIB_PATH=/opt/intel/mkl/10.0.3.020/lib/em64t/
export LD_LIBRARY_PATH=$MKL_LIB_PATH
MKL="-L${MKL_LIB_PATH} -lmkl_gf_lp64 -lmkl_sequential -lmkl_lapack -lmkl_core"
./configure --with-blas="$MKL" --with-lapack
```

The order of the libraries is important. The option '`--with-lapack`' is used since MKL contains a copy of LAPACK as well as BLAS (see Section A.3.2 [LAPACK], page 37).

Threaded MKL may be used by replacing the line defining the variable `MKL` with

```
MKL="-L${MKL_LIB_PATH} -lmkl_gf_lp64 -lmkl_gnu_thread \
     -lmkl_lapack -lmkl_core -liomp5 -lpthread"
```

The default number of threads will be chosen by the OpenMP software, but can be controlled by setting `OMP_NUM_THREADS` or `MKL_NUM_THREADS`.

Static sequential MKL may be used with

```
MKL="   -L${MKL_LIB_PATH}                                  \
        -Wl,--start-group                                  \
                ${MKL_LIB_PATH}/libmkl_gf_lp64.a           \
                ${MKL_LIB_PATH}/libmkl_gnu_thread.a        \
                ${MKL_LIB_PATH}/libmkl_core.a              \
        -Wl,--end-group                                    \
        -lgomp -lpthread"
```

(Thanks to Ei-ji Nakama).

The default linking model, which is also used by version 9 of MKL, can be used by

```
--with-blas="-lmkl -lguide -lpthread"
```

This is multi-threaded, but in version 9 the number of threads defaults to 1. It can be increased by setting `OMP_NUM_THREADS`. (Thanks to Andy Liaw for the information.)

A.3.1.5 Shared BLAS

The BLAS library will be used for many of the add-on packages as well as for R itself. This means that it is better to use a shared/dynamic BLAS library, as most of a static library will be compiled into the R executable and each BLAS-using package.

R offers the option of compiling the BLAS into a dynamic library `libRblas` stored in '*`R_HOME/lib`*' and linking both R itself and all the add-on packages against that library.

This is the default on all platforms except AIX unless an external BLAS is specified and found: for the latter it can be used by specifying the option '`--enable-BLAS-shlib`', and it can always be disabled via '`--disable-BLAS-shlib`'.

This has both advantages and disadvantages.

- It saves space by having only a single copy of the BLAS routines, which is helpful if there is an external static BLAS such as is standard for ATLAS.
- There may be performance disadvantages in using a shared BLAS. Probably the most likely is when R's internal BLAS is used and R is *not* built as a shared library, when it is possible to build the BLAS into '`R.bin`' (and '`libR.a`') without using position-independent code. However, experiments showed that in many cases using a shared BLAS was as fast, provided high levels (e.g., '`-O3`') of compiler optimization are used.
- It is easy to change the BLAS without needing to re-install R and all the add-on packages, since all references to the BLAS go through `libRblas`, and that can be replaced. Note though that any dynamic libraries the replacement links to will need to be found by the linker: this may need the library path to be changed in '*`R_HOME/etc/ldpaths`*'.

Another option to change the BLAS in use is to symlink a dynamic BLAS library (such as ACML or Goto's) to '*`R_HOME/lib/libRblas.so`*'. For example, just

```
mv R_HOME/lib/libRblas.so R_HOME/lib/libRblas.so.keep
ln -s /opt/acml4.4.0/gfortran64_mp/lib/libacml_mp.so R_HOME/lib/libRblas.so
```

will change the BLAS in use to multithreaded ACML. A similar link works for recent versions of the Goto BLAS and perhaps for MKL (provided the appropriate '`lib`' directory is in the run-time library path or `ld.so` cache).

A.3.2 LAPACK

Provision is made for using an external LAPACK library, principally to cope with BLAS libraries which contain a copy of LAPACK (such as `sunperf` on Solaris, `vecLib` on OS X and ACML on '`ix86`'/'`x86_64`' Linux). However, the likely performance gains are thought to be small (and may be negative), and the default is not to search for a suitable LAPACK library, and this is definitely **not** recommended. You can specify a specific LAPACK library or a search for a generic library by the configuration option '`--with-lapack`'. The default for '`--with-lapack`' is to check the BLAS library and then look for an external library '`-llapack`'. Sites searching for the fastest possible linear algebra may want to build a LAPACK library using the ATLAS-optimized subset of LAPACK. To do so specify something like

```
--with-lapack="-L/path/to/libs -llapack -lcblas"
```

since the ATLAS subset of LAPACK depends on `libcblas`. A value for '`--with-lapack`' can be set *via* the environment variable `LAPACK_LIBS`, but this will only be used if '`--with-lapack`' is specified (as the default value is `no`) and the BLAS library does not contain LAPACK.

Since ACML contains a full LAPACK, if selected as the BLAS it can be used as the LAPACK *via* '`--with-lapack`'.

If you do use '`--with-lapack`', be aware of potential problems with bugs in the LAPACK 3.0 sources (or in the posted corrections to those sources). In particular, bugs in `DGEEV` and `DGESDD` have resulted in error messages such as

```
DGEBRD gave error code -10
```

. Other potential problems are incomplete versions of the libraries, seen several times over the years. For problems compiling LAPACK using recent versions of `gcc` on '`ix86`' Linux, see Section C.8 [New platforms], page 53.

Please **do** bear in mind that using '`--with-lapack`' is 'definitely **not** recommended': it is provided **only** because it is necessary on some platforms. Reporting problems where it is used unnecessarily will simply irritate the R helpers.

A.3.3 Caveats

As with all libraries, you need to ensure that they and R were compiled with compatible compilers and flags. For example, this has meant that on Sun Sparc using the native compilers the flag '`-dalign`' is needed so `sunperf` can be used.

On some systems it is necessary that an external BLAS/LAPACK was built with the same FORTRAN compiler used to build R: known problems are with R built with `gfortran`, see Section B.6.1 [Using gfortran], page 42.

Appendix B Configuration on a Unix-alike

B.1 Configuration options

`configure` has many options: running

```
./configure --help
```

will give a list. Probably the most important ones not covered elsewhere are (defaults in brackets)

'`--with-x`'
: use the X Window System [yes]

'`--x-includes=`*DIR*'
: X include files are in *DIR*

'`--x-libraries=`*DIR*'
: X library files are in *DIR*

'`--with-readline`'
: use readline library (if available) [yes]

'`--enable-R-profiling`'
: attempt to compile support for `Rprof()` [yes]

'`--enable-memory-profiling`'
: attempt to compile support for `Rprofmem()` and `tracemem()` [no]

'`--enable-R-shlib`'
: build R as a shared/dynamic library [no]

'`--enable-BLAS-shlib`'
: build the BLAS as a shared/dynamic library [yes, except on AIX]

You can use '`--without-foo`' or '`--disable-foo`' for the negatives.

You will want to use '`--disable-R-profiling`' if you are building a profiled executable of R (e.g. with '`-pg`)'.

Flag '`--enable-R-shlib`' causes the make process to build R as a dynamic (shared) library, typically called '`libR.so`', and link the main R executable '`R.bin`' against that library. This can only be done if all the code (including system libraries) can be compiled into a dynamic library, and there may be a performance[1] penalty. So you probably only want this if you will be using an application which embeds R. Note that C code in packages installed on an R system linked with '`--enable-R-shlib`' is linked against the dynamic library and so such packages cannot be used from an R system built in the default way. Also, because packages are linked against R they are on some OSes also linked against the dynamic libraries R itself is linked against, and this can lead to symbol conflicts.

If you need to re-configure R with different options you may need to run `make clean` or even `make distclean` before doing so.

B.2 Internationalization support

Translation of messages is supported via GNU `gettext` unless disabled by the configure option '`--disable-nls`'. The `configure` report will show `NLS` as one of the 'Additional capabilities' if support has been compiled in, and running in an English locale (but not the `C` locale) will include

```
Natural language support but running in an English locale
```

in the greeting on starting R.

[1] We have measured 15–20% on '`i686`' Linux and around 10% on '`x86_64`' Linux.

B.3 Configuration variables

If you need or want to set certain configure variables to something other than their default, you can do that by either editing the file '`config.site`' (which documents many of the variables you might want to set: others can be seen in file '`etc/Renviron.in`') or on the command line as

```
./configure VAR=value
```

If you are building in a directory different from the sources, there can be copies of '`config.site`' in the source and the build directories, and both will be read (in that order). In addition, if there is a file '`~/.R/config`' (or failing that) `~/.Rconfig`), it is read between the '`config.site`' files in the source and the build directories.

There is also a general `autoconf` mechanism for '`config.site`' files, which are read before any of those mentioned in the previous paragraph. This looks first at a file specified by the environment variable `CONFIG_SITE`, and if not is set at files such as '`/usr/local/share/config.site`' and '`/usr/local/etc/config.site`' in the area (exemplified by '`/usr/local`') where R would be installed.

These variables are *precious*, implying that they do not have to be exported to the environment, are kept in the cache even if not specified on the command line, checked for consistency between two configure runs (provided that caching is used), and are kept during automatic reconfiguration as if having been passed as command line arguments, even if no cache is used.

See the variable output section of `configure --help` for a list of all these variables.

If you find you need to alter configure variables, it is worth noting that some settings may be cached in the file '`config.cache`', and it is a good idea to remove that file (if it exists) before re-configuring. Note that caching is turned *off* by default: use the command line option '`--config-cache`' (or '`-C`') to enable caching.

B.3.1 Setting paper size

One common variable to change is `R_PAPERSIZE`, which defaults to '`a4`', not '`letter`'. (Valid values are '`a4`', '`letter`', '`legal`' and '`executive`'.)

This is used both when configuring R to set the default, and when running R to override the default. It is also used to set the paper size when making DVI and PDF manuals.

The configure default will most often be '`a4`' if `R_PAPERSIZE` is unset. (If the (Debian Linux) program `paperconf` is found or the environment variable `PAPERSIZE` is set, these are used to produce the default.)

B.3.2 Setting the browsers

Another precious variable is `R_BROWSER`, the default HTML browser, which should take a value of an executable in the user's path or specify a full path.

Its counterpart for PDF files is `R_PDFVIEWER`.

B.3.3 Compilation flags

If you have libraries and header files, e.g., for GNU readline, in non-system directories, use the variables `LDFLAGS` (for libraries, using '`-L`' flags to be passed to the linker) and `CPPFLAGS` (for header files, using '`-I`' flags to be passed to the C/C++ preprocessors), respectively, to specify these locations. These default to '`-L/usr/local/lib`' (`LDFLAGS`, '`-L/usr/local/lib64`' on most 64-bit Linux OSes) and '`-I/usr/local/include`' (`CPPFLAGS`) to catch the most common cases. If libraries are still not found, then maybe your compiler/linker does not support re-ordering of '`-L`' and '`-l`' flags (this has been reported to be a problem on HP-UX with the native `cc`). In this case, use a different compiler (or a front end shell script which does the re-ordering).

These flags can also be used to build a faster-running version of R. On most platforms using `gcc`, having '`-O3`' in `CFLAGS` and `FFLAGS` produces worthwhile performance gains with `gcc` and `gfortran`. On systems using the GNU linker (especially those using R as a shared library), it is likely that including '`-Wl,-O1`' in `LDFLAGS` is worthwhile, and ''`-Bdirect,--hash-style=both,-Wl,-O1`'' is recommended at http://lwn.net/Articles/192624/. Tuning compilation to a specific CPU family (e.g. '`-mtune=native`' for `gcc`) can give worthwhile performance gains, especially on older architectures such as '`ix86`'.

B.3.4 Making manuals

The default settings for making the manuals are controlled by `R_RD4PDF`, `R_RD4DVI` and `R_PAPERSIZE`.

B.4 Setting the shell

By default the shell scripts such as '`R`' will be '`#!/bin/sh`' scripts (or using the `SHELL` chosen by '`configure`'). This is almost always satisfactory, but on a few systems '`/bin/sh`' is not a Bourne shell or clone, and the shell to be used can be changed by setting the configure variable `R_SHELL` to a suitable value (a full path to a shell, e.g. '`/usr/local/bin/bash`').

B.5 Using make

To compile R, you will most likely find it easiest to use GNU `make`, although the Sun `make` works on Solaris, as does the native FreeBSD `make`. The native `make` has been reported to fail on SGI Irix 6.5 and Alpha/OSF1 (aka Tru64).

To build in a separate directory you need a `make` that uses the `VPATH` variable, for example GNU `make`, or Sun `make` on Solaris 7 or later.

`dmake` has also been used. e.g, on Solaris 10.

If you want to use a `make` by another name, for example if your GNU `make` is called '`gmake`', you need to set the variable `MAKE` at configure time, for example

```
./configure MAKE=gmake
```

B.6 Using FORTRAN

To compile R, you need a FORTRAN compiler. The default is to search for `f95`, `fort`, `xlf95`, `ifort`, `ifc`, `efc`, `pgf95` `lf95`, `gfortran`, `ftn`, `g95`, `f90`, `xlf90`, `pghpf`, `pgf90`, `epcf90`, `g77`, `f77`, `xlf`, `frt`, `pgf77`, `cf77`, `fort77`, `fl32`, `af77` (in that order)[2], and use whichever is found first; if none is found, R cannot be compiled. However, if `CC` is `gcc`, the matching FORTRAN compiler (`g77` for `gcc` 3 and `gfortran` for `gcc` 4) is used if available.

The search mechanism can be changed using the configure variable `F77` which specifies the command that runs the FORTRAN 77 compiler. If your FORTRAN compiler is in a non-standard location, you should set the environment variable `PATH` accordingly before running `configure`, or use the configure variable `F77` to specify its full path.

If your FORTRAN libraries are in slightly peculiar places, you should also look at `LD_LIBRARY_PATH` or your system's equivalent to make sure that all libraries are on this path.

Note that only FORTRAN compilers which convert identifiers to lower case are supported.

You must set whatever compilation flags (if any) are needed to ensure that FORTRAN `integer` is equivalent to a C `int` pointer and FORTRAN `double precision` is equivalent to a C `double` pointer. This is checked during the configuration process.

[2] On HP-UX `fort77` is the POSIX compliant FORTRAN compiler, and comes after `g77`.

Some of the FORTRAN code makes use of `COMPLEX*16` variables, which is a Fortran 90 extension. This is checked for at configure time[3], but you may need to avoid compiler flags[4] asserting FORTRAN 77 compliance.

For performance reasons[5] you may want to choose a FORTRAN 90/95 compiler.

It may be possible to use `f2c`, the FORTRAN-to-C converter (http://www.netlib.org/f2c), via a script. (An example script is given in '`scripts/f77_f2c`': this can be customized by setting the environment variables `F2C`, `F2CLIBS`, `CC` and `CPP`.) You will need to ensure that the FORTRAN type `integer` is translated to the C type `int`. Normally '`f2c.h`' contains '`typedef long int integer;`', which will work on a 32-bit platform but not on a 64-bit platform. If your compiler is not `gcc` you will need to set `FPICFLAGS` appropriately.

B.6.1 Using gfortran

`gfortran` is the F95 compiler that is part of `gcc` 4.*x.y*. There were problems compiling R with the first release (`gcc` 4.0.0) and more with pre-releases, but these are resolved in later versions.

On Linux '`x86_64`' systems there is an incompatibility in the return conventions for double-complex functions between `gfortran` and `g77` which results in the final example in `example(eigen)` hanging or segfaulting under external BLASs built under `g77`. This should be detected by a `configure` test.

The default `FFLAGS` and `FCFLAGS` chosen (by `autoconf`) for a GNU FORTRAN compiler is '`-g -O2`'. This has caused problems (segfaults and infinite loops) on '`x86_64`' Linux in the past, but seems fine with `gfortran 4.4.4` and later: for `gfortran 4.3.x` set `FFLAGS` and `FCFLAGS` to use at most '`-O`'.

B.7 Compile and load flags

A wide range of flags can be set in the file '`config.site`' or as configure variables on the command line. We have already mentioned

`CPPFLAGS` header file search directory ('`-I`') and any other miscellaneous options for the C and C++ preprocessors and compilers

`LDFLAGS` path ('`-L`'), stripping ('`-s`') and any other miscellaneous options for the linker

and others include

`CFLAGS` debugging and optimization flags, C

`MAIN_CFLAGS`
ditto, for compiling the main program

`SHLIB_CFLAGS`
for shared objects

`FFLAGS` debugging and optimization flags, FORTRAN

`SAFE_FFLAGS`
ditto for source files which need exact floating point behaviour

`MAIN_FFLAGS`
ditto, for compiling the main program

`SHLIB_FFLAGS`
for shared objects

[3] as well as its equivalence to the `Rcomplex` structure defined in '`R_ext/Complex.h`'.

[4] In particular, avoid `g77`'s '`-pedantic`', which gives confusing error messages.

[5] e.g., to use an optimized BLAS on Sun/Sparc

`MAIN_LDFLAGS`
: additional flags for the main link

`SHLIB_LDFLAGS`
: additional flags for linking the shared objects

`LIBnn`
: the primary library directory, '`lib`' or '`lib64`'

`CPICFLAGS`
: special flags for compiling C code to be turned into a shared object

`FPICFLAGS`
: special flags for compiling Fortran code to be turned into a shared object

`CXXPICFLAGS`
: special flags for compiling C++ code to be turned into a shared object

`FCPICFLAGS`
: special flags for compiling Fortran 95 code to be turned into a shared object

`DEFS`
: defines to be used when compiling C code in R itself

Library paths specified as '`-L/lib/path`' in `LDFLAGS` are collected together and prepended to `LD_LIBRARY_PATH` (or your system's equivalent), so there should be no need for '`-R`' or '`-rpath`' flags.

Variables such as `CPICFLAGS` are determined where possible by `configure`. Some systems allows two types of PIC flags, for example '`-fpic`' and '`-fPIC`', and if they differ the first allows only a limited number of symbols in a shared object. Since R as a shared library has about 6200 symbols, if in doubt use the larger version.

To compile a profiling version of R, one might for example want to use '`MAIN_CFLAGS=-pg`', '`MAIN_FFLAGS=-pg`', '`MAIN_LDFLAGS=-pg`' on platforms where '`-pg`' cannot be used with position-independent code.

Beware: it may be necessary to set `CFLAGS` and `FFLAGS` in ways compatible with the libraries to be used: one possible issue is the alignment of doubles, another is the way structures are passed.

On some platforms `configure` will select additional flags for `CFLAGS`, `CPPFLAGS`, `FFLAGS`, `CXXFLAGS` and `LIBS` in `R_XTRA_CFLAGS` (and so on). These are for options which are always required, for example to force IEC 60559 compliance.

Appendix C Platform notes

This section provides some notes on building R on different Unix-alike platforms. These notes are based on tests run on one or two systems in each case with particular sets of compilers and support libraries. Success in building R depends on the proper installation and functioning of support software; your results may differ if you have other versions of compilers and support libraries.

Older versions of this manual (for R < 2.10.0) contain notes on platforms such as HP-UX, IRIX and Alpha/OSF1 for which we have had no recent reports.

C.1 X11 issues

The '`X11()`' graphics device is the one started automatically on Unix-alikes when plotting. As its name implies, it displays on a (local or remote) X server, and relies on the services and in particular the fonts provided by the X server. So if you sometimes use R at a console and sometimes remotely from an X11 session running on a Windows machine, you may have to setup the fonts differently for the two usages.

The 'modern' version of the '`X11()`' device is based on '`cairo`' graphics and (in most implementations) uses '`fontconfig`' to pick and render fonts. This is done on the server, and although there can be selection issues, they are more amenable than the issues with '`X11()`' discussed in the rest of this section.

When X11 was designed, most displays were around 75dpi, whereas today they are of the order of 100dpi or more. If you find that X11() is reporting[1] missing font sizes, especially larger ones, it is likely that you are not using scalable fonts and have not installed the 100dpi versions of the X11 fonts. The names and details differ by system, but will likely have something like Fedora's

```
xorg-x11-fonts-75dpi
xorg-x11-fonts-100dpi
xorg-x11-fonts-ISO8859-2-75dpi
xorg-x11-fonts-Type1
xorg-x11-fonts-cyrillic
```

and you need to ensure that the '`-100dpi`' versions are installed and on the X11 font path (check via `xset -q`). The '`X11()`' device does try to set a pointsize and not a pixel size: laptop users may find the default setting of 12 too large (although very frequently laptop screens are set to a fictitious dpi to appear like a scaled-down desktop screen).

More complicated problems can occur in non-Western-European locales, so if you are using one, the first thing to check is that things work in the `C` locale. The likely issues are a failure to find any fonts or glyphs being rendered incorrectly (often as a pair of ASCII characters). X11 works by being asked for a font specification and coming up with its idea of a close match. For text (as distinct from the symbols used by plotmath), the specification is the first element of the option `"X11fonts"` which defaults to

```
"-adobe-helvetica-%s-%s-*-*-%d-*-*-*-*-*-*-*"
```

If you are using a single-byte encoding, for example ISO 8859-2 in Eastern Europe or KOI8-R in Russian, use `xlsfonts` to find an appropriate family of fonts in your encoding (the last field in the listing). If you find none, it is likely that you need to install further font packages, such as '`xorg-x11-fonts-ISO8859-2-75dpi`' and '`xorg-x11-fonts-cyrillic`' shown in the listing above.

Multi-byte encodings (most commonly UTF-8) are even more complicated. There are few fonts in '`iso10646-1`', the Unicode encoding, and they only contain a subset of the available

[1] for example, `X11 font at size 14 could not be loaded`.

glyphs (and are often fixed-width designed for use in terminals). In such locales *fontsets* are used, made up of fonts encoded in other encodings. If the locale you are using has an entry in the '`XLC_LOCALE`' directory (typically '`/usr/share/X11/locale`', it is likely that all you need to do is to pick a suitable font specification that has fonts in the encodings specified there. If not, you may have to get hold of a suitable locale entry for X11. This may mean that, for example, Japanese text can be displayed when running in '`ja_JP.UTF-8`' but not when running in '`en_GB.UTF-8`' on the same machine (although on some systems many UTF-8 X11 locales are aliased to '`en_US.UTF-8`' which covers several character sets, e.g. ISO 8859-1 (Western European), JISX0208 (Kanji), KSC5601 (Korean), GB2312 (Chinese Han) and JISX0201 (Kana)).

On some systems scalable fonts are available covering a wide range of glyphs. One source is TrueType/OpenType fonts, and these can provide high coverage. Another is Type 1 fonts: the URW set of Type 1 fonts provides standard typefaces such as Helvetica with a larger coverage of Unicode glyphs than the standard X11 bitmaps, including Cyrillic. These are generally not part of the default install, and the X server may need to be configured to use them. They might be under the X11 '`fonts`' directory or elsewhere, for example,

```
/usr/share/fonts/default/Type1
/usr/share/fonts/ja/TrueType
```

C.2 Linux

Linux is the main development platform for R, so compilation from the sources is normally straightforward with the standard compilers.

Remember that some package management systems (such as RPM and deb) make a distinction between the user version of a package and the developer version. The latter usually has the same name but with the extension '`-devel`' or '`-dev`': you need both versions installed. So please check the `configure` output to see if the expected features are detected: if for example '`readline`' is missing add the developer package. (On most systems you will also need '`ncurses`' and its developer package, although these should be dependencies of the '`readline`' package(s).)

When R has been installed from a binary distribution there are sometimes problems with missing components such as the FORTRAN compiler. Searching the '`R-help`' archives will normally reveal what is needed.

It seems that '`ix86`' Linux accepts non-PIC code in shared libraries, but this is not necessarily so on other platforms, in particular on 64-bit CPUs such as '`x86_64`'. So care can be needed with BLAS libraries and when building R as a shared library to ensure that position-independent code is used in any static libraries (such as the Tcl/Tk libraries, `libpng`, `libjpeg` and `zlib`) which might be linked against. Fortunately these are normally built as shared libraries with the exception of the ATLAS BLAS libraries.

The default optimization settings chosen for `CFLAGS` etc are conservative. It is likely that using '`-mtune`' will result in significant performance improvements on recent CPUs (especially for '`ix86`'): one possibility is to add '`-mtune=native`' for the best possible performance on the machine on which R is being installed: if the compilation is for a site-wide installation, it may still be desirable to use something like '`-mtume=core2`'. It is also possible to increase the optimization levels to '`-O3`': however for many versions of the compilers this has caused problems in at least one CRAN package.

For platforms with both 64- and 32-bit support, it is likely that

```
LDFLAGS="-L/usr/local/lib64 -L/usr/local/lib"
```

is appropriate since most (but not all) software installs its 64-bit libraries in '`/usr/local/lib64`'. To build a 32-bit version of R on '`x86_64`' with Fedora 14 we used

```
CC="gcc -m32"
CXX="g++ -m32"
```

```
F77="gfortran -m32"
FC=${F77}
OBJC=${CC}
LDFLAGS="-L/usr/local/lib"
LIBnn=lib
```

64-bit versions of Linux are built with support for files > 2Gb, and 32-bit versions will be if possible unless '`--disable-largefile`' is specified.

To build a 64-bit version of R on '`ppc64`' (also known as '`powerpc64`') with `gcc` 4.1.1, Ei-ji Nakama used

```
CC="gcc -m64"
CXX="gxx -m64"
F77="gfortran -m64"
FC="gfortran -m64"
CFLAGS="-mminimal-toc -fno-optimize-sibling-calls -g -O2"
FFLAGS="-mminimal-toc -fno-optimize-sibling-calls -g -O2"
```

the additional flags being needed to resolve problems linking against '`libnmath.a`' and when linking R as a shared library.

C.2.1 Intel compilers

Intel compilers have been used under '`ix86`' and '`x86_64`' Linux. Brian Ripley used version 9.0 of the compilers for '`x86_64`' on Fedora Core 5 with

```
CC=icc
CFLAGS="-g -O3 -wd188 -ip -mp"
F77=ifort
FLAGS="-g -O3 -mp"
CXX=icpc
CXXFLAGS="-g -O3 -mp"
FC=ifort
FCFLAGS="-g -O3 -mp"
ICC_LIBS=/opt/compilers/intel/cce/9.1.039/lib
IFC_LIBS=/opt/compilers/intel/fce/9.1.033/lib
LDFLAGS="-L$ICC_LIBS -L$IFC_LIBS -L/usr/local/lib64"
SHLIB_CXXLD=icpc
```

`configure` will add '`-c99`' to `CC` for C99-compliance. This causes warnings with `icc` 10 and later, so use `CC="icc -std=c99"` there. The flag '`-wd188`' suppresses a large number of warnings about the enumeration type '`Rboolean`'. Because the Intel C compiler sets '`__GNUC__`' without complete emulation of `gcc`, we suggest adding `CPPFLAGS=-no-gcc`.

To maintain correct IEC 60559 arithmetic you most likely need add flags to `CFLAGS`, `FFLAGS` and `CXXFLAGS` such as '`-mp`' (shown above) or '`-fp-model precise -fp-model source`', depending on the compiler version.

For some comments on building on an Itanium ('`ia64`') Linux system with `gcc` or the Intel compilers see http://www.nakama.ne.jp/memo/ia64_linux/.

Others have reported success with versions 10.x and 11.x.

C.2.2 Oracle Solaris Studio compilers

Brian Ripley tested the Sun Studio 12 compilers, since renamed to Oracle Solaris Studio, (http://developers.sun.com/sunstudio/index.jsp) On '`x86_64`' Linux with

```
CC=suncc
CFLAGS="-xO5 -xc99 -xlibmil -nofstore"
```

```
CPICFLAGS=-Kpic
F77=sunf95
FFLAGS="-O5 -libmil -nofstore"
FPICFLAGS=-Kpic
CXX="sunCC -library=stlport4"
CXXFLAGS="-xO5 -xlibmil -nofstore -features=tmplrefstatic"
CXXPICFLAGS=-Kpic
FC=sunf95
FCFLAGS=$FFLAGS
FCPICFLAGS=-Kpic
LDFLAGS=-L/opt/sunstudio12.1/rtlibs/amd64
SHLIB_LDFLAGS=-shared
SHLIB_CXXLDFLAGS=-G
SHLIB_FCLDFLAGS=-G
SAFE_FFLAGS="-O5 -libmil"
```

'`-m64`' could be added, but was the default. Do not use '`-fast`': see the warnings under Solaris. (The C++ options are also explained under Solaris.)

Others have found on at least some versions of '`ix86`' Linux that the configure flag '`--disable-largefile`' was needed (since '`glob.h`' on that platform presumed `gcc` was being used).

C.3 FreeBSD

Rainer Hurling has reported success on '`amd64`' FreeBSD 9.0 (and on earlier versions in the past), and Brian Ripley tested '`amd64`' FreeBSD 8.2. Since Darwin (the base OS of (Mac) OS X) is based on FreeBSD we find testing on Darwin tends to pick up most potential problems on FreeBSD. However, FreeBSD lacks adequate collation support for multi-byte locales (but a port of ICU is available), and does not yet implement C99 complex math functions (for which R's substitutes are used).

The native BSD `make` suffices to build R but a number of packages require GNU `make`, despite the recommendations of the "Writing R Extensions" manual. (The BSD version is `bsdmake` on Darwin.)

The simplest way to get the additional software needed to build R is to install a pre-compiled version first, e.g. by

```
pkg_add -r R
```

(on the system this was tested on, this installed Tcl, Tk, blas, lapack and `gcc-4.5.3` which includes `gfortran45`). A listing of dependencies (not necessarily for current R) can be found at http://www.freebsd.org/ports/lang.html: you will however also need a TEX system[2] to build the manuals.

Then R itself can be built by something like

```
./configure CC=gcc45 F77=gfortran45 CXX=g++45 FC=gfortran45
```

There are also FreeBSD packages for a small eclectic collection of CRAN packages.

C.4 (Mac) OS X

You can build R as a Unix application on OS X using the Apple Developer Tools ('Xcode') and `gfortran`. You may also need to install an X sub-system or configure with '`--without-x`'. The X window manager is part of the standard OS X distribution since OS X version 10.3 (Panther),

[2] TeXLive is recommended.

but it is typically not pre-installed prior to 10.5 (Leopard). You will also need `gfortran` and `libreadline` (or configure with '`--without-readline`').

For more information on how to find these tools please read the R for Mac OS X FAQ.

The `vecLib` library can be used *via* the (default) configuration options

```
--with-blas="-framework vecLib" --with-lapack
```

to provide higher-performance versions of the BLAS and LAPACK routines. Building R without these options *via*

```
--without-blas --without-lapack
```

can be done (and is provided as an alternative in the binary distribution).

C.4.1 64-bit builds

64-bit builds are supported on 10.5.x (Leopard) and later. All that is needed is to select suitable compiler options, e.g. for most Intel Macs

```
CC='gcc -arch x86_64'
CXX='g++ -arch x86_64'
F77='gfortran -arch x86_64'
FC='gfortran -arch x86_64'
OBJC='gcc -arch x86_64'
```

in '`config.site`' or on the `configure` command line. (Or on Leopard specify '`gcc-4.2`' etc: those versions are the default on Snow Leopard.)

C.4.2 Snow Leopard

On Snow Leopard you will most likely to need to specify '`-arch i386`' for a 32-bit build and '`-arch x86_64`' for a 64-bit build: the Apple compilers default to 64-bit but the `gfortran` supplied on `http://r.research.att.com` defaults to 32-bit.

Another quirk is that the X11 libraries are not in the default linking path, so something like '`LIBS=-L/usr/X11/lib`' may be required (or you can use the `configure` options '`--x-includes=/usr/X11/include --x-libraries=/usr/X11/lib`.').

So for a 64-bit build of R you need a file '`config.site`' in the top-level build directory containing

```
CC='gcc -arch x86_64'
CXX='g++ -arch x86_64'
F77='gfortran -arch x86_64'
FC='gfortran -arch x86_64'
OBJC='gcc -arch x86_64'
LIBS=-L/usr/X11/lib
```

whereas for a 32-bit build it should have

```
CC='gcc -arch i386'
CXX='g++ -arch i386'
F77='gfortran -arch i386'
FC='gfortran -arch i386'
OBJC='gcc -arch i386'
LIBS=-L/usr/X11/lib
```

C.5 Solaris

R has been built successfully on Solaris 10 (both Sparc and '`x86`') using the (zero cost) Oracle Solaris Studio 12 compilers: there has been some success with `gcc` 4/`gfortran`, mainly on Sparc. Sun packages for R are available from `http://www.sunfreeware.com/` for both 32-bit

architectures, but recently have often been several versions old.[3] (Recent Sun machines are AMD Opterons or Intel Xeons ('amd64') rather than 'x86', but 32-bit 'x86' executables are the default.)

There were also reports of success on OpenSolaris (also known as Solaris Express Community Edition, and dropped in August 2010) on 'x86'.

The Solaris versions of several of the tools needed to build R (e.g. make, ar and ld) are in '/usr/ccs/bin', so if using those tools ensure this is in your path. A version of the preferred GNU tar is (if installed) in '/usr/sfw/bin', as sometimes are tools like makeinfo. It may be necessary to avoid the tools in '/usr/ucb': in particular '/usr/ucb/tr' has caused problems in the past. POSIX-compliant versions of some tools can be found in '/usr/xpg4/bin' and '/usr/xpg6/bin'.

A large selection of Open Source software can be installed from http://www.opencsw.org (and also http://www.blastwave.org) via pkg-get, by default installed under '/opt/csw'.

You will need GNU libiconv and readline: the Solaris version of iconv is not sufficiently powerful.

The native make suffices to build R but a small number of packages require GNU make (most without good reason and without declaring it as 'SystemRequirements' in the 'DESCRIPTION' file).

Some people have reported that the Solaris libintl needs to be avoided, for example by using '--disable-nls' or '--with-included-gettext' or using libintl from OpenCSW.

When using the Oracle compilers[4] do *not* specify '-fast', as this disables IEEE arithmetic and make check will fail.

For the Solaris Studio compilers a little juggling of paths was needed to ensure GNU libiconv (in '/usr/local') was used rather than the Solaris iconv:

```
CC="cc -xc99"
CFLAGS="-O -xlibmieee"
F77=f95
FFLAGS=-O4
CXX="CC -library=stlport4"
CXXFLAGS=-O
FC=f95
FCFLAGS=$FFLAGS
FCLIBS="-lfai -lfsu"
R_LD_LIBRARY_PATH="/usr/local/lib:/opt/csw/gcc4/lib:/opt/csw/lib"
```

For a 64-bit target add '-m64' to the compiler macros and use something like LDFLAGS=-L/usr/local/lib/sparcv9 or LDFLAGS=-L/usr/local/lib/amd64 as appropriate.

With Solaris Studio 12.2 on Sparc, FCLIBS needs to be

```
FCLIBS="-lfai -lfai2 -lfsu"
```

(and possibly other Fortran libraries, but this suffices for the packages currently on CRAN).

Currently 'amd64' and 'sparcv9' builds work out-of-the-box with Sun Studio 12u1 but not Solaris Studio 12.2: 'libRblas.so' and 'lapack.so' are generated with code that causes relocation errors (which is being linked in from the Fortran libraries). This means that building R as a shared library may be impossible with Solaris Studio 12.2. For a standard build the trick seems to be manually set FLIBS. For example, on 'amd64', set in 'Makeconf'

```
FLIBS_IN_SO = -R/opt/solstudio12.2/lib/amd64 \
  /opt/solstudio12.2/lib/amd64/libfsu.so
```

[3] For example, 2.11.1 at the time of release of 2.13.0.

[4] including gcc for Sparc from Oracle.

and set FLIBS in 'etc/Makeconf' to the same value.

For 64-bit Sparc, use something like

```
FLIBS = -R/opt/solstudio12.2/prod/lib/sparc/64 -lifai -lsunimath -lfai -lfai2 \
 -lfsumai -lfprodai -lfminlai -lfmaxlai -lfminvai -lfmaxvai -lfui \
 -lsunmath -lmtsk /opt/solstudio12.2/prod/lib/sparc/64/libfsu.so.1
```

in 'Makeconf' and 'etc/Makefconf'.

By default the Solaris Studio compilers do not by default conform to the C99 standard (appendix F 8.9) on the return values of functions such as log: use '-xlibmieee' to ensure this.

You can target specific Sparc architectures for (slightly) higher performance: Oracle recommend

```
32-bit: -xtarget=ultra3 -xarch=v8plusa
64-bit: -xtarget=ultra3 -xarch=sparcvis2
```

(in CFLAGS etc.) as a good compromise for recent Sparc chipsets. Using -libmil in CFLAGS allows more system mathematical functions to be inlined.

On 'x86' you will get marginally higher performance *via*

```
CFLAGS="-O5 -xc99 -xlibmieee -xlibmil -nofstore -xtarget=native"
FFLAGS="-O5 -libmil -nofstore -xtarget=native"
CXXFLAGS="-O5 -xlibmil -nofstore -xtarget=native"
SAFE_FFLAGS="-libmil -fstore -xtarget=native"
```

but the use of -nofstore can be less numerically stable.

The Solaris Studio compilers provide several implementations of the C++ standard which select both the set of headers and a C++ runtime library. These are selected by the '-library' flag, which as it is needed for both compiling and linking is best specified as part of the compiler. The examples above use 'stlport4', currently the most modern of the options: the default (but still needed to be specified as it is needed for linking) is 'Cstd': see http://developers.sun.com/solaris/articles/cmp_stlport_libCstd.html. Note though that most external Solaris C++ libraries will have been built with 'Cstd' and so an R package using such libraries also needs to be. Occasionally the flag '-library=stlport4,Crun' is needed.

Several CRAN packages using C++ need the more liberal interpretation given by adding

```
CXXFLAGS="-features=tmplrefstatic"
```

The performance library sunperf is available for use with the Solaris Studio compilers. If selected as a BLAS, it must also be selected as LAPACK *via* (for Solaris Studio 12.2)

```
./configure --with-blas='-library=sunperf' --with-lapack
```

This has often given test failures in the past, in several different places. For R 2.13.0 it fails in 'tests/reg-BLAS.R', and on some builds, including for 'amd64', it fails in example(eigen).

C.5.1 Using gcc

If using gcc, ensure that the compiler was compiled for the version of Solaris in use. (This can be ascertained from gcc -v.) gcc makes modified versions of some header files, and several reports of problems were due to using gcc compiled on one version of Solaris on a later version.

Compilation for a 32-bit Sparc target with gcc 4.3.3 (the most recent version on OpenCSW) needed

```
CPPFLAGS=-I/opt/csw/include
LDFLAGS="-L/opt/csw/gcc4/lib -L/opt/csw/lib"
```

and for a 64-bit Sparc target

```
CC="gcc -m64"
F77="gfortran -m64"
```

```
CXX="g++ -m64"
FC=$F77
CPPFLAGS=-I/opt/csw/include
LDFLAGS="-L/opt/csw/gcc4/lib/sparcv9 -L/opt/csw/lib/sparcv9"
```

Note that paths such as '`/opt/csw/gcc4/lib/sparcv9`' may need to be in the `LD_LIBRARY_PATH` during configuration.

Tests with `gcc`[5] for '`x86`' and '`amd64`' have been less successful.

'`x86`' builds have built successfully but failed or gave incorrect results on various tests unless built without any optimization in `CFLAGS`: the issue seems to be that `gcc` and the system libraries differ over the use of SSE registers.

For '`amd64`' the builds have failed to complete in several different ways, most recently with relocation errors for '`libRblas.so`'.

C.6 AIX

We no longer support AIX prior to 4.2, and `configure` will throw an error on such systems.

Ei-ji Nakama was able to build under AIX 5.2 on '`powerpc`' with GCC 4.0.3 in several configurations. 32-bit versions could be configured with '`--without-iconv`' as well as '`--enable-R-shlib`'. For 64-bit versions he used

```
OBJECT_MODE=64
CC="gcc -maix64"
CXX="g++ -maix64"
F77="gfortran -maix64"
FC="gfortran -maix64"
```

and was also able to build with the IBM `xlc` and Hitachi `f90` compilers by

```
OBJECT_MODE=64
CC="xlc -q64"
CXX="g++ -maix64"
F77="f90 -cpu=pwr4 -hf77 -parallel=0 -i,L -O3 -64"
FC="f90 -cpu=pwr4 -hf77 -parallel=0 -i,L -O3 -64"
FLIBS="-L/opt/ofort90/lib -lhf90vecmath -lhf90math -lf90"
```

Some systems have `f95` as an IBM compiler that does not by default accept FORTRAN 77. It needs the flag '`-qfixed=72`', or to be invoked as `xlf_r`.

The AIX native `iconv` does not support encodings '`latin1`' nor '`""`' and so cannot be used. (As far as we know GNU `libiconv` could be installed.)

Fan Long reports success on AIX 5.3 using

```
OBJECT_MODE=64
LIBICONV=/where/libiconv/installed
CC="xlc_r -q64"
CFLAGS="-O -qstrict"
CXX="xlC_r -q64"
CXXFLAGS="-O -qstrict"
F77="xlf_r -q64"
AR="ar -X64"
CPPFLAGS="-I$LIBICONV/include -I/usr/lpp/X11/include/X11"
LDFLAGS="-L$LIBICONV/lib -L/usr/lib -L/usr/X11R6/lib"
```

On one AIX 6.x system it was necessary to use `R_SHELL` to set the default shell to be Bash rather than Zsh.

[5] specifically `gcc` 4.3.3 and `gcc` 3.4.6 from OpenCSW.

Kurt Hornik and Stefan Theussl at WU (Wirtschaftsuniversität Wien) successfully built R on a 'powerpc' (8-CPU Power6 system) running AIX 6.1, configuring with or without '--enable-R-shlib' (Ei-ji Nakama's support is gratefully acknowledged).

It helps to describe the WU build environment first. A small part of the software needed to build R and/or install packages is available directly from the AIX Installation DVDs, e.g., Java 6, X11, and Perl. Additional open source software (OSS) is packaged for AIX in '.rpm' files and available from both IBM's "AIX Toolbox for Linux Applications" (http://www-03.ibm.com/systems/power/software/aix/linux/) and http://www.oss4aix.org/download/. The latter website typically offers more recent versions of the available OSS. All tools needed and libraries downloaded from these repositories (e.g., GCC, Make, libreadline, etc.) are typically installed to '/opt/freeware', hence corresponding executables are found in '/opt/freeware/bin' which thus needs to be in PATH for using these tools. As on other Unix systems one needs GNU libiconv as the AIX version of iconv is not sufficiently powerful. Additionally, for proper Unicode compatibility one should install the corresponding package from the ICU project (http://www.icu-project.org/download/), which offers pre-compiled binaries for various platforms which in case of AIX can be installed via unpacking the tarball to the root file system. For full LaTeX support one can install the TeX Live DVD distribution (http://www.tug.org/texlive/): it is recommended to update the distribution using the tlmgr update manager. For 64-bit R builds supporting Tcl/Tk this needs to installed from the sources as available pre-compiled binaries supply only 32-bit shared objects.

The recent WU testing was done using compilers from both the GNU Compiler Collection (version 4.2.4) which is available from one of the above OSS repositories, and the IBM C/C++ (XL C/C++ 10.01) as well as FORTRAN (XL Fortran 12.01) compilers (http://www14.software.ibm.com/webapp/download/byproduct.jsp#X).

To compile for a 64-bit 'powerpc' (Power6 CPU) target one can use

```
CC ="gcc -maix64 -pthread"
CXX="g++ -maix64 -pthread"
FC="gfortran -maix64 -pthread"
F77="gfortran -maix64 -pthread"
CFLAGS="-O2 -g -mcpu=power6"
FFLAGS="-O2 -g -mcpu=power6"
FCFLAGS="-O2 -g -mcpu=power6"
```

for the GCC and

```
CC=xlc
CXX=xlc++
FC=xlf
F77=xlf
CFLAGS="-qarch=auto -qcache=auto -qtune=auto -O3 -qstrict -ma"
FFLAGS="-qarch=auto -qcache=auto -qtune=auto -O3 -qstrict"
FCFLAGS="-qarch=auto -qcache=auto -qtune=auto -O3 -qstrict"
CXXFLAGS="-qarch=auto -qcache=auto -qtune=auto -O3 -qstrict"
```

for the IBM XL compilers. For the latter, it is important to note that the decision for generating 32-bit or 64-bit code is done by setting the OBJECT_MODE environment variable appropriately (recommended) or using an additional compiler flag ('-q32' or '-q64'). By default the IBM XL compilers produce 32 bit code. Thus, to build R with 64-bit support one needs to either export OBJECT_MODE=64 in the environment or, alternatively, use the '-q64' compiler options.

It is strongly recommended to install Bash and use it as the configure shell, e.g., via setting CONFIG_SHELL=/usr/bin/bash in the environment, and to use GNU Make (e.g., via (MAKE=/opt/freeware/bin/make).

Further installation instructions to set up a proper R development environment can be found in the "R on AIX" project on R-Forge (http://R-Forge.R-project.org/projects/aix/).

C.7 Cygwin

The Cygwin emulation layer on Windows can be treated as a Unix-alike OS. This is unsupported, but experiments have been conducted and a few workarounds added. R 2.13.0 requires C99 complex type support, which is available as from Cygwin 1.7.8 (March 2011). However, the (new) implementation of `cacos` gives incorrect results, so `HAVE_CACOS` is undefined in '`src/main/complex.c`'.

Currently Cygwin 1.7.9-1 is unable to build R: the support for dynamic loading of DLLs is broken.

Only building as a shared library can possibly work,[6] so use e.g

```
./configure --disable-nls --enable-R-shlib
make
```

NLS does work if required, although adding '`--with-included-gettext`' is preferable. You will see many warnings about the use of auto-import.

Note that this gives you a command-line application using `readline` for command editing. The '`X11`' graphics device will work if a suitable X server is running, and the standard Unix-alike ways of installing source packages work. There was a bug in the '`/usr/lib/tkConfig.sh`' script in the version we looked at, which needs to have

```
TK_LIB_SPEC='-ltk84'
```

The overhead of using shell scripts makes this noticeably slower than a native build of R on Windows.

Even with Cygwin 1.7.8, not all the tests passed: there were incorrect results from wide-character regular expressions code and from sourcing CR-delimited files.

C.8 New platforms

There are a number of sources of problems when installing R on a new hardware/OS platform. These include

Floating Point Arithmetic: R requires arithmetic compliant with IEC 60559, also known as IEEE 754. This mandates the use of plus and minus infinity and `NaN` (not a number) as well as specific details of rounding. Although almost all current FPUs can support this, selecting such support can be a pain. The problem is that there is no agreement on how to set the signalling behaviour; Sun/Sparc, SGI/IRIX and '`ix86`' Linux require no special action, FreeBSD requires a call to (the macro) `fpsetmask(0)` and OSF1 requires that computation be done with a '`-ieee_with_inexact`' flag etc. On a new platform you must find out the magic recipe and add some code to make it work. This can often be done via the file '`config.site`' which resides in the top level directory.

Beware of using high levels of optimization, at least initially. On many compilers these reduce the degree of compliance to the IEEE model. For example, using '`-fast`' on the Solaris Studio compilers has caused R's `NaN` to be set incorrectly.

Shared Objects: There seems to be very little agreement across platforms on what needs to be done to build shared objects. there are many different combinations of flags for the compilers and loaders. GNU libtool cannot be used (yet), as it currently does not fully support FORTRAN: one would need a shell wrapper for this). The technique we use is to first interrogate the X window system about what it does (using `xmkmf`), and then override this in situations where we

[6] Windows DLLs need to have all links resolved at build time and so cannot resolve against '`R.bin`'.

know better (for tools from the GNU Compiler Collection and/or platforms we know about). This typically works, but you may have to manually override the results. Scanning the manual entries for `cc` and `ld` usually reveals the correct incantation. Once you know the recipe you can modify the file '`config.site`' (following the instructions therein) so that the build will use these options.

It seems that `gcc` 3.4.x and later on '`ix86`' Linux defeat attempts by the LAPACK code to avoid computations entirely in extended-precision registers, so file '`src/modules/lapack/dlamc.f`' may need to be compiled without optimization. Set the configure variable `SAFE_FFLAGS` to the flags to be used for this file. If configure detects GNU FORTRAN it adds flag '`-ffloat-store`' to `FFLAGS`. (Other settings are needed when using `icc` on '`ix86`' Linux, for example.)

If you do manage to get R running on a new platform please let us know about it so we can modify the configuration procedures to include that platform.

If you are having trouble getting R to work on your platform please feel free to use the '`R-devel`' mailing list to ask questions. We have had a fair amount of practice at porting R to new platforms . . .

Appendix D The Windows toolset

If you want to build R or add-on packages from source in Windows, you will need to collect, install and test an extensive set of tools. See http://www.murdoch-sutherland.com/Rtools/ for the current locations and other updates to these instructions. (Most Windows users will not need to build add-on packages from source; see Chapter 6 [Add-on packages], page 18 for details.)

We have found that the build process for R is quite sensitive to the choice of tools: please follow our instructions **exactly**, even to the choice of particular versions of the tools.[1] The build process for add-on packages is somewhat more forgiving, but we recommend using the exact toolset at first, and only substituting other tools once you are familiar with the process.

This appendix contains a lot of prescriptive comments. They are here as a result of bitter experience. Please do not report problems to the R mailing lists unless you have followed all the prescriptions.

We have collected most of the necessary tools (unfortunately not all, due to license or size limitations) into an executable installer named[2] '`Rtools213.exe`', available from http://www.murdoch-sutherland.com/Rtools/. You should download and run it, choosing the default "Package authoring installation" to build add-on packages, or the "full installation" if you intend to build R.

You will need the following items to build R and packages. See the subsections below for detailed descriptions.

- The command line tools (in '`Rtools*.exe`')
- The MinGW toolchain to compile C, Fortran and C++.

For installing simple source packages containing data or R source but no compiled code, none of these are needed. Perl is no longer needed to build R nor to install nor develop source packages.

A complete build of R including PDF manuals, and producing the installer will also need the following:

- LaTeX
- The Inno Setup installer
- (optional) `qpdf`

It is important to set your `PATH` properly. The installer '`Rtools*.exe`' optionally sets the path to components that it installs.

Your `PATH` may include '`.`' first, then the '`bin`' directories of the tools, the compiler toolchain and LaTeX. Do not use filepaths containing spaces: you can always use the short forms (found by `dir /x` at the Windows command line). Network shares (with paths starting `\\`) are not supported.

For example for a 32-bit build, all on one line,

```
PATH=c:\Rtools\bin;c:\Rtools\MinGW\bin;c:\MiKTeX\miktex\bin;
     c:\R\bin\i386;c:\windows;c:\windows\system32
```

It is essential that the directory containing the command line tools comes first or second in the path: there are typically like-named tools[3] in other directories, and they will **not** work. The ordering of the other directories is less important, but if in doubt, use the order above. If you install both the 32 bit and 64 bit toolchains from '`Rtools*.exe`', it should work to list

[1] For example, the Cygwin version of `make 3.81` fails to work correctly.

[2] for R 2.13.x.

[3] such as `sort`, `find` and `make`.

both in the path (as the automatic installer does), and this is needed if you want to install or build binary packages containing both architectures using the `--merge-multiarch` option as described in Section 6.3.1 [Windows packages], page 19. Other toolchains are more likely to conflict.

Our toolset contains copies of Cygwin DLLs that may conflict with other ones on your system if both are in the path at once. The normal recommendation is to delete the older ones; however, at one time we found our tools did not work with a newer version of the Cygwin DLLs, so it may be safest not to have any other version of the Cygwin DLLs in your path.

D.1 LaTeX

The '`MiKTeX`' (`http://www.miktex.org/`) distribution of LaTeX includes a suitable port of `pdftex`. The 'basic' version of '`MiKTeX`' almost suffices (the **grid** vignettes need '`fancyvrb.sty`'), but it will install the 15Mb '`lm`' package if allowed to (although that is not actually used). The '`Rtools*.exe`' installer does *not* include any version of LaTeX.

It is also possible to use the TeXLive distribution from `http://www.tug.org/texlive/`.

Please read Section 2.3 [Making the manuals], page 4 about how to make '`refman.pdf`' and set the environment variables `R_RD4DVI` and `R_RD4PDF` suitably; ensure you have the required fonts installed.

D.2 The Inno Setup installer

To make the installer package ('`R-2.13.1-win.exe`') we currently require Inno Setup 5.3.7 or later (including 5.4.x) from `http://jrsoftware.org/`. We use the Unicode version and recommend that version be used. This is *not* included in '`Rtools*.exe`'.

Copy file '`src/gnuwin32/MkRules.dist`' to '`src/gnuwin32/MkRules.local`' and edit it to set `ISDIR` to the location where Inno Setup was installed.

D.3 The command line tools

This item is installed by the '`Rtools*.exe`' installer.

If you choose to install these yourself, you will need suitable versions of at least `basename`, `cat`, `cmp`, `comm`, `cp`, `cut`, `date`, `diff`, `du`, `echo`, `expr`, `gzip`, `ls`, `make`, `makeinfo`, `mkdir`, `mv`, `rm`, `rsync`, `sed`, `sh`, `sort`, `tar`, `texindex`, `touch` and `uniq`; we use those from the Cygwin distribution (`http://www.cygwin.com/`) or compiled from the sources. You will also need `zip` and `unzip` from the Info-ZIP project (`http://www.info-zip.org/`). All of these tools are in '`Rtools*.exe`'.

Beware: 'Native' ports of make are **not** suitable (including that called 'MinGW make' at the MinGW SourceForge site). There were also problems with other versions of the Cygwin tools and DLLs. To avoid frustration, please use our tool set, and make sure it is at the front of your path (including before the Windows system directories). If you are using a Windows shell, type `PATH` at the prompt to find out.

You may need to set the environment variable `CYGWIN` to a value including '`nodosfilewarning`' to suppress messages about Windows-style paths.

D.4 The MinGW toolchain

This is the only step which differs between 32- and 64-bit builds. Technically you need more than just a compiler so the set of tools is referred to as a 'toolchain'.

In each case you need to put the '`bin`' directory of the toolchain early in your `PATH`: if you use the recommended toolchains you can have both in the path (in either order), and need to if you want to install 32/64-bit packages.

D.4.1 32-bit toolchain

Note that the recommended toolchain for R >= 2.12.0 differs from that for earlier versions, and there are binary incompatibilities, especially where C++ code is involved. So we recommend that all packages are built with the current toolchain.

The 32-bit build for this version of R is set up to use `gcc 4.5.0`. The toolchain included in '`Rtools213.exe`', and was collected from the files at http://sourceforge.net/projects/mingw/files/, specifically

```
binutils-2.20.51-1-mingw32-bin.tar.lzma
gcc-c++-4.5.0-1-mingw32-bin.tar.lzma
gcc-core-4.5.0-1-mingw32-bin.tar.lzma
gcc-fortran-4.5.0-1-mingw32-bin.tar.lzma
libgcc-4.5.0-1-mingw32-dll-1.tar.lzma
libgmp-5.0.1-1-mingw32-dll-10.tar.lzma
libgomp-4.5.0-1-mingw32-dll-1.tar.lzma
libmpc-0.8.1-1-mingw32-dll-2.tar.lzma
libmpfr-2.4.1-1-mingw32-dll-1.tar.lzma
libssp-4.5.0-1-mingw32-dll-0.tar.lzma
mingwrt-3.18-mingw32-dev.tar.gz
mingwrt-3.18-mingw32-dll.tar.gz
w32api-3.15-1-mingw32-dev.tar.lzma
```

with the import libraries in '`lib/gcc/mingw32/4.5.0`' moved to a subdirectory '`dlls`' so that '`libgfortran`' and '`libstdc++`' are linked statically.

It is also possible to use the 32-bit toolchain from the `mingw-w64` project and options to use that are documented in file '`MkRules.dist`'.

D.4.2 64-bit toolchain

Several versions of the MinGW-w64 toolchain are available: use one of them to replace (or supplement) the 32-bit toolchain's '`bin`' directory in your path. Then copy file '`src/gnuwin32/MkRules.dist`' to '`src/gnuwin32/MkRules.local`' and edit it to set `WIN=64` and `BINPREF64` and `SYMPAT64` appropriate to your toolchain, then make R in the usual way.

The toolchain we use is technically a cross-compiler: the tools run under 32-bit Windows but produce code to run under 64-bit Windows.[4] This is based on files from http://sourceforge.net/projects/mingw-w64/files/Toolchains%20targetting%20Win64/Automated%20Builds/ and has a name like '`mingw-w64-1.0-bin_i686-mingw_20100917.zip`', including a pre-release of `gcc 4.5.2`. However, this has been patched to use static Fortran and C++ runtime libraries: a suitable toolchain is included in '`Rtools213.exe`'.

NB: 64-bit toolchains after late April 2010 at that site generate objects with symbols without a leading underscore. Only such toolchains are supported (and not that used to build R 2.11.1 on 64-bit Windows).

Other toolchains are available: for example under http://sourceforge.net/projects/mingw-w64/files/Toolchains%20targetting%20Win64/Personal%20Builds/ contains a native x64 Windows toolchain by 'sezero' that may be useful when building external software.

[4] As the R build process itself runs R, it has to be done under 64-bit Windows. So does installation of source packages unless '`--no-test-load`' is used.

D.5 Useful additional programs

The process of making the installer will make use of `qpdf` to compact some of the package vignettes, if it is available. MinGW binaries of `qpdf` are available from http://sourceforge.net/projects/qpdf/files/. Set the path to the `qpdf` installation in file '`MkRules.local`'.

Developers of packages will find some of the 'goodies' at http://www.stats.ox.ac.uk/pub/Rtools/goodies useful.

There is a version of the `file` command that identifies the type of files, and is used by `Rcmd check` if available.

The file '`xzutils.zip`' contains the program `xz` which can be used to (de)compress files with that form of compression.

Function and variable index

Concept index

Environment variable index