Merja Anneli Oja

Work address

Helsinki Institute of Information Technology   Phone: +358-9-451 4351
Lab. of Computer and Information Science   Fax: +358-9-451 3277
Helsinki University of Technology (TKK)   `merja.oja@tkk.fi`
P.O. Box 5400, 02015 TKK, Finland   `http://www.cis.hut.fi/moja/`

Home address

Rummunlyöjänkatu 2 A 29,   Mobile: +358-41-522 6544
02600 Espoo, Finland

Personal Information

Born January 30th, 1979 in Espoo, Finland. Married to Nuutti Hyvönen, since January 2007.

Language skills

Finnish (native language), English (excellent), Swedish (basic), French, German, Russian & Japanese (elementary).

Education

PhD (Doctor of Science in Technology)   2007 (Expected)
Department of Computer Science, Helsinki University of Technology.
Major: information science, minor: genetics.

MSc (Master of Science in Technology)   2002
Department of Engineering Physics and Mathematics, Helsinki University of Technology.
Major: information science, minor: information processing science.

Matriculation examination   1997
Tapiolan lukio (Tapiola senior high school), Espoo.

Professional Experience

Researcher   1.7.2007-Present
Lab. of Computer and Information Science   Helsinki University of Technology
Member of professor Samuel Kaski's research group *Statistical Machine Learning and Bioinformatics* (`www.cis.hut.fi/projects/mi`) since summer of 2000.
Currently, finished writing the PhD thesis and starting to work on generative models of network structure.

Bioinformatics Specialist   1.4.2006-Present
Genolyze Ltd
Analyzing microarray data (part time), see `www.genolyze.com`.

Researcher/Research Assistant   1.10.2004-30.6.2007
Department of Computer Science   University of Helsinki
Analyzing human endogenous retrovirus (HERV) data. Development of a special hidden Markov mixture model for estimating the activities of HERVs from databases of expressed sequences.

Researcher/Research Assistant   24.5.2000-30.9.2004
Lab. of Computer and Information Science   Helsinki University of Technology
Applying self-organizing maps to the analysis of various data sources: scientific abstracts, gene

expression data, and human endogenous retrovirus sequences. Development of methods for assessing the reliability of SOM visualizations.

Exercise assistant   Fall 1998 - Spring 2000
Institute of Mathematics   Helsinki University of Technology
Giving exercises in the first year mathematics courses and grading exams (part time).

Trainee   Summers of 1998 and 1999
Nokia Telecommunications
Testing software and implementing software for testing mobile telephone exchanges.

Teaching experience

Course assistant, High-throughput bioinformatics, TKK, 2007.

Course assistant, Practical course in microarray data analysis, UH, 2007.

Course assistant, Computational modeling methods in bioinformatics (seminar), TKK, 2006.

Course assistant, Machine Learning in Bioinformatics, UH, 2005.

Course assistant, From gene expression to regulation of cell function, TKK, 2003.

Exercise assistant, Basic course in mathematics C1/C2/S1/S2, TKK, 1998-2000.

UH=University of Helsinki, TKK = Helsinki University of Technology

Research Interests

Bioinformatics, machine learning, exploratory data analysis, probabilistic modeling, analysis of microarray data, study of transcriptional regulation of genes, analysis of genomic data, unsupervised learning, clustering and visualization.

Publications

1.Merja Oja. In Silico Expression Profiles of Human Endogenous Retroviruses. In Jagath C. Rajapakse, Bertil Schmidt, and Gwenn Volkert, editors, *Proceedings of Workshop on Pattern Recognition in Bioinformatics PRIB 2007, Lecture Notes in Bioinformatics*, pages 253-263, 2007. In press.

2.Merja Oja, Jaakko Peltonen, Jonas Blomberg, and Samuel Kaski. Methods for estimating human endogenous retrovirus activities from EST databases. *BMC Bioinformatics.* 8(Suppl 2):S11, 2007.

3.Merja Oja, Jaakko Peltonen, and Samuel Kaski. Estimation of human endogenous retrovirus activities from expressed sequence databases. In Juho Rousu, Samuel Kaski, and Esko Ukkonen, editors, *Probabilistic Modeling and Machine Learning in Structural and Systems Biology (PMSB 2006), Tuusula, Finland, 17-18 June*, workshop proceedings, pages 50-54, Helsinki University Printing House, 2006.

4.Merja Oja, Göran O. Sperber, Jonas Blomberg, and Samuel Kaski. Self-organizing map-based discovery and visualization of human endogenous retroviral sequence groups. *International Journal of Neural Systems*, 15(3):163–179. 2005.

5.Merja Oja, Göran O. Sperber, Jonas Blomberg, and Samuel Kaski. Grouping and Visualizing Human Endogenous Retroviruses by Bootstrapping Median Self-organizing Maps. In *Proceedings of IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology. San Diego, California, USA, 7-8 October, 2004.* pages 95-101, 2004.

6.Samuel Kaski, Janne Nikkilä, Merja Oja, Jarkko Venna, Petri Törönen, and Eero Castrén. Trustworthiness and metrics in visualizing similarity of gene expression. *BMC Bioinformatics*, 4:48, 2003.

7.Merja Oja, Panu Somervuo, Samuel Kaski and Teuvo Kohonen Clustering of Human Endogenous Retrovirus Sequences with Median Self-Organizing Map. In *WSOM'03 – Workshop on Self-Organising Maps, Hikibino, Kitakyushu, Japan, 11-14 September, 2003.* Proceedings on CD-ROM. 2003.

8.Merja Oja, Samuel Kaski, and Teuvo Kohonen, Bibliography of Self-Organizing Map (SOM) Papers: 1998-2001 Addendum, *Neural Computing Surveys*, Volume 3, pages 1-156, 2003.

9.Merja Oja, Janne Nikkilä, Petri Törönen, Eero Castrén, and Samuel Kaski. Learning metrics for visualizing gene functional similarities. In Pekka Ala-Siuru and Samuel Kaski, editors, *STeP 2002 - Intelligence, The Art of Natural and Artificial. The 10th Finnish Artificial Intelligence Conference, Oulu, Finland, 15-17 December, 2002*, pages 31-40, 2002.

10.Merja Oja, Janne Nikkilä, Petri Törönen, Garry Wong, Eero Castrén, and Samuel Kaski. Exploratory clustering of gene expression profiles of mutated yeast strains. In Wei Zhang and Ilya Shmulevich, editors, *Computational And Statistical Approaches To Genomics.* Kluwer Acadmic Publishers, 2002.

Theses

1.Merja Oja. PhD Thesis: Methods for exploring genomic data sets: Application to human endogenous retroviruses. Helsinki University of Technology, submitted.

2.Merja Oja. Diploma Thesis: Gene expression data mining using self-organizing maps and learning metrics (in Finnish), Helsinki University of Technology, April 2002.

Referee work

I have reviewed journal articles for *BMC Bioinformatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics, IEEE Transactions on Neural Networks, Computer-Aided Civil and Infrastructure Engineering* and *IEEE Engineering in Medicine and Biology Magazine.*

In addition, I have reviewed articles for the following international conferences: Asia Pacific Bioinformatics Conference (APBC'08). Workshop on Self-Organizing Maps (WSOM'07), (WSOM'05). Workshop on Pattern Recognition in Bioinformatics (PRIB'07). International Conference on Natural Computation (ICNC'07), (ICNC'05). IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, (CIBCB'07; program committee), (CIBCB'06), (CIBCB'05). Scandinavian Conference on Artificial Intelligence (SCAI'06). Workshop on Computational Systems Biology (WCSB'06). European Symposium on Artificial Neural Networks (ESANN'06). International Symposium on Neural Networks (ISNN'06). European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'05). IEEE Machine Learning for Signal Processing Workshop (MLSP'05). International Conference on Intelligent Data Engineering and Automated Learning (IDEAL'04).

Oral presentations at international conferences

1.In Silico Expression Profiles of Human Endogenous Retroviruses. In *Workshop on Pattern Recognition in Bioinformatics 2007 (PRIB). Singapore, October 1-2, 2007.* Publication: [1].
2.Estimation of human endogenous retrovirus activities from expressed sequence databases. In *Probabilistic Modeling and Machine Learning in Structural and Systems Biology (PMSB 2006), Tuusula, Finland, 17-18 June.* Publication: [3]
3.Grouping and Visualizing Human Endogenous Retroviruses by Bootstrapping Median Self-organizing Maps. In *IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). San Diego, California, USA, 7-8 October, 2004.* Publication: [5].
4.Clustering of Human Endogenous Retrovirus Sequences with Median Self-Organizing Map. In *WSOM'03 – Workshop on Self-Organising Maps, Hikibino, Kitakyushu, Japan, 11-14 September, 2003.* Publication: [7].
5.Learning metrics for visualizing gene functional similarities. In *STeP 2002 – The 10th Finnish Artificial Intelligence Conference, Oulu, Finland, 15-17 December, 2002.* Publication: [9].

Posters presentations at international conferences

1.Merja Oja, Jaakko Peltonen, Jonas Blomberg and Samuel Kaski. Estimating human endogenous retrovirus activities in various tissues with a hidden Markov mixture model. Poster in *Intelligent Systems for Molecular Biology & European Conference on Computational Biology 2007 (ISMB/ECCB)*. Vienna, Austria, July 21-25, 2007.

2.Merja Oja, Jaakko Peltonen, and Samuel Kaski. A hidden Markov model mixture for estimating human endogenous retrovirus activities from expressed sequence databases. Poster in *European Conference on Computational Biology 2006 (ECCB)*, Eilat, Israel, January 21-24, 2007.

3.Merja Oja, Jonas Blomberg, and Samuel Kaski. Class discovery and visualization for human endogenous retroviruses by bootstrapping Median Self-organizing Maps. Poster in *Bioinformatics 2004*, Linköping, Sweden, June 3-6, 2004.

4.Merja Oja, Petri Törönen, Janne Nikkilä, Eero Castrén, and Samuel Kaski. Learning metrics for SOM-based clustering and visualization of yeast gene expression data. Poster in *Bioinformatics 2002*, Bergen, Norway, April 4-7, 2002.

Research visits abroad

Three research visits to the laboratory of our collaborator, professor Jonas Blomberg, the Section of Virology, Department of Medical Sciences, Uppsala University. October 2005 (one week), December 2004 (one week), March 2004 (one month).

Computer and programming skills

Operating systems: Linux/unix, Mac OS, Windows

Programming languages: R, perl, Matlab, html, C/C++, Java.

Awards and grants

Grant of 16000 euros, Finnish cultural foundation, 2006.

Travel award of $1100, the International Society for Computational Biology (ISCB), 2006.

Grant of 4000 euros, Tekniikan edistämissäätiö, 2004.

Travel award of 1500 SEK, the Nordic Bioinformatics Network, 2004.

Travel award of 1500 SEK, the Nordic Bioinformatics Network, 2002.

Grant of 2000 FIM, The Student Union of Helsinki University of Technology, 1999.

Positions of Trust

Board Member, Finnish society for Bioinformatics, 2006.

Treasurer, Jämerän naiset ry, 2003.

Secretary, Iteraatio (The Club of Students of Computer and Information Science), 2002-2003.

Board member, Fyysikkokilta (Guild of Physics Students), 1999.

Secretary, Fyysikkokilta, 1998.

Hobbies

Reading, traveling, aerobics.