Workshop on Graphical models for Genomics

September 18th, 2007

GIGA-R, Sart-Tilman Campus, B4000 Liège, Belgium

The research unit in Bioinformatics and Modeling of the GIGA-Research centre invites you to attend a half-day workshop which aims at presenting recent results concerning the use of graphical probabilistic models in genomics and fostering the discussions among biologists and computer scientists about relevant research topics or directions in this area.

Program:

8:30 AM	Welcome Louis Wehenkel, GIGA-R/EECS, Liège
9:00 AM	**Reverse-engineering gene co-expression networks as multivariate normal models from microarray data.** *Robert Castelo. Pompeu Fabra University (UPF), Barcelona*
10:00 AM	Coffee break
10:30 AM	**Graphical models for genetics applications.** *Dan Geiger. Technion, Haifa*
11:30 AM	Open discussions
12:30 AM	Closing

Important information:

Please register by email to Raphael.Maree@ULg.ac.be, with 'Subject: GMG07 Registration'. We can only accommodate for a limited number of attendees.

The workshop will take place in the room R3 at the Montefiore Institute (B28, Parking P32). Access maps are available on http://www.montefiore.ulg.ac.be/location.php.

Sandwiches and drinks will be served after the workshop.

Visit our website: http://www.montefiore.ulg.ac.be/bioinformatics/

Université de Liège

Workshop on Graphical models for Genomics

September 18th, 2007

GIGA-R, Sart-Tilman Campus, B4000 Liège, Belgium

REVERSE-ENGINEERING GENE CO-EXPRESSION NETWORKS AS MULTIVARIATE NORMAL MODELS FROM MICROARRAY DATA, *by Robert Castelo.*

A straightforward and widely used approach to reverse-engineer gene co-expression networks from microarray data is to estimate Pearson correlation coefficients (PCCs) and their significance among every pair of genes. Then, by using some model selection strategy, like False Discovery Rate (FDR), one builds a co-expression network out of the subset of selected significant PCCs. Such a bivariate model, however, can potentially hide indirect associations between the genes resulting in a substantial number of spurious co-expression interactions in the resulting network. A more sensible approach would be to try to select a multivariate normal model of the network that would provide us the so-called full-order partial correlation coefficients (PACs) which can describe an association between two genes taking into account the rest of them (i.e., PACs become zero when the association is not direct). However, the dimension of microarray data, where the number of genes is much larger than the number of experiments, precludes the direct use of standard multivariate techniques. We have developed a new procedure based on limited-order PACs that deals with microarray data based on the search for an appropriate lower-order q of the associations. If such a q is found then, because we have reduced the initial dimension of the problem, we can estimate PACs and their significance with standard methods. We show that selecting associations using a FDR strategy on the significance levels of the PACs we attain a lower proportion of spurious co-expression associations than applying the same FDR strategy on PCCs.

GRAPHICAL MODELS FOR GENETICS APPLICATIONS, *by Dan Geiger.*

This talk introduces several algorithms and software tools for mapping disease genes. These applications extend the need for efficient exact and approximate inference algorithms for graphical models. In particular, I will focus on three mapping methods, each provides mapping results in different resolution and using different methods but all three use graphical models as the underlying engine.