

Keith Noto

knoto@cs.ucsd.edu · (608) 695-3215

Education

Postdoc, University of California—San Diego, 2007-2009.

PhD, University of Wisconsin—Madison, Computer Sciences, 2007.

Thesis Title: *Learning Expressive Computational Models of Gene Regulatory Sequences and Responses*

MS, University of Wisconsin—Madison, Computer Sciences, 2002.

BA, University of Wisconsin—Madison, Computer Sciences and Philosophy, 1998.

Selected Peer-Reviewed Publications

K. Noto, M. H. Saier, Jr. and C. Elkan.

Learning to Find Relevant Biological Articles Without Negative Training Examples.

Twenty-First Australasian Joint Conference on Artificial Intelligence (AI 2008) (acceptance rate 29%),
In *Lecture Notes in Bioinformatics* 5360:202-213. Springer-Verlag.

M. H. Saier, Jr., M. R. Yen, K. Noto, D. G. Tamang, and C. Elkan.

The Transporter Classification Database: Recent Advances.

Nucleic Acids Research 37:(Database issue), D274-D278, 2009.

C. Elkan and K. Noto.

Learning Classifiers from Only Positive and Unlabeled Data.

Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008), 213-220 (acceptance rate 18.6%).

K. Noto and M. Craven.

Learning Hidden Markov Models for Regression using Path Aggregation.

Proceedings of the 24th Uncertainty in Artificial Intelligence Conference (UAI 2008), 444-451 (acceptance rate 13%).

K. Noto and M. Craven.

Learning Probabilistic Models of *cis*-Regulatory Modules that Represent Logical and Spatial Aspects.

Proceedings of the 2006 European Conference on Computational Biology (acceptance rate 18%),
In *Bioinformatics* 23(2):e156-162.

K. Noto and M. Craven.

A Specialized Learner for Inferring Structured *cis*-Regulatory Modules.

BMC Bioinformatics 7:528, 2006.

K. Noto and M. Craven.

Learning Regulatory Network Models that Represent Regulator States and Roles.

RECOMB 2004 Workshop on Regulatory Genomics.

In *Lecture Notes in Bioinformatics* 3318, pages 52-64. Springer-Verlag, 2004.

Presentations and Posters

Paper presentation, 2008 Conference on Uncertainty in Artificial Intelligence (UAI) in Helsinki, Finland.

Invited speaker, National Library of Medicine (NLM) Informatics Training Conference 2006, hosted by Vanderbilt University.

Paper presentation, 2006 European Conference on Computational Biology (ECCB) in Eilat, Israel.

K. Noto and M. Craven, Poster: A Specialized Learner for Inferring Structured *cis*-Regulatory Modules. 2005 International Conference on Intelligent Systems for Molecular Biology (ISMB) in Detroit, MI.

Paper presentation, RECOMB 2004 Workshop on Regulatory Genomics, hosted by the University of California—San Diego.

Research Experience

(The following briefly highlights my research interests and experience. Please see my research statement for details.)

Research Interests

- Machine learning algorithm development and application.
- Probabilistic methods/probabilistic graphical models (Bayesian networks, hidden Markov models).
- Applying machine learning to biological domains (bioinformatics).
- Development of software beyond research level to create real-world tools.

Research Projects

- Inferring gene regulatory networks from differential gene expression (microarray) and cellular conditions (RECOMB workshop 2004).
- Discovering modules of transcription factor binding sites, including spatial and logical relationships between sites, primarily from sequence and expression data (ECCB 2006, BMC Bioinformatics 2006).
- Adapting a sequence classification algorithm for regression tasks using hidden Markov models (UAI 2008) and support vector machines (2009; in preparation).
- Biomedical text classification for a specialized database (KDD 2008, AJCAI 2008, NAR 2009).

Teaching Experience

Taught *Introduction to Artificial Intelligence* at the University of Wisconsin—Madison, 2002.

- Taught about 30 undergraduate students taking this "core" undergraduate course.
- Prepared and gave lectures.
- Prepared written and programming homework assignments.
- Wrote and administered exams.

Teaching assistant at the University of Wisconsin—Madison.

- *Machine Learning* (Fall 2001, Prof. Jude Shavlik).
- *Introduction to Algorithms* (Spring 2002, Prof. Deborah Joseph).
- *Advanced Artificial Intelligence* (Fall 2002, Prof. David Page).
- *Introduction to Artificial Intelligence* (Fall 2002, Prof. Jude Shavlik).

Co-Mentor of an undergraduate student for the Computational Biology and Biostatistics (CBB) summer research program at the University of Wisconsin—Madison, June-August, 2006.

- Designed a two-month computational biology project for the student.
- Advised on all aspects of the project: Code design, computational experiments, analysis of results and a final report.
- Attended a seminar course for mentors.

Guest lecturer for

- *Introduction to Algorithms* (University of Wisconsin—Madison).
- *Advanced Artificial Intelligence* (University of Wisconsin—Madison).
- *Advanced Bioinformatics* (University of Wisconsin—Madison).
- *Introduction to Artificial Intelligence* (University of California—San Diego).

Professional Service

I have reviewed papers for *Machine Learning*, the *Journal of Machine Learning Research*, the *IBM Journal of Research and Development*, and *Data Mining and Knowledge Discovery*

I was on the 2008 program committee for the 2008 AAAI Conference on Artificial Intelligence.

I served on an NSF grant review panel in the area of *robust intelligence* in 2009.

Awards

Computation and Informatics in Biology and Medicine (CIBM) fellowship, April 2003-2006 (<http://www.cibm.wisc.edu>).

CIBM fellowship (second award), April 2006-2007

Travel fellowship award for ECCB 2006 from Biosapiens, funded by the EU.

Travel fellowship award for the RECOMB 2004 workshop.

Distinctive scholastic achievement (top 20% of graduating class), B.A. from the University of Wisconsin—Madison, 1998.

Software

Rupa (Regression Using Path Aggregation), a package for learning a regression model (based on hidden Markov models) from the real-valued response associated with each of a set of training sequences (written in C and Python) (<http://www.cs.ucsd.edu/~knoto/rupa>).

Scrm (Structured *cis*-Regulatory Modules, versions 1 and 2), packages for learning the arrangements of multiple DNA-binding sites from sequence data (written in C++) (<http://pages.cs.wisc.edu/~noto/crm>).

Previous Employment

September 1998 - January 2002: Software engineer,
Legislative Technology Services Bureau (Wisconsin state government).

- Developed software used by the Wisconsin state representatives, senators and staff, including several features of the Assembly and Senate voting systems.
- Developed software that communicates with Wisconsin's "InSession" sites.
 - <http://www.legis.state.wi.us/insession/assembly>
 - <http://www.legis.state.wi.us/insession/senate>

References

Prof. Mark Craven (Ph.D. advisor)

University of Wisconsin—Madison, Dept. of Biostatistics and Medical Informatics
1300 University Avenue #6730 MSC, Madison WI 53706

E-mail: craven@biostat.wisc.edu

Phone: (608) 265-6181

Prof. Charles Elkan (Postdoc Mentor)

University of California at San Diego, Dept. of Computer Science and Engineering
9500 Gilman Drive (EBU3B, MC0404), La Jolla, CA 92093

E-mail: elkan@cs.ucsd.edu

Phone: (858) 534-8897

Prof. Milton H. Saier, Jr.

University of California at San Diego, Dept. of Biology

9500 Gilman Drive, La Jolla, CA 92093

E-mail: saier@biomail.ucsd.edu

Phone: (858) 534-4084

Prof. Jude Shavlik

University of Wisconsin—Madison, Dept. of Computer Sciences

1210 West Dayton Street, Madison WI 53706

E-mail: shavlik@cs.wisc.edu

Phone: (608) 263-7625

Prof. David Page

University of Wisconsin—Madison, Dept. of Biostatistics and Medical Informatics

1300 University Avenue #6743 MSC, Madison WI 53706

E-mail: page@biostat.wisc.edu

Phone: (608) 265-6186

Jeffrey Ylvisaker

Director, Legislative Technology Services Bureau (Wisconsin)

17 W. Main St. Suite 200, Madison WI 53703

E-mail: jeff.ylvisaker@legis.state.wi.us

Phone: (608) 267-9528