S-PLUS Package System

Stephen Kaluzny

spk@insightful.com

Insightful Corp. Seattle, WA USA

Overview

- Introduction
- Purpose
- Implementation
- CSAN
- Porting Packages from R
- Issues / Future Work
- Wrap Up

Introduction

- The R package system has been extremely successful
  - one of the key features that has contributed to R's growth
- Next release of S-PLUS will include a package system not unlike the package system in R
- Goal is a common system for distributing new functionality and data in R and S-PLUS

History

- We have wanted a package-like system in S-PLUS for years
- Original StatLib "package" idea - Unix specific with shar archives
- The CHAPTER system with SV4 was never quite fully implemented
  - mixed software development with distribution
  - Bell Labs version was Unix specific
  - S-PLUS chapter system on Windows implemented and documented poorly

Purpose of S-PLUS Package System

- Allow users to share code and data
- Allow researchers to "publish" their work
- Allow users to tap into the experts of the R/S community
- Replace StatLib shar system
- Replace CHAPTER system

S-PLUS Package System Implementation

- Adapted from R Package System
- Same functions e.g.
  - `install.packages()`
  - `available.packages()`
  - `update.packages()`
- Same scripts - Perl, shell e.g.
  - INSTALL, build, check
- Package related functions implemented in Rcompat library section (package)
  - Installed separately from S-PLUS
  - Bootstrap the system with `getRcompat()`

Functions Added for Package System

- Low level functions, many for package manipulation
- Adding additional arguments to existing functions
- Renaming of existing functions
- Additional functions as we come across them in porting packages

File / Directory Functions Added

- dir.create
- file.splitpath
- dirname / basename
- list.files / dir
- file.create / file.remove
- file.copy / file.append / file.rename
- file.info
- file.access
- system.file
- file.show

Character / String Functions Added

- toupper / tolower
- squote / dquote
- gettext / ngettext / gettextf
- sprintf
- formatC - a wrapper for sprintf that mimics R's formatC
- strsplit

Misc. Functions Added to S-PLUS

- log1p
- expm1
- which.min, which.max
- colnames / rownames
- is.R
- stopifnot
- upper.tri

Functions with New Name To Match R

- Sys.getenv - same as getenv
- Sys.putenv - different argument list than setenv
- Sys.getpid - returns process.id()$pid
- getOption
- Sys.sleep - same as sleep

Modified Functions

- grep - added value argument and pass ... on to regexpr
- mkdir / unlink - added recursive argument
- gsub / substituteString / regexpr - added ignore.case, extended, fixed, and perl arguments
- file.exists - added ...
- sapply - added USE.NAMES argument
- scale - added attributes scaled:center and scaled:scale

Help System for Packages

- Uses .Rd format
- Convert Rd to html on Unix/Linux
  - .Rd to sgml and then continue with current help system
- Convert Rd to html and then create CHM on Windows
- (New help system is needed)

Windows Package System

- Expect most Windows users will consume, not produce packages
  - Binary package installations currently work
- Source builds supports Microsoft Visual C and Visual Fortran
- Also support MinGW compiler
- Requires other software as R on Windows
  - Perl
  - Cygwin tools
  - HTML Help Compiler

CSAN

- Comprehensive S-PLUS Archive Network
- Modeled after CRAN
- Currently only available at csan.insightful.com
- Plans for mirrors
- Current Packages
  - calibration
  - glars
  - gtools
  - randomForest
  - rmeta
  - rpart

Packages and S-PLUS Workbench

- S-PLUS now includes an IDE based on Eclipse called the S-PLUS Workbench
- Next release will include an S source debugger and profiler in the workbench
- Workbench will include ability to build and install packages

Graphics Enhancements in S-PLUS 8.0

- Adding support for named colors (color name or RGB)
  - associated functions colornames, rgb
  - color palettes
  - backward compatible - col=n works as before
- Adding for support for vectorized high level parameters
  - pch, col, cex, etc
- Makes porting calls to graphics in packages easier

Portable Packages

- Packages that work in both R and S-PLUS
  - source level only
- Use `is.R()` to if-def code for R and S-PLUS
- Packages could be ported to S-PLUS or to R
  - Research libraries from Insightful
  - StatLib code
  - S-PLUS users have asked for a system to create a single package for both systems
- Some of the CRAN packages are already portable packages

Porting Packages From R to S-PLUS

- Problem Areas
  - lexical scoping
  - namespaces
  - environments
  - Internal R code (.Call / .Internal)
  - Rda binary data
  - Unavailable interfaces e.g. tcl/tk

Tools To Help Porting Packages

- `set.parse.mode("R")`
  - allows _ in variable names
- `unresolvedGlobalReferences()`
  - finds most scoping problems in apply family calls
- Include files that map to S-PLUS includes:
  - R.h
  - Rmath.h
  - R-ext/*.h

Porting Your Package to S-PLUS

- Free copy of S-PLUS for porting work
- Mail listserver for discussion of porting issues
- Support from Insightful
- If interested, please contact me

S-PLUS 8.0 Timeline

- S-PLUS 8.0 will include the S-PLUS Package System
- Currently in Beta 1
- Beta 2 starts in mid-July
- FCS October 2006

DESCRIPTION file

- New tags
  - Dialect - R, Splus
  - Splus maintainer
- Version numbers
  - R Version
  - Splus Version

Future Development

- Single package archive or sharing of packages
- Meta data for packages
  - Package rating
  - Package reviews
  - Requirements - C / Fortran
  - License classes
  - Application area

Wrap Up

- Packages in R and S-PLUS
  - Packages have been a good thing for R
  - Packages will be a good thing for Splus
  - Packages in both R and S-PLUS will be a good thing for the statistical community