

Mayday RLink – The best of two worlds

Florian Battke[†], Stephan Symons, and Kay Nieselt

Center for Bioinformatics Tübingen, University of Tübingen, Sand 14, 72076 Tübingen, Germany

[†] Corresponding author: battke@informatik.uni-tuebingen.de

Keywords: microarray analysis, visual analytics, R integration

DNA Microarrays are the standard method for large scale analyses of gene expression and epigenomics. Analysis software must keep pace with the increasing complexity of generated data. MAYDAY [1] is a free and flexible graphical workbench for visualization and analysis of microarray data. It is written in Java and can be used as fully-functional WebStart application on every major computing platform without any installation. New challenges can swiftly be met due to MAYDAY's plugin interface. Currently, MAYDAY includes a large variety of plugins for visual data exploration, clustering, machine learning and classification, as well as Gene Set Enrichment Analysis. MAYDAY can import data from several file formats, database connectivity is included for efficient data organization. Numerous interactive visualization tools, including box plots, profile plots, principal component plots, our enhanced heatmap [2], the use of metadata to enhance plots as well as the possibility to create publication quality images make MAYDAY a power analysis tool for microarray data.

MAYDAY offers an intuitive interface to work on experimental data and many methods for interactive *visual* data exploration. R, on the other hand, can be used to quickly create unconventional visualizations, it offers an extremely versatile shell allowing e.g. fast filtering of data by arbitrarily complex criteria as well as methods for interactive *computational* exploration of data. The wealth of R packages available, such as those from the BioConductor project, is an immensely valuable resource and one of the reasons R is used throughout the microarray community.

With RLink we provide an interactive and integrated approach to harness the power of R within the framework provided by MAYDAY, further increasing the power of our visual analytics platform. Using the `rJava` package [3] for R, we have integrated an interactive R shell into MAYDAY. Within this shell, users can directly work on MAYDAY's core data structures, and apply methods provided by R or R packages such as LIMMA. Results can either be passed back to MAYDAY in the form of new datasets or attached to the original data as meta-information. Plugins offered by MAYDAY can also be called from the R shell allowing sophisticated analyses by combining the methods both programs offer, manual as well as scripted. The implementation hides the syntactic complexities involved in using native Java objects within R, all interaction is done via R objects that behave like normal R vectors resp. matrices.

The combination of both programs integrates the best of both worlds providing researchers with a background in R the opportunity to quickly test new hypotheses or find out details about their data that MAYDAY's user interface does not provide. While for these needs new plugins could of course be developed for MAYDAY, this is often not desirable for one-time analyses specific to only one particular dataset. Thus, using the RLink shell alongside MAYDAY's graphical user interface, scientists can quickly gain a deeper understanding of their data.

MAYDAY and RLink are available at <http://www.zbit.uni-tuebingen.de/pas/mayday/>

References

- [1] J Dietzsch, et al. (2006) MAYDAY – a microarray data analysis workbench. *Bioinformatics* 22:1010.
- [2] N Gehlenborg, et al. (2005) A framework for visualization of microarray data and integrated meta information. *Information Visualization* 4:164.
- [3] S Urbanek. `rJava`: Low-level R to Java interface