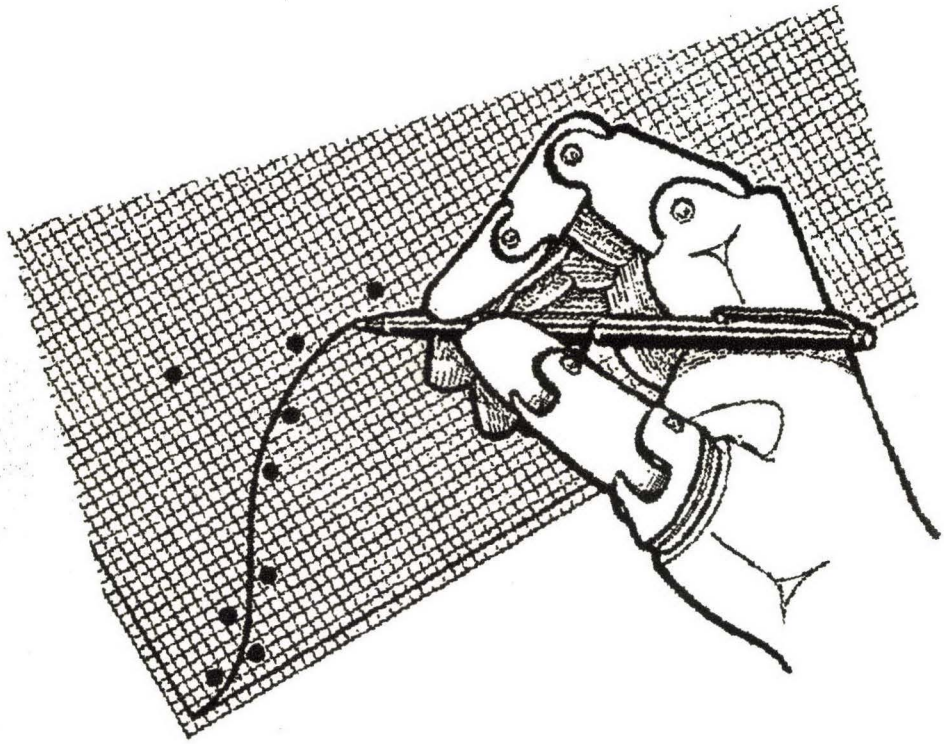
Preliminary Papers

of the

Sixth International Workshop on Artificial Intelligence and Statistics



January 4-7, 1997

Fort Lauderdale, Florida

Preface

Welcome to the Sixth International Workshop on Artificial Intelligence and Statistics.

This workshop is unique in bringing together statisticians and researchers in AI to discuss "common ground" in an interdisciplinary manner. The workshop series began in 1985 when Bill Gale and Daryl Pregibon then of AT&T conducted the first small workshop in the series. Workshops have subsequently taken place every 2 years, moving from the wintry woods of New Jersey south to the warmer beachfronts of Fort Lauderdale. Whether it was the locale or the workshop content, interest and participation have grown since that first workshop. Nonetheless, a key feature of this workshop is that it has remained relatively small and informal. Through the years this has made this workshop one of the more enjoyable and simultaneously more worthwhile workshops on what can be a very busy workshop/conference calendar for many of us.

The technical focus of the workshop can be hard to describe precisely, and indeed it has evolved over the years. One good example of the AI-Statistics intersection, at the last two workshops in particular, is graphical models. Graphical models are an excellent example of how Statistics and AI researchers can both contribute substantially to a field but in complementary ways. There are a wide variety of other topics being discussed in the various papers at this workshop, including model complexity, Markov models, neural networks, decision trees, Bayesian inference, causality, information retrieval, classification, and data analysis modeling. While the technical content of the papers is broad and varied there is an ever-present common thread, namely, the investigation of intelligence and automation within a statistical context. This reflects ongoing shifts in research thinking in both AI and Statistics. In Statistics, computational techniques are revolutionizing the field: a fundamental issue in this context is the interface between the data analyst and automated inference methods. Likewise, the field of Artificial Intelligence has in the last 10 years realized that uncertainty is critically important (necessary, but usually not sufficient) for building intelligent systems which interact with the real world. Thus, probabilistic and statistical methods are an increasingly important component in AI approaches to learning, planning, reasoning, vision, and so forth. The papers at this meeting reflect the dynamism and activity in research at the intersection of AI and Statistics.

This year we have 18 plenary papers and 38 poster papers being presented from January 5th to 7th. There is also a day of what promise to be excellent tutorials on January 4th. We thank all the paper, poster, and tutorial presenters for their participation. Thanks also to the Program Committee members for reviewing papers and providing feedback to the authors.

Finally, we would like to acknowledge the contributions of various people behind the scenes whose assistance was invaluable in making this workshop happen. Daryl Pregibon, both as treasurer and as unofficial local arrangements gopher, made the general chair's job much easier by taking care of many of the financial and contractual details. Doug Fisher, general chair of the last meeting, provided a broad range of advice and help on various logistical items. The Information and Computer Science Department at UC Irvine (Mike Pazzani, Department Chair and Bernie Bender, Administration) donated administrative support to handle registration, badges, etc.

David Madigan, Program Chair
Padhraic Smyth, General Chair

Programme Committee

General Chair:	Padhraic Smyth	U.C. Irvine and JPL
Programme Chair:	David Madigan	U. Washington

Members:

Russell Almond,	Educational Testing Service
Wray Buntine,	Ultimode Systems, LLC
Peter Cheeseman,	NASA
Paul Cohen,	University of Massachusetts
Greg Cooper,	University of Pittsburgh
Bill DuMouchel,	AT&T Labs
Doug Fisher,	Vanderbilt University
Dan Geiger,	Technion
Clark Glymour,	Carnegie-Mellon University
David Hand,	Open University
Steve Hanks,	University of Washington
Trevor Hastie,	Stanford University
David Haussler,	UC Santa Cruz
David Heckerman,	Microsoft
Paula Hietala,	University of Tampere
Geoff Hinton,	University of Toronto
Mike Jordan,	MIT
Hans Lenz,	Free University of Berlin
David Lewis,	AT&T Labs
Andrew Moore,	Carnegie-Mellon University
Radford Neal,	University of Toronto
Jonathan Oliver,	Monash University
Steve Omohundro,	NEC Research, Princeton
Judea Pearl,	UCLA
Daryl Pregibon,	AT&T Labs
Ross Shachter,	Stanford University
Glenn Shafer,	Rutgers University
Prakash Shenoy,	University of Kansas
David Spiegelhalter,	MRC, Cambridge
Peter Spirtes,	Carnegie-Mellon University

Reviewer Acknowledgement

Craig Aumann	U. Washington
David Dowe	Monash University

Sponsors

Society for Artificial Intelligence and Statistics
International Association for Statistical Computing

AISTATS-97 Program

Saturday Jan 4

8:30-11:30 Tutorial A

Conditional Independence for Statistics and AI
A. P. Dawid, University College London

12:30-3:30 Tutorial B

Bayesian Time Series Analysis and Forecasting
Mike West, Duke University

12:30-3:30 Tutorial B

Learning in Information Agents
Tom M. Mitchell, Carnegie Mellon University

4:00-7:00 Tutorial C

Graphical models, neural networks and machine learning algorithms
Michael Jordan, MIT

Sunday Jan 5

7:30 to 8:45 CONTINENTAL BREAKFAST/REGISTRATION

8:45 to 9:00 OPENING COMMENTS

9:00 to 10:30 SESSION 1:

- *A Bayesian approach to CART*
  Hugh Chipman, Edward I George, & Robert E. McCulloch
- *A comparison of scientific and engineering criteria for Bayesian model selection*
  David Heckerman & David Maxwell Chickering
- *Strategies for model mixing in generalized linear models*
  Merlise Clyde

10:30 to 11:00 COFFEE BREAK

11:00 to 12:00 SESSION 2:

- *Variational inference for continuous sigmoidal belief networks*
  Brendan J. Frey
- *Extensions of undirected and acyclic, directed graphical models*
  Thomas Richardson

12:00 to 1:00 LUNCH (provided)

1:00 to 4:00 BREAK

4:00 to 6:00 POSTER SUMMARIES

6:00 to 7:00 DINNER

7:00 to 9:30 POSTER SESSIONS

Monday Jan 6

8:00 to 9:00 CONTINENTAL BREAKFAST

9:00 to 10:30 SESSION 3:

A note on cyclic graphs and dynamical feedback systems
Thomas Richardson, Peter Spirtes, & Clark Glymour

Estimating Latent Causal Inferences: Tetrad II model selection and Bayesian parameter estimation
Richard Scheines

Using classification trees to improve causal inferences in observational studies
Louis Anthony Cox

10:30 to 11:00 COFFEE BREAK

11:00 to 12:30 SESSION 4:

Building an EDA Assistant: A Progress Report
Robert St. Amant & Paul R. Cohen

Mixed memory Markov models
Lawrence K. Saul & Michael I Jordan

Wavelet based random densities
David Rios Insua & Brani Vidakovic

12:30 to 2:00 LUNCH (provided)

2:00 to 3:30 SESSION 5:

Using Prediction to Improve Combinatorial Optimization Search
Justin A. Boyan & Andrew W. Moore

Inference using Probabilistic Concept Trees
Doug Fisher & Doug Talbert

The Effects of Training Set Size on Decision Tree Complexity
Tim Oates & David Jensen

3:30 to 4:30 BREAK

4:00 to 5:00 BUSINESS MEETING

Tuesday Jan 7

8:30 to 9:00 CONTINENTAL BREAKFAST

9:00 to 10:30 SESSION 6:

WWW Cache Layout to Ease Network Overload
Kenichi Yoshida

PAC learning with constant-partition classification noise and applications to decision tree induction
Scott E. Decatur

Graphical Model Based Computer Adaptive Testing
Russell G. Almond & Robert J. Mislevy

10:30 to 11:00 COFFEE BREAK

11:00 to 12:00 SESSION 7:

Asessing and Improving Classification Rules
David J. Hand, Keming Yu, & Niall Ada

A variational approach to Bayesian logistic regression models and their extensions
Tommi S. Jaakkola & Michael I. Jordan

12:00 to 12:15 CLOSING COMMENTS

Table of Contents

Intelligent Support of Secondary Data Analysis 1
Russell G. Almond

Graphical Model Based Computer Adaptive Testing 11
Russell G. Almond & Robert J. Mislevy

Finding Overlapping Distributions with MML 23
Rohan A. Baxter & Jonathan J. Oliver

Markov chain Monte Carlo methods for decision analysis 31
Concha Bielza, Peter Muller, & David Rios Insua

A Comparison of decision trees, influence diagrams, & valuation networks for Asymmetric Decision Problems 39
Concha Bielza & Prakash P.Shenoy

Integrating Signal and Language Context to Improve Handwritten Phrase Recognition: Alternative Approaches 47
Djamel Bouchaffra, Eugene Koontz, V. Krpasundar, Rohini K. Srihari, & Sargur N. Srihari,

Using Prediction to Improve Combinatorial Optimization Search 55
Justin A. Boyan & Andrew W. Moore

Comparing Tree-Simplification Procedures 67
Leonard A. Breslow & David W. Aha

A Forward Monte Carlo Method for Solving Influence Diagrams using local Computation 75
John M. Charnes & Prakash P. Shenoy

An algorithm for Bayesian network construction from data 83
Jie Cheng, David A. Bell, & Weiru Liu

A Bayesian approach to CART 91
Hugh Chipman, Edward I George, & Robert E. McCulloch

Strategies for model mixing in generalized linear models 103
Merlise Clyde

Overfitting Explained 115
Paul R. Cohen & David Jensen

Using classification trees to improve causal inferences in observational studies 123
Louis Anthony Cox

Dataset cataloging metadata for machine learning applications research 139
Sally Jo Cunningham

PAC learning with constant-partition classification noise and applications to decision tree induction 147
Scott E. Decatur

Bayesian model averaging in rule induction 157
Pedro Domingos

Memory Based Stochastic Optimization for Validation and Tuning of Function Approximators 165
Artur Dubrawski & Jeff Schneider

Inductive Inference of First-Order Models from Numeric-Symbolic Data 173
Floriana Esposito, Sergio Caggese, Donato Malerba, & Giovanni Semeraro

Learning Influence Diagram from Data 183
Kazuo J. Ezawa & Narendra K. Gupta

Inference using Probabilistic Concept Trees 191
Doug Fisher & Doug Talbert

A Characterization of Bayesian Network Structures and its Application to Learning 203
James I.G. Forbes

Variational inference for continuous sigmoidal belief networks 211
Brendan J. Frey

Multivariate Density Factorization for Independent Component Analysis: An Unsupervised Artificial Neural Network Approach 223
Mark Girolami & Colin Fyfe

Intelligent Assistant for Computational Scientists: Integrated modelling, experimentation, and analysis 231
Dawn E. Gregory & Paul R. Cohen

On Predictive Classification of Binary Vectors 239
Mats Gyllenberg & Timo Koski

Asessing and Improving Classification Rules 243
David J. Hand, Keming Yu, & Niall Ada

Robust interpretation of neural network models 255
Orna Intrator & Nathan Intrator

Wavelet based random densities 263
David Rios Insua & Brani Vidakovic

A comparison of scientific and engineering criteria for Bayesian model selection 275
David Heckerman & David Maxwell Chickering

A variational approach to Bayesian logistic regression models and their extensions 283
Tommi S. Jaakkola & Michael I. Jordan

Adjusting for Multiple Testing in Decision Tree Pruning 295
David Jensen

Bayesian Information Retrieval: Preliminary Evaluation 303
Michelle Keim, David D. Lewis, & David Madigan

Comparing predictive inference methods for discrete domains 311
Petri Kontkanen, Petri Myllymaki, Tomi Silander, Henry Tirri, & Peter Grunwald

Approximate Inference and Forecast Algorithms in Graphical Models for Partially Observed Dynamic Systems 319
Alberto Lekuona, Beatrix Lacruz, & Pilar Lasala

Unsupervised clustering with numeric and nominal mixed data - a new similarity based system 327
Cen Li & Gautam Biswas

How to Find Big-OH in Your Data Set (and How Not To) 347
C.C. McGeoch & P.R. Cohen

An objective function for belief net triangulation 355
Marina Meila & Michael I. Jordan

Combining neural network regression estimates using principal components 363
Christopher J Merz & Michael J. Pazzani

A family of algorithms for finding temporal structure in data 371
Tim Oates, Matthew J. Schmill, David Jensen, & Paul R. Cohen

The Effects of Training Set Size on Decision Tree Complexity 379
Tim Oates & David Jensen

Case-based Probability Factoring in Bayesian Belief Networks 391
Luigi Portinale

Robust parameter learning in Bayesian networks with missing data 399
Marco Ramoni & Paola Sebastiani

Extensions of undirected and acyclic, directed graphical models 407
Thomas Richardson

A note on cyclic graphs and dynamical feedback systems 421
Thomas Richardson, Peter Spirtes, & Clark Glymour

Applying a Gaussian-Bernoulli Mixture Model Network to Binary and Continuous Missing Data in Medicine 429
David B. Rosen & Harry B. Burke

Mixed memory Markov models 437
Lawrence K. Saul & Michael I Jordan

Estimating Latent Causal Inferences: Tetrad II model selection and Bayesian parameter estimation 445
Richard Scheines

A Distance Metric for Classification Trees 457
William D. Shannon & David Banks

An Incremental Construction of a Nonparametric Regression Models 465
Jan Smid and Petr Volf

Cross-validated likelihood for model selection in unsupervised learning 473
Padhraic Smyth

Heuristic greedy search algorithms for latent variable models 481
Peter Spirtes, Thomas Richardson, & Christopher Meek

A polynomial time algorithm for determining DAG equivalence in the presence of latent variables and selection bias 489
Peter Spirtes & Thomas Richardson

Building an EDA Assistant: A Progress Report 501
Robert St. Amant & Paul R. Cohen

On the error probability of model selection for classification 513
Joe Suzuki

Statistical Aspects of Classification in Drifting Populations 521
C.C. Taylor, G. Nakhaeizadeh, & G. Kunisch

MML Mixture modelling of multi-state, Poisson, vonMises circular, & Gaussian distributions 529
Chris S. Wallace & David L. Dowe

WWW Cache Layout to Ease Network Overload 537
Kenichi Yoshida