Group of Computational Biology
http://zcam.tsinghua.edu.cn/Compbio/

Metabolite Identification through Machine Learning

Juho Rousu
Aalto University

Juho Rousu is Associate Professor at the Department of Information and Computer Science, Aalto University, where he leads the research group Kernel Machines, Pattern Analysis and Computational Biology. Juho has deep expertise on algorithms and machine learning method development, and their application to systems biology, resulting in a significant number of original computational methods, including metabolic flux analysis, metabolic network reconstruction, enzyme function prediction, and mass spectrometric data analysis. Current research focus of Dr. Rousu's group is in silico analysis and reconstruction of biological networks and prediction of structured data. In education, Dr. Rousu is in charge of the euSYSBIO MSc programm in Computational and Systems Biology at Aalto, a program organized in partnership with KTH Royal Institute of Technology, Instituto Superior Tecnico Lisbon and Tsinghua University.

Abstract : Metabolite identification from tandem mass spectrometric data is a key task in metabolomics. Various computational methods have been proposed for the identification of metabolites from tandem mass spectra. Fragmentation tree methods explore the space of possible ways the metabolite can fragment, and base the metabolite identification on scoring of these fragmentation trees. Machine learning methods have been used to map mass spectra to molecular fingerprints; predicted fingerprints, in turn, can be used to score candidate molecular structures. Here I will present our recent work, combining fragmentation tree computation with kernel-based machine learning to predict molecular fingerprints and identify molecular structures. We introduce a family of kernels capturing the similarity of fragmentation trees, and combine these kernels using recently proposed multiple kernel learning approaches. Experiments on two large reference datasets show that the new methods significantly improve molecular fingerprint prediction accuracy. These improvements result in better metabolite identification, doubling the number of metabolites ranked at the topposition of the candidates list.
The work is joint with Huibin Shen, Kai Duehrkop and Sebastian Boecker and will be presented at the ISMB-2014 conference in Boston in July.

The talk will run at **10am** on **June 27(Friday)**, in **the Auditorium on the third floor of the Physics Building** of Tsinghua University.