Paul Komarek

Last Updated: 2013-08-09
komarek.paul@gmail.com

San Jose, CA
412.983.1284
U.S. Citizen

Objectives

I am a project-oriented scientist and software engineer that enjoys solving large and complex technical challenges in a team environment.

Education

Ph.D (Algorithms, Combinatorics, and Optimization) **Carnegie Mellon University**, advised by Andrew Moore, May 2004.

M.S (Algorithms, Combinatorics, and Optimization) **Carnegie Mellon University**, advised by Andrew Moore, May 1998.

B.S (Mathematics) **Western Washington University**, Magna Cum Laude, Graduation with Distinction in Mathematics, June 1997

Awards

- **Google OC Award**, for improvements to cluster utilitzation, 2009
- **Google OC Award**, for Google Custom Search, 2009
- **NASA Space Grant Scholarship** Pennsylvania Space Grant Consortium, December 2001
- **Outstanding Participant** CMU Center for Nonlinear Analysis Summer Undergraduate Applied Mathematics Institute, July 1996
- **Outstanding Mathematics Graduate** Western Washington University, June 1997

Employment

Software Engineer	Google, **Site Reliability Engineering**, December 2007 through present, Technical Lead (2009-) and Manager(2011-) for two production automation environments.
Software Engineer	Google, **Search Quality**, November 2005 through December 2007, Ranking (2005-2006) and Co-op/Custom Search (2006-2007)

Education Related Activities and Employment

Postdoctoral Fellow	Carnegie Mellon University, Robotics Institute, **Auton Lab**, May 2004 through September 2005
Pfizer Collaboration Lead	Carnegie Mellon University, Robotics Institute, **Auton Lab**, January 2003 through present, coordinated development of Auton software products for the Auton-Pfizer collaboration
Consultant	**Aethon Inc.**, January 2002 to April 2002, researched wireless communications, developed an elevator interface protocol and accompanying software, provided some systems support, and delivered finished functional prototypes for these systems (Aethon develops a mobile robot for hospital use)
Systems Administrator	Carnegie Mellon University, Robotics Institute, **Auton Lab**, 2000 through present, including planning, acquisition, deployment, maintenance and security of computing resources and servers; also helped hire, train, and supervise three systems administrators.

Research Assistant	Carnegie Mellon University, **Department of Mathematical Sciences**, August 1997 through May 2004, except when teaching (see below)
Undergraduate Researcher	Western Washington University, **Department of Mathematics**, September 1996 to June 1997, developed Maple-based two- and three-dimensional tomographic reconstruction software for convex polytopes
Mathematics Fellow	Western Washington University, **Department of Mathematics' Math Center**, September 1994 to June 1997, upper-division mathematics tutor

Deployed Hardware and Software Systems

Internal Google Software	*Jan 2006*- I have design, built, and deployed several internal and customer-facing software systems at Google, related to ranking, Google Custom Search, and automation tools.
LR-TRIRLS	*May 2005*- First source release for my Logistic Regression with Truncated Regularized Iteratively Re-weighted Least Squares software. Licensed under the GNU General Public License (GPL), available at `http://www.autonlab.org` and `http://komarix.org/ac/lr`.
AFC Active Learning	*September 2004*- Active learning software for scheduling roboticized pharmaceutical experiments (see AFC, below). I am responsible for the design, implementation, and maintenance of this software. This software has been delivered to the sponsor, and will be maintained and distributed contingent on future contracts.
Auton Fast Classifiers (AFC)	*April 2002*- Fast classification software for high-dimensional datasets. I provided new algorithms and eventually took over the entire software system, including the user interface, learner and and dataset framework, performance evaluation, and documentation. This software is still in use by the sponsor, is maintained regularly, and has been widely distributed.
Aethon Elevator Controller	*January 2002*- Software for managing a single-board computer and serial interface board connected to an elevator's control system. I developed a protocol and daemons for bidirectional communication between a mobile robot and a passenger elevator. Aethon's current elevator controller is a derivative of my prototype.
Aethon Wireless Relays	*January 2002*- Stand-alone devices for ad-hoc relaying of communications between mobile robots and an elevator controller (above). I selected the embedded hardware, created a small GNU/Linux operating system and installation utilities adapted to compact flash, and wrote the message relay software. Aethon's current version of this device is a derivative of my prototype.
Auton Build System	-*July 2001*- I maintain the makefile and scripts used for building all Auton software on various compilers, microprocessors, and POSIX-ish environments.
Auton Compute Infrastructure	-*January 2000*- Servers, compute machines, storage and services used by Auton lab members for research. My responsibilities include • hardware and software selection, procurement, and deployment • maximizing performance for niche scientific needs on a limited capital budget in a university environment • maintaining vendor relationships and negotiating affordable prices • understanding the current high-performance and consumer computing markets, both for our needs and for occasional advisement of clients and other academics. • maintaining software, security, and services This collection of user and server systems is used daily and maintained constantly. Some responsibilities have been shared with additional admins since Spring 2002.

Tomographic Reconstruction Software

Fall 1996 Software package for reconstructing a density function over a convex polytope using only information from (n-1)-dimensional integrals ("x-rays"). I developed this software for a math professor to use as part of his Geometric Tomography classes.

Publications and Talks

- Paul Komarek and Andrew Moore, *Making Logistic Regression A Core Data Mining Tool with TR-IRLS*, **International Conference on Data Mining**, 2005 (ICDM 2005)
- Paul Komarek, *Logistic regression for fast, accurate, and parameter free data mining*, Invited talk at **Google Inc.**, July 2005.
- Paul Komarek and Andrew Moore, *Making Logistic Regression A Core Data Mining Tool: A Practical Investigation of Accuracy, Speed, and Simplicity*, Technical Report TR-05-27 at the **Robotics Institute, Carnegie Mellon University**, May 2005.
- Paul Komarek, *Autonomous Fast Classifiers for Pharmaceutical Data Sets*, Invited talk at **Applied Biosystems Inc.**, July 2004
- Paul Komarek, *Autonomous Fast Classifiers for Pharmaceutical Data Sets*, Invited talk at the **Midwest Biopharmaceutical Statistics Workshop** 2004 (MBSW 2004)
- Paul Komarek, *Logistic Regression for Data Mining and High-Dimensional Classification*, **Doctoral Thesis**, 2004
- Alex Gray, Paul Komarek, Ting Liu, and Andrew Moore, *High-Dimensional Probabilistic Classification for Drug Discovery*, **Computational Statistics**, 2004 (CompStat 2004)
- Anya Goldenberg, Paul Komarek, Jeremy Kubica, Andrew Moore, and Jeff Schneider *A Comparison of Statistical and Machine Learning Algorithms on the Task of Link Completion*, **Knowledge Discovery in Databases**, 2003 (KDD 2003)
- Paul Komarek and Andrew Moore, *Fast Logistic Regression for Large Sparse Datasets with Binary Outputs*, **Artificial Intelligence and Statistics**, 2003 (AISTAT 2003)
- Paul Komarek and Andrew Moore, A Dynamic Adaptation of AD-trees for Efficient Machine Learning on Large Data Sets, **International Conference on Machine Learning**, 2000 (ICML 2000)
- Paul Komarek, Canonical Ramsey Numbers–A New Lower Bound for Off-Diagonal Ramsey Numbers, **Joint Mathematics Meetings (AMS/MAA)**, 1997

Other Professional Activities

- Mentoring Auton graduate students, Jan 2002 to Dec 2002 and Jan 2005 to present.
- Advising an undergraduate intern in the Auton lab, June 2005 to present.
- Supervising part-time system administration employees, March 2002 to June 2005.
- Research advisement of graduate and undergraduate students in the Auton Lab, at the Robotics Institute, Carnegie Mellon University, May 2004 to present.
- Refereeing submissions to the *Information Systems* journal and several conferences, including the IEEE *Transactions on Knowledge and Data Engineering* (TKDE), *Uncertainty in Artificial Intelligence* (UAI), *Knowledge Discovery and Data Mining* (KDD), and *Neural Information Processing Systems* (NIPS).
- Teaching Assistant for the Department of Mathematical Sciences, Carnegie Mellon University, August 1998 to December 1998.
- Participant in the Center for Nonlinear Analysis' Summer Undergraduate Applied Mathematics Institute at Carnegie Mellon University, Summer 1996.

Personal

When outdoors I enjoy soccer, hiking, travel and photography. Indoors, I dabble in electronics and use my embedded computing experience for entertainment. I like to combine software, hardware, woodworking and metalworking to complete "essential" upgrades to our home, including a small home theater.

References

Available on request.