Carnegie Mellon University
Pittsburgh, PA 15213
brigham@cmu.edu

6543 Dalzell Place
Pittsburgh, PA 15217

Brigham Anderson

Goals

My recent research has had the major themes of *fast* inference and *active* inference in graphical models. Fast inference is clearly useful for real-world applications such as recommenders and personalization. The promise of active inference can be expressed intuitively as introducing *curiosity* into machine learning. This allows for closed-loop learning, in which machines are more autonomous and learn faster because they actively seek the most useful data.

Education

Ph.D. (Robotics), **Carnegie Mellon University**, August 2003. Thesis title: "Nonparametric Optimization and Galactic Morphology." Thesis Advisor: Andrew W. Moore.

M.S. (Organizational Theory), **Carnegie Mellon University**, June 1999. Thesis title: "Interaction of Size and Density with Graph Level Indices." Thesis advisor: Kathleen M. Carley

B.S. (Biochemistry), **University of Washington**, June 1994, Honors
B.S. (Biology), **University of Washington**, June 1994, Cum Laude

Research Employment

05/2005-present

Postdoctoral fellowship, Biosurveillance project, **Carnegie Mellon.** This project is creating an early-warning system for monitoring multiple data streams (emergency room visits, over-the-counter sales of various items, webMD usage, etc.) and noticing patterns indicative of different outbreak types (anthrax, botulism, influenza, cryptosporidium, etc.) Challenging aspects of the problem include properly fusing these disparate data streams while also controlling for environmental variables such as the "weekend-effect" for ER visits. We have employed two approaches to the problem. This first uses dynamic Bayes nets to model the entire system explicitly, and the second uses a large mixture model to estimate the baseline signal

each day, in conjunction with a hidden Markov model to perform inference.

2003-2005

Postdoctoral fellowship, CALO project, **Carnegie Mellon.** This DARPA-funded project aims to create an automated desktop assistant that learns to predict the needs of its user. The AUTON lab's role was to model the user's mental/emotional state from keyboard, mouse, motion, and sound sensors. I directed three full-time programmers in development of a real-time system of sensors and communications protocols. My research developed novel active learning algorithms for optimally determining what the system should be curious about at any given time (to enable information-seeking behavior.) This research culminated in the first linear-time all-pairs value of information algorithm, several orders of magnitude faster than the state of the art.

2001-03 Graduate Research Assistantship, Astronomy project, **Carnegie Mellon.** Given a dataset of millions of images of galaxies, determine the morphology (shape parameters) of each. The problem is especially difficult because the images are noisy, distorted by atmospheric and lens imperfections, and the dataset is massive, containing tens of millions of images. The state of the art algorithm took about 3 minutes per galaxy image. Our algorithm, based on nearest neighbor regression in eigenspace, did the same analysis in about 1 second.

2002 Graduate Research Assistantship, **Caterpillar Inc.** Optimizing engine performance and emissions under different speed and torque conditions by automating engine timing adjustment.

2001 Graduate Research Assistantship, **3M Corp.** Automated experimentation in research and development of adhesives.

1999 Graduate Research Assistantship, **Carnegie Mellon, Robotics Institute.** Automated experimental design for mobile robots. Applied active learning techniques to control of a small mobile robot.

1998 Personal Data Assistant project. **Carnegie Mellon, Robotics Institute.** Learning rules to predict patterns of human activity from data gathered from a Personal Digital Assistant (PDA.)

1997 Graduate Research Assistantship, Organizational Simulation Project. **Carnegie Mellon, Social and Decision Sciences.** Directed by Prof. Kathleen Carley. Developed models for simulating organizations as networks of learning machines.

1993-94 Undergraduate research, Werner Syndrome Gene Project, **University of Washington.** Directed by Dr. Jerry Schellenberg. Investigated causes of premature aging at the site of the Werner syndrome gene.

Teaching Employment

2000 TA, **Artificial Intelligence, CMU**, graduate-level and undergraduate-level. Taught labs, created assignments and tests, held office hours.

1996,98 TA, **Economic Policy, CMU**, created tutorial software using economic simulations, e.g., supply-demand curve emergence.

1996,97 TA, **Experimental Research Methods, CMU**, Taught labs

1995 TA, **Organizational Theory, CMU**, office hours, grading.

Industry Employment

1997 Consultant, **Elanex Pharmaceuticals**, Vancouver, Canada. Optimized the production of protein using Markov models of bioreactor production process.

1994-95 Lab Technician, Process Development, Research & Development, **Bristol-Myers Squibb**, Seattle, WA. Optimizing production of experimental drugs.

Recent Presentations

2006 "Probabilistic Methods and Data mining", To be presented as a workshop at the Center for Discrete Mathematics and Computer Science (DIMACS), Rutgers, New Jersey.

2005 "Efficient Value of Information for Graphical Models", Neural Information Processing Systems conference, Vancouver, Canada.

"Bayesian Networks for Customs Security", Customs and Border Patrol, Washington DC.

"Active Learning for Hidden Markov Models", International Conference on Machine Learning, Bonn, Germany.

"Optimal Hidden Markov Model Learning", Machine Learning Lunch, Carnegie Mellon University, PA.

2004 "Astrophysics and Massive Data-mining", Service Composition for Data Exploration in the Virtual Observatory (SC4DEVO) conference, Caltech, CA.

Professional Activities

2006 Program Committee
Knowledge Discovery and Data Mining conference.

2005 Reviewer
Journal of Machine Learning Research.

2005 Reviewer
Knowledge Discovery and Data Mining conference.

Publications

[1] **B. Anderson**, S. Siddiq, A. Moore, "Sequence Selection for Active Learning," Submitted to *International Conference on Machine Learning*, 2006.

[2] **B. Anderson**, A. Moore, "Efficient Value of Information for Graphical Models," *Neural Information Processing Systems*, 2005. (Main-Track Presentation.)

[3] **B. Anderson**, A. Moore, "Active Learning for Hidden Markov Models: Objective Functions and Algorithms," *International Conference on Machine Learning*, 2005.

[4] **B. Anderson**, A. Moore, A. Connolly, & B. Nichol, "Eigengalaxies for Fast Galaxy Morphology," *Knowledge Discovery and Data Mining*, 2004.

[5] **B. Anderson**, "Nonparametric Optimization and Galactic Morphology," Doctoral Dissertation, CMU-RI-TR-03-17, Carnegie Mellon University, Pittsburgh, PA, 2003.

[6] **B. Anderson**, A. Moore, & D. Cohn, "A Nonparametric Approach to Noisy and Costly Optimization," *International Conference on Machine Learning,* 2000.

[7] **B. Anderson**, C. Butts, & K. Carley, "The Interaction of Size and Density with Graph Level Indices," *Social Networks*, 21(3), 239-267, 1999.

[8] **B. Anderson**, & A. Moore, "ADtrees for Fast Counting and for Fast Learning of Association Rules," *Knowledge Discovery and Data Mining*, New York, 1998.

Languages

C, C++, Java, Matlab, Perl, VB

References

Andrew Moore, Director, Google Labs Pittsburgh
Carnegie Mellon University
Pittsburgh, PA 15213
(412) 334-2063
awm@cs.cmu.edu

Matt Mason, Director, Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
(412) 268-8804
matt.mason@cs.cmu.edu

Andrew Connolly, Associate Professor
401c Allen Hall
Department of Physics and Astronomy
University of Pittsburgh
Pittsburgh, PA 15260
(412) 624-1345
ajc@phyast.pitt.edu