from which I gather that the (already rather extended) dialect of SQL can be extended by user-defined functions that call R to compute new columns from old ones.

We are exploring R–DBMS interfaces in data mining applications. There databases can be large but are perhaps growing slower than computing power. For example, insurance databases already cover around 30 million drivers. The idea is to use the DBMS not just to extract subsets for analysis but also perform cross-tabulations and search for exceptions.

Brian D. Ripley
University of Oxford, UK
ripley@stats.ox.ac.uk

Rcgi 4: Making Web Statistics Even Easier

by M.J. Ray

Webservers allow browsers to run a selection of programs chosen by the webmaster, via the Common Gateway Interface which defines how inputs are passed from browser to program and output passed back again. Naturally, it is possible to make R one of the available programs by using a "wrapper script" to translate between standard R input and outputs and CGI. Rcgi is such a script and has just been revised for a new, more powerful and easier-to-install release.

Benefits

Rcgi has been developed in response to a perceived need at the University of East Anglia. Even though we can provide students with their own copies of the R software for free, there are two principal reasons why it is still useful to provide access to our installation over the internet.

The first is that not all students have their own workstations. While this may change eventually, many students use the campus's open access computing facilities. Our department can manage software on only a small proportion of these machines, so for the others a web interface is simpler to access and insulates us from configuration changes beyond our control. As long as the computer can still run a web browser, they can use Rcgi.

Feedback from students on our third-year course (who were the first to use the system) suggests that the web front-end is popular partly because of its "batch mode" of operation, running some commands and returning the output, together with the commands for editing and resubmission.

The second and increasingly important benefit of Rcgi is the ability for the lecturer to provide worked examples to the students with the option of leaving spaces for the students to contribute their own data to the examples. Rather than having to write their own CGI programs for each example, lecturers need only write the HTML for the page (which many know already) and the program in the R language.

Example

Take the following snippet of R code, which defines a vector of numbers and generates the basic summary statistics for them:

```
test <- c(1,45,2,26,37,35,32,7,
          4,8,42,23,32,27,29,20)
print(summary(test))
```

which has the output

```
Min. 1st Qu  Median  Mean 3rd Qu.  Max.
1.00   7.75   26.50 23.13   32.75 45.00
```

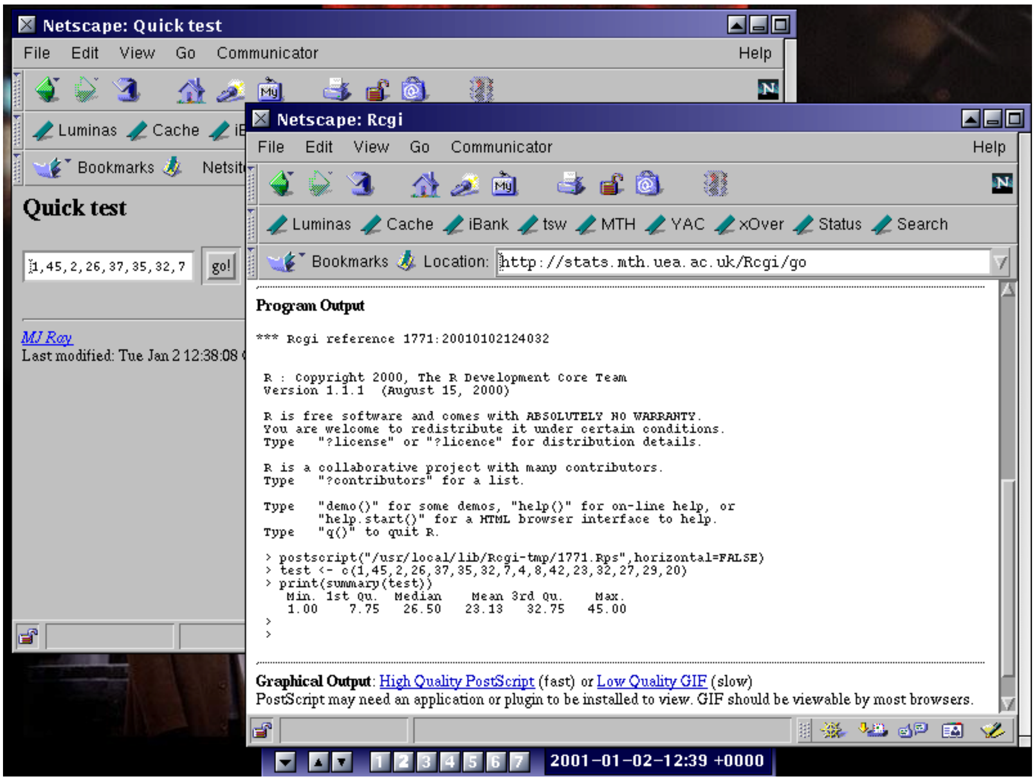
For an elementary statistics course, the lecturer may wish to provide this as a worked example, but allow the students to change the values in the "test" list. The HTML code to do this is:

```
<form  method='post'
       action='/cgi-bin/Rcgi'>
<input type='hidden' name='script'
       value='test <- c(' />
<input type='text' name='script'
value='1,45,2,26,37,35,32,7,
       4,8,42,23,32,27,29,20' />
<input type='hidden'
       name='script' value=')
       print(summary(test))
       ' />
<input type='submit' value='go!' />
</form>
```

and a view of the data entry page and the results returned from Rcgi are included in figure 4. Hopefully by the time that this article appears, this example will be back online at the Rcgi site (address below).

Note that the code executed is displayed directly below the output, offering the student the chance to examine and modify the R program. Hopefully, they can learn from this experimentation in a similar way to having the entire system on their own computer.

The most commonly suggested point for improvement was the installation procedure, which was previously completely manual. The usual Unix make program is used to install the new release, with a script that attempts to locate the various



Netscape: Quick test
File Edit View Go Communicator Help
Luminas Cache
Bookmarks Netsite
Quick test
1,45,2,26,37,35,32,7
go!
MJ Ray
Last modified: Tue Jan 2 12:38:08
Netscape: Rcgi
File Edit View Go Communicator Help
Luminas Cache iBank tsw MTH YAC xOver Status Search
Bookmarks Location: http://stats.mth.uea.ac.uk/Rcgi/go
1 2 3 4 5 6 7
2001-01-02-12:39 +0000

Program Output

```
*** Rcgi reference 1771:20010102124032


 R : Copyright 2000, The R Development Core Team
 Version 1.1.1  (August 15, 2000)

 R is free software and comes with ABSOLUTELY NO WARRANTY.
 You are welcome to redistribute it under certain conditions.
 Type   "?license" or "?licence" for distribution details.

 R is a collaborative project with many contributors.
 Type   "?contributors" for a list.

 Type   "demo()" for some demos, "help()" for on-line help, or
        "help.start()" for a HTML browser interface to help.
 Type   "q()" to quit R.

 > postscript("/usr/local/lib/Rcgi-tmp/1771.Rps",horizontal=FALSE)
 > test <- c(1,45,2,26,37,35,32,7,4,8,42,23,32,27,29,20)
 > print(summary(test))
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    1.00    7.75   26.50   23.13   32.75   45.00
 >
 >
```

Graphical Output: High Quality PostScript (fast) or Low Quality GIF (slow)
PostScript may need an application or plugin to be installed to view. GIF should be viewable by most browsers.

Figure 4: Screenshots of the results of the example HTML

programs it requires and build a module called 'Rcgi::SystemPrograms' detailing the location of some system programs required. The installer is still advised to review the scripts to make sure the details found are correct, though.

Other forthcoming improvements include better security checking (which is fairly basic at the moment) and better load balancing and session management. However, it is already a useful and popular system with a growing user community, with a mailing list available for help and development. The download site is http://stats.mth.uea.ac.uk/Rcgi/.

M.J. Ray
University of East Anglia, Norwich, UK
mjr@stats.mth.uea.ac.uk

Omegahat Packages for R

by John M. Chambers and Duncan Temple Lang

Overview

As part of the Omegahat project (http://www.omegahat.org/), we are developing a collection of packages to support new directions in programming in the S language, as implemented in R (http://www.r-project.org/) or in S-Plus (http://www.insightful.com/products/splus/). The packages we describe here illustrate ways to communicate between R and other languages and applications. You may find the packages useful if you want to access code from R that is written in another language, or call R functions from another programming language or application. This includes the notion of embedding R in such a package, or vice versa.

We also comment on future approaches to writing such software, with particular reference to making it available in both R and S-Plus.

We start with a very brief description of each of the available packages. More details can be found at the Omegahat web site (http://www.omegahat.org/). You can find a slightly lengthier description of how the packages relate to each other and the current tools in the sections below. We also plan to describe individual packages in future editions of the R newsletter. Please note that all of these packages are work-in-progress, at various different stages of evolution. We would be very appreciative of any com-