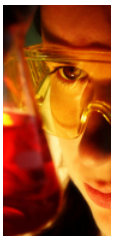
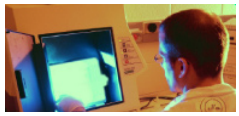



Vol. 1 No. 4
April-May 2008

Bioinformatics Organization

In this issue

Franklin Award ceremony

The 2008 Benjamin Franklin Award in the Life Sciences was presented to Robert Gentleman of the Fred Hutchinson Cancer Research Center during the 2008 Annual Meeting of the Bioinformatics Organization.

Robert Gentleman (left) with J.W. Bizzaro

The Meeting took place at the Bio-IT World Conference & Expo in Boston, late April. Organization president J.W. Bizzaro presented the Award to Robert during the plenary keynote session, which was attended by several hundred people.

Robert's laureate seminar included acknoweldgments to the contributors of the R and Bioconductor projects for which he received the Award. He also spoke about version 2.2 of Bioconductor, which includes some performance enhancements for handling high-throughput sequence data and some string-handling features.

Robert also spoke about "reproducible research," a topic that addresses the reproducibility of computational analyses that have been published in scientific journals: an approach that involves packaging all of the data, software and parameters needed to produce the figures provided in articles. He also addressed "literate programming," which is a means of blending text with source code for the production of articles. These efforts of his in sharing data-transformation methods, as well as the underlying data, are also why he received the Award this year.

New to the Board of Directors:

Prashanth Suravajhala

At the 2008 Bioinformatics Organization Annual Meeting, the Board of Directors elected Prashanth Suravajhala to join the Board.

Prashanth Suravajhala is a PhD fellow in Prof. Lene Juel Rasmussen's laboratory at Department of Science, Systems and Models, Roskilde University, Denmark. After he finished his Masters in Biotechnology, he was trained in Centre for Cellular and Molecular Biology, Hyderabad, India. He later worked for Cenitech and Biosolutions, Hyderabad as a trainer and researcher there before he joined the present lab. Computational Biology fascinated him since his early under graduate studies. His interests are Systems Biology of hypothetical proteins and Mitochondrial Biology. Prash has been a member of various organizations, to mention a few: Bioinformatics Society of India, ISCB Student Council. He has been a member of the Bioinformatics Organization since 2001. His team, through Bioclues.org and Roskilde University, developed a database of hypothetical proteins human (Hypo, hype and "hyp" human proteins, Bioinformation 2007, 2(1):31-33), which is due to surpass the recent annotation of human draft (In 2006, he founded Bioclues.org through which he mentors students in the area of Computational Biology by providing them virtual projects.)

Besides his zeal in learning how to accomplish research tasks, Prash loves traveling while he mulls networking as one of the traits required for an active researcher to be a go-getter in this competitive world.

His profile can be found at http://ruc.dk/~prash

Unveiling Pipet (part three)

J.W. Bizzaro

We've been looking at the use of Pipet to implement the example "`ls -l ./`" Unix command. In this issue of the newsletter, I'll introduce the concept of "cell compilation" (CC), which is a fundamental feature of Pipet.

An important consideration in Pipet is that you wouldn't want Web users entering arbitrary commands and options on your computer. So, at this point, we can add various attributes to the `<cell>` and `<value>` elements in the PCML, such as whether or not the user can change or even see a value.

As seen in Figure 4, we can hide everything but the option to specify the directory name. We can also have an attribute that sets this value as an item in a list (with checkboxes or radio buttons) so that the user cannot enter an arbitrary value at all.

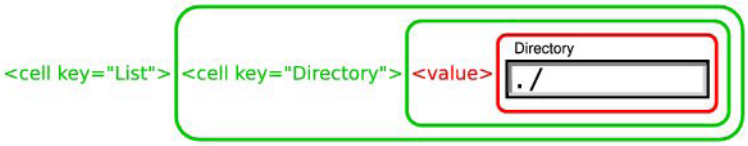


Figure 4: The GUI skin for the directory listing command, with only the directory name being visible. The red and green outlines are used here to represent the nesting of the value within the cells, respectively, and also the key-value relationship in PCML.

To accomplish this, Pipet "compiles" cells from

Continued on page 3.

Continued from page 2.

their PCML descriptions: when the PCML tree is traversed, the character data in the `<value>` elements are summed to determine key-value relationships, and then the element attributes are applied to determine skin and script types (affecting both element and contents). The former operation is performed in preorder fashion (parent element first), and the latter is performed in postorder fashion (child element first). And the process is pretty much the same whether cells are nested directly in the same PCML file or called via the "Wiki template" method mentioned earlier.

In the next installment, we'll look into the use of Pipet to create research and analysis groups. If you have any questions or would like to get involved in Pipet, please contact jeff@bioinformatics.org.

Career Center

Job search

Looking for a job in the bioinformatics field?

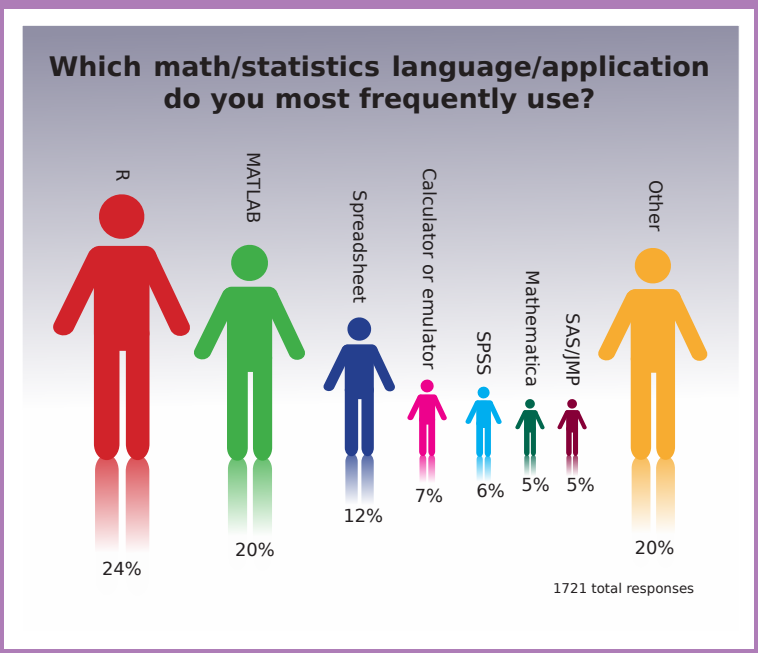
The Bioinformatics Section at Lundbeck forms an integrated part of Discovery Biology Research and supports bioinformatics analysis of a wide variety of data such as HTS and genomics data. The main task of the department is to provide novel perspectives and analysis to frontline central nervous system research through analysis of high content data analysis. The team consists of two PhDs and interacts with students and all departments within biological research.

High contents data acquisition methodologies such as microarray expression analysis, genome wide association analysis, proteomics analysis, epigenetic profiling are increasingly used in our exciting early discovery research to understand the underlying mechanisms for complex neurological and neuropsychiatric disorders such as Alzheimers disease, Parkinsons disease, Schizophrenia and mood disorder. A key task for the Bioinformatics section is to link complex analysis to systems biology and an understanding of pathways involved in CNS disorders.

You will be responsible for applying a long range of bioinformatics analysis supporting early target identification and validation projects. In collaboration with our drug discovery informatics department you will be involved in the establishment of an infrastructure for the storage and analysis of high contents biological data. You will be directly involved in drug discovery projects.

You will advise colleagues on novel bioinformatics analysis methods and algorithm, perform the detailed analysis and communicate the results to progress our research projects.

Full announcement:
http://bifx.org/forums/forum.php?forum_id=6534



Upcoming events

CS101B R for Biologists, Level 1; Monday, June 2 » Friday, June 6

Taught in the context of biological research, this course helps biologists learn how to use the statistical scripting language R for data analysis. Armed with some knowledge and hands-on experience with a programming or scripting language, scientists taking a CS101 course at Bioinformatics.Org will be able to perform basic software development tasks and phrase research questions in the context of the language.

http://wiki.bioinformatics.org/CS101B_R_for_Biologists,_Level_1

CS102A Perl for Biologists, Level 2; Monday, June 9 » Friday, June 13

Taught in the context of biological research, this course shows biologists how to use the scripting language Perl to automate certain tasks. It is a continuation of CS101A Perl for Biologists, Level 1 and covers advanced topics and projects.

http://wiki.bioinformatics.org/CS102A_Perl_for_Biologists,_Level_2

BI201C dChip for Gene Expression & SNP Genotyping; Monday, June 16 » Friday, June 20

This course will teach the workings of dChip, for doing Gene Expression and SNP Genotyping analysis. We will cover topics such as importing arrays, performing normalization, model based expression calculations, gene and snp filtering, clustering, linkage and LOH analysis, and much more.

http://wiki.bioinformatics.org/BI201C_dChip_for_Gene_Expression_and_SNP_Genotyping

CS102B R for Biologists, Level 2; Monday, June 23 » Friday, June 27

Taught in the context of biological research, this course helps biologists understand the data analysis and visualization language R. It is a continuation of CS101B R for Biologists, Level 1 and covers advanced topics and projects.

http://wiki.bioinformatics.org/CS102B_R_for_Biologists,_Level_2

CS103A Perl for Biologists, Level 3; Monday, June 30 » Thursday, July 3

This course aims at teaching Biologists how to use Perl as a programming language to automate certain tasks in their research. Advanced topics in regular expressions, objects, modules will be covered, along with tips and tricks to fine-tune programs and resolve bugs.

http://wiki.bioinformatics.org/CS103A_Perl_for_Biologists,_Level_3

BI211A Protein-Protein Interactions; Monday, July 7 » Friday, July 11

This course is a starter with introduction to various tools and some examples to find true positive interactors from web searches and web interfaces. The exercises and take-home messages should probably encourage participants to give some thought to the topics covered and hopefully build some excitement about the top-down approach of systems biology involving protein-protein interactions.

http://wiki.bioinformatics.org/BI211A_Protein-Protein_Interactions

Bioinformatics Organization

17 Pope Street
Hudson, MA 01749
USA