Despite this power, the main use I make of regular expressions is in simple script processing. For example, in porting a package written for S-PLUS recently I wrote a function to find all the SGML documentation files in the package, strip the version control comments from the beginning of each file, and feed them to `R CMD Sd2Rd` to make Rd files. This could have been done in Perl or probably with a simple shell script, but it seemed easier just to use R.

An interesting exercise for the reader: Creating a namespace for an existing package that uses S3 methods requires registering the methods in the `NAMESPACE` file. That is, for a function such as `print.coxph`, the `NAMESPACE` file should contain `S3method(print, coxph)`. How would you write a function that looked for methods for a list of common generic functions and created the necessary `S3method` calls? You would probably need to use the functions either `strsplit` and `paste` or `regexpr` and `substr`, and would need to remember that the generic name need not be a single word (`is.na.Surv` springs to mind). Some hand editing would still be required, for example to determine whether `t.test.formula` was a method for `t` or `t.test` or neither.

References

Leisch, F. (2002) Sweave, Part I: Mixing R and LaTeX. *R News* 2(3): 28–31.

Thomas Lumley
Biostatistics, University of Washington
tlumley@u.washington.edu

Recent Events

R at the Statistics Canada Symposium

R was one of the software packages demonstrated at the 2003 Statistics Canada Symposium, an annual event organized by Statistics Canada and presented near Ottawa, Ontario. The event is attended largely by the methodologists (statisticians) who work here and at other national statistical agencies, such as the US Census Bureau. This year's symposium was the 20th edition of the conference. The four day event had two days of software demonstrations with different packages being shown each day. Out of the 10 different packages being demonstrated, 7 were developed at Statistics Canada for very specific tasks.

Although the software demonstrations were hidden in a room out of sight from the main proceedings, there were still several curious visitors to the R demo computer. We showed some basic demos of R at work in sampling and in simulation problems, emphasizing both the simplicity of the code and the results that can be produced. We also used the built-in graphics demo, which impressed many people. As most people working for Statistics Canada use SAS, many of our discussions were about why one would use R instead of SAS (besides the fact that R is free). We had many discussions on programming structure, creating plots, dealing with large amounts of data and general ease of use. All our visitors, when shown the software, seemed to like it and were not adverse to the idea of using it instead of SAS. In fact, they couldn't understand why we were not already using it more widely.

Krisztina Filep
Social Survey Methods Division
Statistics Canada
Ottawa, Ontario, Canada
Krisztina.Filep@statcan.ca