# Conference on Automated Learning and Discovery

*June 11-13, 1998*
*Carnegie Mellon University, Pittsburgh, PA*

http://www.cs.cmu.edu/~conald

# The CONALD questions:

- STATE OF THE ART

  What is the state of the art? What are examples of successful systems?

- GOALS AND IMPACT

  What are the long-term goals of the field? What will be the most likely future impact of the area?

- PROMISING RESEARCH TOPICS

  What are examples of the most promising research topics that should be pursued in the next three to five years?

- OPPORTUNITIES FOR CROSS-DISCIPLINARY RESEARCH

  Which are the most significant opportunities for cross-cutting research (e.g., involving AI, databases, statistics, etc.)?

# Schedule of Events

| | Day One<br>Thursday, June 11 | Day Two<br>Friday, June 12 | Day Three<br>Saturday, June 13 |
|---|---|---|---|
| 8:30 AM | Coffee & Donuts – Connan Room<br>(8:55) Opening | Coffee & Donuts – Connan Room | Coffee & Donuts – Connan Room |
| 9:00<br>9:30 | 9:00–9:45, McConomy Auditorium<br>*Robert Tibshirani* | 9:00–9:45, McConomy Auditorium<br>*Herb Simon* | 9:00–9:45, McConomy Auditorium<br>*Stuart Geman* |
| 10:00 | 9:45–10:30, McConomy Auditorium<br>*Tom Dietterich* | 9:45–10:30, McConomy Auditorium<br>*Daryl Pregibon* | 9:45–10:30, McConomy Auditorium<br>Plenary Session I |
| 10:30 | Break – Outside of Connan | Break – Outside of Connan | Break – Inside Connan |
| 11:00<br>11:30 | 11:00–12:30<br><u>Workshop</u> — <u>Room</u><br>Learning Causal Bayesian Networks — Rangos 1<br>Mixed Media Databases — Rangos 2<br>Machine Learning & Reinf. Learning — Connan<br>Large-Scale Consumer Databases — Dowd | 11:00–12:30<br><u>Workshop</u> — <u>Room</u><br>Visual Methods for Massive Data — Connan<br>Learning from Text and the Web — Rangos 1<br>Robot Exploration and Learning — Rangos 2<br>Neural Basis of Learning — Dowd | 11:00–11:45, McConomy Auditorium<br>Plenary Session II |
| 12:00 PM | | | 11:45–12:30, McConomy Auditorium<br>*David Heckerman* |
| 12:30 | Lunch – Weigand Gym, University Center | Lunch – Weigand Gym, University Center | Closing |
| 1:00<br>1:30 | | | |
| 2:00<br>2:30<br>3:00 | 2:00–3:30<br>Workshops continued<br>(See above for locations) | 2:00–3:30<br>Workshops continued<br>(See above for locations) | Post-Conference Local Activities |
| 3:30 | Break – Outside of Rangos 1 & 2 | Break – Outside of Rangos 1 & 2 | |
| 4:00<br>4:30<br>5:00 | 4:00–5:25<br>Workshops continued<br>(See above for locations) | 4:00–5:25<br>Workshops continued<br>(See above for locations) | |
| 5:30<br>6:00<br>6:30 | 5:30–6:30, McConomy Auditorium<br>*Michael Jordan* | <u>River Cruise Banquet Transportation</u><br>6:40 leaves CMU University Center<br>6:40 leaves Holiday Inn, Lytton Ave.<br>7:00 boarding boat at Station Square Dock<br>10:45 Transportation back to Oakland | |

## Scientific Program Committee

*Sebastian Thrun (chair)*
*Christos Faloutsos*
*Stephen Fienberg*
*Tom Mitchell*
*Andrew Moore*
*Roni Rosenfeld*
*Richard Scheines*
*Teddy Seidenfeld*
*Kannan Srinivasan*
*Larry Wasserman*

---

## Workshop Organizers

*Richard Scheines*
*Larry Wasserman*
*Shumeet Baluja*
*Christos Faloutsos*
*Alex Hauptmann*
*Sridhar Mahadevan*
*Andrew W. Moore*
*Mike Meyer*
*Teddy Seidenfeld*
*Kannan Srinivasan*
*Bill Eddy*
*Steve Eick*
*Jaime Carbonell*
*Steve Fienberg*
*Tom Mitchell*
*Yiming Yang*
*Howie Choset*
*Maja Matarić*
*Sebastian Thrun*
*Michael Witbrock*
*David Touretzky*
*Tai Sing Lee*
*Bill Skaggs*

---

## Local Arrangements Committee

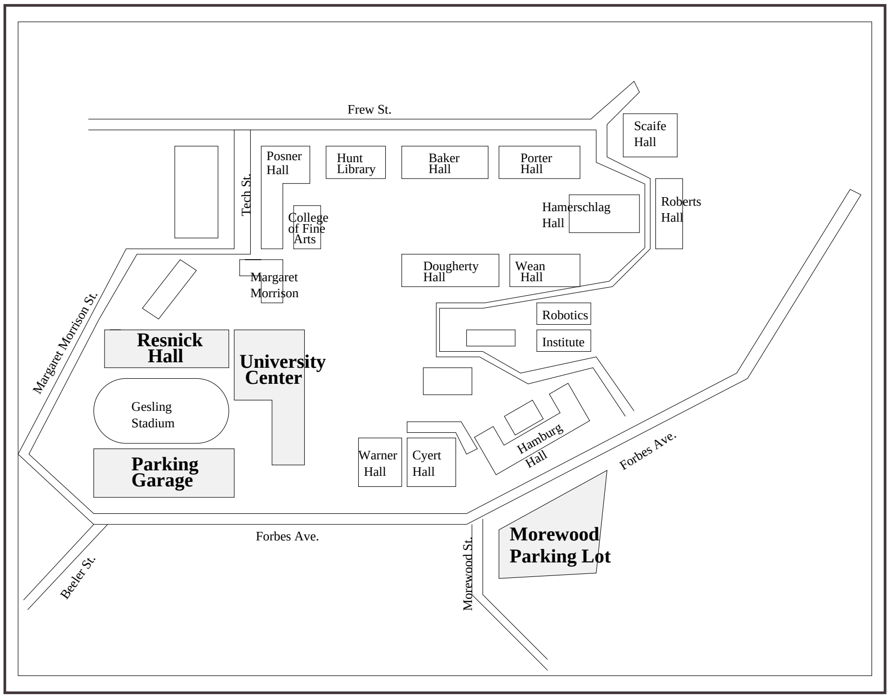
*Diane Stidle*
*Michael Bowling*
*Mark Craven*
*Frank Dellaert*
*Thomas Nichols*
*Nicholas Roy*
*Stella Maris Salvatierra*
*Ashish Sanil*
*Sean Slattery*
*Joseph O'Sullivan*
*Belinda Thom*

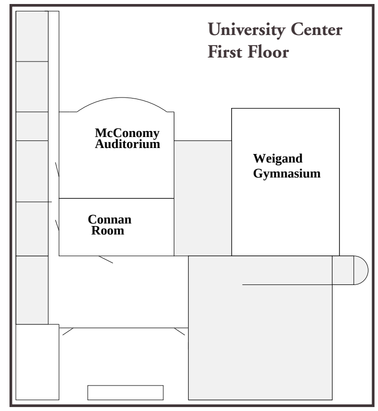
---

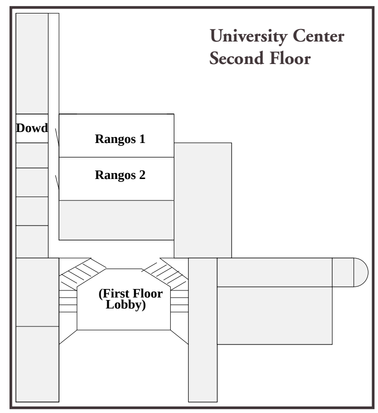
## Sponsored by

*Center for Automated Learning and Discovery*
*National Science Foundation*
*Caterpillar Inc.*
*Exxon Paramins*
*Giant Eagle, Inc.*
*KeyBank*
*Mellon Bank*
*Schlumberger Cambridge Research*
*Siemens Corporate Research, Inc.*
*VSI Satech*
*Westinghouse Science and Technology Center*

# Campus Map and University Center




# Navigating around Pittsburgh

To get to ⑤, University Center, site of CONALD, from:

① The University Club

② Holiday Inn Select University Center

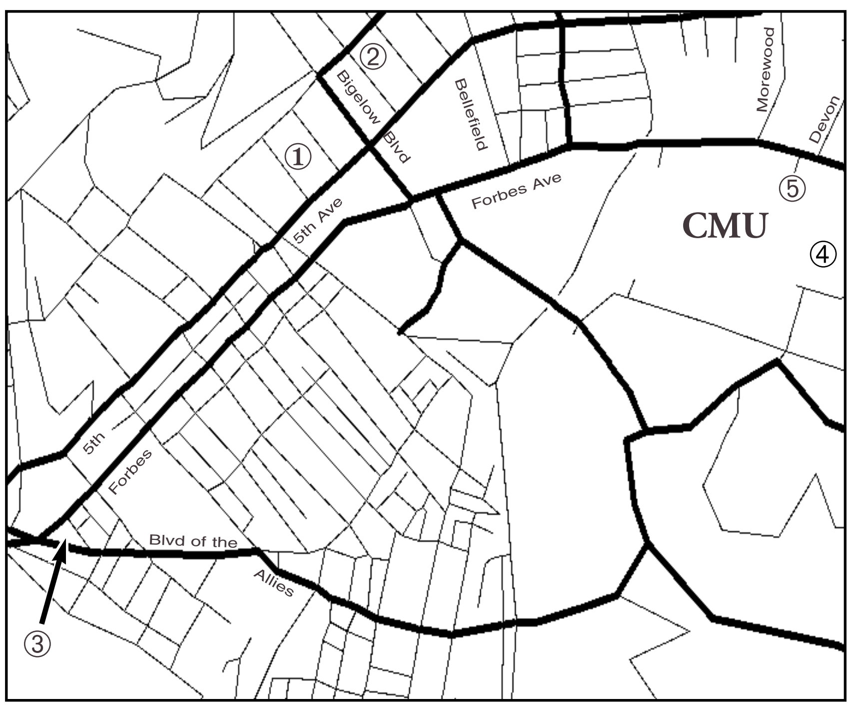
Proceed to Fifth Avenue, turn east (left) and walk to any of the intersections. Turn north (right) onto Bellefield or N. Craig Street and walk to Forbes Avenue. Turn east (left) onto Forbes Avenue and cross over Forbes. Continue up Forbes past the Exxon on the left, National City Bank on the right, across a small bridge, continue past Hamburg Hall, Campus Police and Cyert Hall. Once past Cyert Hall you will walk up the steps by the mail box veering left. You will see University Center, yellow brick with green windows. Walk along the side of the building and enter in the middle of the building.

③ Hampton Inn

Take a cab or hotel shuttle to Carnegie Mellon University Center

④ Resnick Hall Campus Dorms

Leave dorm, heading west. The University Center is the next building.



# Table of Contents

"Information Retrieval by Image Content and its Associated Text in a Digital Library"
*Victoria Meza, Jesus Favela*

"Generating a Large Database of Named Faces in CNN News Video" *Ricky A. Houghton*

"Automated Visual Search Tools Based on Integration of Multi-Level, Multi-Media Features"
*Shih-Fu Chang*

"A Hidden Markov Model Framework for Video Segmentation Using Audio and Image Features"
*Boreczky & Wilcox*

"When Listening is Not Enough: Potential Uses for Vision for a Reading Tutor that Listens"
*John Kominek, Gregory Aist, Jack Mostow*

"Classification Based Medical Case Retrieval: Using Radiology Images and Patient Text Information" *Yanxi Liu, Frank Dellaert, William E. Rothfus,MD*

"Discrete MMI Signatures for Media Classification and Clustering" *Jonathan T. Foote*

## W3: Machine Learning and Reinforcement Learning for Manufacturing

(Chairs: Sridhar Mahadevan and Andrew Moore)

"General Paradigm for Applying ML in Automated Manufacturing processes"
*Wei Zhang and Rod Tjoelker*

"Designing Experiments for Learning with Neural Networks"
*Özden Gür Ali, Yu-To Chen, Angela Neff*

"Reinforcement Learning for Realistic Manufacturing Processes"
*Stephan ten Hagen, Ben Kröse*

"Average Reward Reinforcement Learning for Optimizing Factories" *Prasad Tadepalli*

"High Quality Thermostat Control by Reinforcement Learning – A Case Study"
*Martin Riedmiller*

"Reinforcement Learning for Intelligent Building Control"
*Jamieson Schulte, Sebastian Thrun*

"Neuro-fuzzy Computing for Feature Evaluation, Classification and Rule Generation"
*Rajat K. De*

"Qualitatively Constrained Reinforcement Learning"
*Samuel Matsushima, Achim Hoffmann*

"Relevant Feature Selection for Large Learning Problems" *William D. Smart*

"Learning Segmentation of Time Series" *Patrick Feucht, Robert Suna*

## W4: Large-Scale Consumer Databases

(Chairs: Mike Meyer, Teddy Seidenfeld and Kannan Srinivasan)

"A Formal Satistical Approach to Collaborative Filtering" *Lyle Ungar, Dean Foster*

"Automated Discovery of Discriminant Rules for a Group of Objects in Databases"
*Tae-wan Ryu, Christoph F. Eick*

"Mining Data to Target Market Consumers From a Large Relational Database"
*Peter Johnson, Jim Delaney*

## W5: Visual Methods for Massive Data

(Chairs: Bill Eddy and Steve Eick)

"A Beautiful Diamond has Many Facets" *Graham Wills*

"The Interplay of Large-Scale Data Analysis and Data Visualization" *Andreas Buja*

"When Speech Input is Not an Afterthought: A Reading Tutor that Listens" *Jack Mostow, Gregory Aist*

"Information visualization: the broad product design problem" *Steve Roth*

## W6: Learning from Text and the Web

(Chairs: Jaime Carbonell, Steve Fienberg, Tom Mitchell and Yiming Yang)

"Wrapping Semistructured Web Pages with Finite-State Transducers" *Chun-Nan Hsu, Ming-Tzung Dung*

"Building Intelligent Agents for Web-Based Tasks: A Theory-Refinement Approach" *Jude Shavlik, Tina Eliassi-Rad*

"Learning to Exploit Document Relationships and Structure: The Case for Relational Learning on the Web" *Sean Slattery, Mark Craven*

"Wrapper Induction for Semistructured, Web-based Information Sources" *Ion Muslea, Steve Minton, Craig Knoblock*

"CiteSeer: An Autonomous System for Processing and Organizing Scientific Literature on the Web" *Kurt Bollacker, Steve Lawrence, C. Lee Giles*

"Simultaneous Word and Document Clustering" *John Lafferty, Peter Venable*

"Learning and Representing Topic, A Hierarchical Mixture Model for Word Occurrences in Document Databases" *Thomas Hofmann*

"Event Detection" *Yiming Yang, Tom Pierce, Jaime Carbonell*

"Text Mining with Decision Rules and Decision Trees" *Chid Apte, Fred Damerau, Sholom Weiss*

"Feature Selection for Classification Based on Text Hierarchy" *Dunja Mladenic, Marko Grobelnik*

"Estimating the Effectiveness of Conversational Behaviors in a Reading Tutor that Listens" *Gregory Aist, Jack Mostow*

"A Collaborative Filtering Agent System for Dynamic Virtual Communities on the Web" *O. de Vel, S. Nesbitt*

"Automatic Learning of a Syntactical Government Patterns Dictionary from Web-Retrieved Texts" *A.F. Gelbukh, I.A. Bolshakov, S.N. Galicia-Haro*

"Automated Concept Extraction From Plain Text" *Boris Gelfand, Marilyn Wulfekuhler, William F. Punch III*

"A Probabilistic Approach to Full-Text Document Clustering" *Moises Goldszmidt, Mehran Sahami*

"Text Mining the Web: Extracting Chemical Compound Names" *Richard D. Hull, Eugene M. Fluder*

"Automated Building of a Database of Neural Networks Papers" *Ath. Kehagias, Vas. Petridis*

"The Effect of Relevant Input Information in ID3's Learning Performance" *Young S. Kwon, Nam H. Kim*

"Information Extraction & Heuristics for Human-like Browsing" *Zoé Lacroix, Arnaud Sahuguet*

"Active Learning with Committees: Preliminary Results in Comparing Winnow and Perceptron in Text Categorization" *Ray Liere, Prasad Tadepalli*

"Tool For Measurement of Statistical Properties of Text" *Dr. P.P. Makagonov, Dr. M.A. Alexandrov*

"Learning for Information Extraction, Querying, and Recommending" *Raymond J. Mooney*

"Pool-Based Active Learning for Text Classification" *Kamal Nigam, Andrew McCallum*

"Learning Correlations between Linguistic Indicators and Semantic Constraints: Reuse of Context-Dependent Descriptions of Entities" *Dragomir R. Radev*

## W7: Robot Exploration and Learning

(Chairs: Howie Choset, Maja Matarić and Sebastian Thrun)

"Team-Partitioned, Opaque-Transition Reinforcement Learning" *Peter Stone, Manuela Veloso*

"A Learning Algorithm for Cooperating Autonomous Agents" *Christopher Zach*

"Learning Approaches in Cooperative Robotics: A Few Facts" *Claude Touzet*

"Reinforcement Learning of 3-D Object Recognition from Appearance" *L. Paletta, M. Prantl, A. Pinz*

"Robotic Exploration Via Partially Observable Markov Decision Processes" *Michael Littman*

"Learning Exploration Policies with Models" *Marco Wiering and Jürgen Schmidhuber*

"Learning Stationary Temporal Probabilistic Networks" *Daniel Nikovski*

"Socially Embedded Learning: A Learning Framework using Natural Communication between Robots and Humans" *Hideki Asoh and Toshihiro Matsui*

"Comparison of Reinforcement Learning Techniques for Automatic Behavior Programming" *J. Andrew Bagnell, Keith L. Doty, A. Antonio Arroyo*

"The Robot With Seven Brains: Modularity and Re-Use in Reinforcement Learning" *Malcolm Ryan*

"Contextual Control Policy Selection" *Jefferson A. Coelho Jr., Elizeth G. Araujo, Manfred Huber, Roderic A. Grupen*

"Innate Behaviors Can Train Supervised Neural Network Controllers" *Noel E. Sharkey*

"Efficient Representation of Environments from Collection of Images" *Ben Kröse*

"Path planning for unknown, partially-known, and changing environments" *Tony Stentz*

"Getting Reinforcement Learning to Work on Real Robots" *William Smart, Leslie Kaelbling*

"Learning, Hierarchies, Abstraction and Exploration in Robotics" *Bruce Digney*

"Exploration and Spatial Learning Research at NCARAI" *Brian Yamauchi, Alan Schultz, William Adams, and Kevin Graves*

"Localization with the Generalized Voronoi Graph" *Howie Choset, Keiji Nagatani*

"Distributed Map-making Using Online Generalized Voronoi Graphs" *James Jennings, Chris Kirkwood-Watts, Craig Tanis*

"On-board Science Understanding: NASA Ames' Efforts" *Ted L. Roush, Peter Cheeseman, Virginia Gulick, David Wolf, Paul Gazis, Gretchen Benedix, Wray Buntime, Clark Glymour, Liam Pedersen, Mark Ruzon, Jon Oliver*

## W8: The Neural Basis of Learning

(Chairs: David Touretzky, Tai Sing Lee and Bill Skaggs)

No Papers

# Conference on
# Automated Learning and Discovery

**June 11-13, 1998    Carnegie Mellon University**

## Invited plenary speakers

- Tom Dietterich
- Stuart Geman
- David Heckerman
- Michael Jordan
- Daryl Pregibon
- Herb Simon
- Robert Tibshirani

## Workshops

Learning Causal Bayesian Networks
*organized by Richard Scheines and Larry Wasserman*

Mixed-Media Databases
*organized by Shumeet Baluja, Christos Faloutsos, Alex Hauptmann and Michael Witbrock*

Machine Learning and Reinforcement Learning for Manufacturing
*organized by Sridhar Mahadevan and Andrew Moore*

Large-Scale Consumer Databases
*organized by Mike Meyer, Teddy Seidenfeld and Kannan Srinivasan*

Visual Methods for the Study of Massive Data Sets
*organized by Bill Eddy and Steve Eick*

Learning from Text and the Web
*organized by Jaime Carbonell, Steve Fienberg, Tom Mitchell and Yiming Yang*

Robot Exploration and Learning
*organized by Howie Choset, Maja Matarić and Sebastian Thrun*

The Neural Basis of Learning
*organized by David Touretzky, Tai Sing Lee and Bill Skaggs*