VASANT G. HONAVAR

Dr. Vasant Honavar
Department of Computer Science
226 Atanasoff Hall
Iowa State University
Ames, Iowa 50011-1040.

honavar@cs.iastate.edu
Voice: 515 294 1098
Fax: 515 294 0258
http://www.cs.iastate.edu/~honavar/

EDUCATION

University of Wisconsin-Madison, USA	Computer Science (Advisor: Leonard Uhr)	Ph.D., 1990
University of Wisconsin-Madison, USA	Computer Science	M.S., 1989

PRIMARY TEACHING AND RESEARCH INTERESTS

Artificial Intelligence, Machine Learning, Bioinformatics and Computational Systems Biology, Data Mining, Information Integration, Knowledge Representation and Inference, Semantic Web.

RESEARCH AND TEACHING POSITIONS HELD

2001-	Professor of Computer Science	Iowa State University
2005-	Director, Center for Computational Intelligence, Learning, and Discovery	Iowa State University
2002	Visiting Professor, Medical Informatics & Biological Statistics	University of Wisconsin
1996-2001	Associate Professor, Computer Science	Iowa State University
1998	Visiting Professor, Computer Science	Carnegie Mellon University
1990-1996	Assistant Professor, Computer Science	Iowa State University

INTERDEPARTMENTAL AND INTERDISCIPLINARY PROGRAM PARTICIPATION

2004-	Faculty member, Center for Computational Intelligence, Learning, & Discovery	Iowa State University
2003-2005	Chair, Bioinformatics and Computational Biology Graduate Program	Iowa State University
1999-	Faculty member, Bioinformatics & Computational Biology Graduate Program	Iowa State University
1999-	Faculty member, Laurence H. Baker Center for Bioinformatics and Biological Statistics	Iowa State University
1999-	Faculty member, Computational Molecular Biology Training Group (Supported by an NSF IGERT award)	Iowa State University
2002-2005	Faculty member, Computational Biology for Animal Agriculture Training Group (Supported by a USDA MGET award)	Iowa State University
2007-	Faculty Member, Institute for Combinatorial Discovery	Iowa State University
2007-	Faculty Member, Cyberinnovation Institute	Iowa State University
2003-	Faculty Member, Human Computer Interaction Graduate Program	Iowa State University
2001-	Faculty Member, Information Assurance Masters Program	Iowa State University
2001-	Faculty Member, Information Assurance Center	Iowa State University
2002-	Member, Institute for Science and Society	Iowa State University
1992-	Faculty member, Neuroscience	Iowa State University

My primary responsibilities at Iowa State University during 1990-2007 have been Research (50%), Teaching (50%), and Service.

CAREER HIGHLIGHTS

- Significant research contributions in Computer Science in Machine learning, Data Mining, and semantic web and interdisciplinary contributions in Bioinformatics and Computational Biology reported in 200 refereed journal and conference papers including several that are "highly cited" "highly viewed" or recipients of best paper awards, 20 book chapters, and 1 research monograph during 1990-2007
- Founding the Center for Computational Intelligence, Learning and Discovery at Iowa State University focused on fundamental and applied research in informatics infrastructure for data integration, data mining, and services for emerging data-rich applications e.g., in e-science, and serving as its director during 2005-present
- Securing approximately $18 million in research and training grants as PI or co-PI during 1990-2007; Currently serving as PI or Co-PI on approximately $7.2 million in research and training grants (PI on $2 million)
- Mentoring 12 Ph.D. graduates (6 employed as tenure-track or tenured faculty in academia, 2 employed in academic research positions, and 4 employed in industrial research and development), 20 MS graduates (all employed in industry); Currently mentoring or jointly mentoring 18 Ph.D. students.
- Establishing together with colleagues Drena Dobbs and Dan Voytas, the interdepartmental graduate program in Bioinformatics and Computational Biology at ISU, currently one of the top Bioinformatics programs in the US, with support from an NSF Integrative Graduate Education and Research Training (IGERT) award in 2000, renewed for a second 5-year period in 2005, and serving as associate chair of the program during 2001-2003 and chair during 2003-2005.
- Leading a major curriculum development effort focused on required core courses for the graduate program in Bioinformatics and Computational Biology, culminating in the creation of a Bioinformatics graduate core curriculum comprising of 4 core courses covering Genomics to Systems Biology
- Establishing, working with colleagues Drena Dobbs, Karin Dorman, Duane Enger, Stephen Willson, and David Fernandez-Baca, an interdepartmental undergraduate degree program in Bioinformatics and Computational Biology in 2007.
- Serving as charter member of NIH study section on Biological Data Management and Analysis during 2004-2007 and as a member of an ad hoc special study section with a similar focus during 2001-2004.
- Establishing and over the years, refining and broadening graduate and undergraduate curricula in Artificial Intelligence and Machine Learning at Iowa State University.
- Serving on editorial boards of several journals, organizing several scientifc conferences and workshops, and serving on the scientific program committees of major conferences in artificial intelligence, machine learning, data mining, semantic web, and bioinformatics

BIOGRAPHICAL SUMMARY

Dr. Vasant Honavar received a B.E. in Electronics Engineering from Bangalore University, India, an M.S. in Electrical and Computer Engineering from Drexel University, and an M.S. and a Ph.D. in Computer Science from the University of Wisconsin, Madison. He founded (in 1990) and has been the director of the Artificial Intelligence Research Laboratory at Iowa State University (ISU) where he is currently a professor of Computer Science. He directs the Center for Computational Intelligence, Learning & Discovery which he founded in 2004. Honavar is on the faculty of interdepartmental graduate programs in Bioinformatics and Computational Biology, Human-Computer Interaction, Neuroscience, and Information Assurance. He has served as the associate chair (2001-2003) and chair (2003-2005) of the Bioinformatics and Computational Biology Graduate Program which he helped establish at ISU with support from an Integrative Graduate Education and Research Training (IGERT) award from the National Science Foundation. Twelve Ph.D. students and twentyone M.S. students have graduated under Honavar's supervision.

Honavar's research and teaching interests include Artificial Intelligence, Machine Learning, Bioinformatics, Computational Molecular Biology, Intelligent Agents and Multi-agent systems, Collaborative Information Systems, Semantic Web, Environmental Informatics, Security Informatics, Social Informatics, Neural Computation, Systems Biology, Data Mining, Knowledge Discovery and Visualization. Honavar has published over 180 research articles in refereed journals, conferences and books, and has co-edited several books.

Honavar is a co-editor-in-chief of the Journal of Cognitive Systems Research. He has served on, or currently serves on the editorial boards of the Machine Learning Journal, the Journal of Bioinformatics and Biology Insights, the International Journal of Semantic Web and Information Systems, the International Journal of Computational Biology and Drug Design, the International Journal of Computer and Information Security, and the International Journal of Data Mining and Bioinformatics. Honavar has served on the program committees of several major conferences in artificial intelligence, data mining, and bioinformatics including in particular, International Conference on Machine Learning (ICML), ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), SIAM Conference on Data Mining (SDM), IEEE Conference on Data Mining (ICDM), IEEE Conference on Tools With Artificial Intelligence (ICTAI), the National Conference on Artificial Intelligence (AAAI), ACM/IEEE Conference on Intelligent Agent Technology (IAT), Intelligent Systems in Molecular Biology (ISMB), IEEE Conference on Bioinformatics and Bioengineering (BIBE), IEEE Conference on Bioinformatics and Biomedicine (BIBM), Workshop on Algorithms in Bioinformatics (WABI), among others.

Honavar has served as a PI or Co-PI on research grants totaling approximately $18.0 million from various sources including the National Science Foundation, the National Institutes of Health million during 1990-2007. He currently serves as a PI or co-PI on grants totaling approximately $7.2 million (PI on $2 million). Honavar has extensive curriculum development and teaching experience in Artificial Intelligence, Machine Learning, and Bioinformatics. Honavar is a member of the National Institutes of Health study section on Biological Data Management and Analysis. Prof. Honavar is a senior member of the Association for Computing Machinery (ACM), Association for Advancement of Artificial Intelligence (AAAI), International Society for Computational Biology (ISCB) and the New York Academy of Sciences, and a senior member of the Institute of Electrical and Electronic Engineers (IEEE).

EDITORIAL BOARD MEMBERSHIP

2007-	Member of Editorial Board	Journal of Bioinformatics and Biology Insights
2007-	Member of Editorial Board	International Journal of Computational Biology and Drug Design
2007-	Member of Review Board	Applied Intelligence Journal
2006-	Member of Editorial Board	International Journal of Semantic Web and Information Systems
2006-	Member of Editorial Board	Springer-Verlag Book Series on Advanced Information and Knowledge Processing
2002-2005	Member of Editorial Board	Machine Learning Journal
1999-	Co-Editor-in-Chief	Cognitive Systems Research
2005-	Member of Editorial Board	International Journal of Data Mining and Bioinformatics
2004-	Member of Editorial Board	International Journal of Information and Computer Security
2001	Guest Editor	Machine Learning Journal (Grammar Inference)

STUDY SECTION AND REVIEW PANEL MEMBERSHIP

2004-2007	Member of Study Section, Biological Data Management & Analysis	National Institutes of Health
2006	CISE Intelligent Information Systems Review Panel	National Science Foundation
2005	CISE Intelligent Information Systems Review Panel	National Science Foundation
2004	Member of Special Study Section, NIH Roadmap Initiative	National Institutes of Health
2004	Member of Special Study Section, Data Sharing Initiative	National Institutes of Health
2001-2004	Member of Special Study Section, Bioinformatics	National Institutes of Health
2003	CISE Intelligent Information Systems Review Panel	National Science Foundation
2003	CISE CAREER Review Panel	National Science Foundation
2003	SBIR Review Panel	National Science Foundation
2003	SBIR Review Panel	National Science Foundation
1999	IGERT Review Panel	National Science Foundation
2002	CISE Intelligent Information Systems Review Panel	National Science Foundation
1994	Research Instrumentation Review Panel	National Science Foundation

HONORS AND AWARDS

2007	Regents Award for Faculty Excellence, Board of Regents, Iowa
2006	Best Paper Award, Asian Semantic Web Conference (ASWC 2006)
2006	Best Paper Award, IEEE ICTAI 2006
1994-99	Research Initiation Award, National Science Foundation
1994	Who's who in Science and Engineering
1992	Elected Member, Sigma Xi
1990-	Associate, Behavior and Brain Sciences
1990	Elected Member, New York Academy of Sciences
1990	Fellow, Workshop on Human and Machine Cognition
1989	Fellowship, Summer Institute in Parallel Computing, Argonne National Laboratory
1989	Student Fellowship, International Joint Conference on Artificial Intelligence (IJCAI)
1989	Fellowship, McDonnell Summer Institute in Cognitive Neuroscience, Dartmouth
1988	Fellowship, Connectionist Models Summer School, Carnegie Mellon University
1982	Gold medal for academic excellence, Bangalore University, India
1975-82	National Merit Scholar, India

PROFESSIONAL AFFILIATIONS

1986-	Member, AAAI (Association for Advancement of Artificial Intelligence)
2007-	Senior Member, ACM (Association for Computing Machinery)
2005-	Senior Member, IEEE (Institute of Electrical and Electronic Engineers)
1990-	Member, Institution of Electrical and Electronic Engineers (IEEE)
1990-	Member, IEEE Computer Society
1990-2007	Member, Association of Computing Machinery (ACM)
1992-	Member, American Association for Advancement of Science
1998-	Member, Cognitive Science Society
2003-	Member, International Society for Computational Biology

ADMINISTRATIVE EXPERIENCE

2005-	Director	Center for Computational Intelligence, Learning, and Discovery		Iowa State University
2003-2005	Chair	Interdepartmental Bioinformatics & Computational Biology Graduate Program		Iowa State University
2004-2005	Founder and Director	Computational Intelligence, Learning, & Discovery Program		Iowa State University
2003-2004	Chair	Research Space Committee	Computer Science	Iowa State University
2002-	Director of Research	Computer Science		Iowa State University
2001-2002	Chair	Graduate Admissions Committee	Computer Science	Iowa State University
2001-2003	Associate Chair	Bioinformatics & Computational Biology Graduate Program		Iowa State University
2000-2001	Chair	Graduate Admissions Committee	Computer Science	Iowa State University
2001	Co-Chair	Strategic Planning Committee	Computer Science	Iowa State University
2000-2001	Chair	Graduate Admissions Committee	Bioinformatics & Computational Biology Program	Iowa State University
1999-2000	Chair	Graduate Admissions Committee	Bioinformatics & Computational Biology Program	Iowa State University
1999-2000	Chair	Graduate Admissions Committee	Computer Science	Iowa State University
1999-2000	Chair	Faculty Search Committee	Computer Science	Iowa State University
1990-	Founder and Head	Artificial Intelligence Research Laboratory		Iowa State University

CONSULTING EXPERIENCE

Scientific consulting on data mining, artificial intelligence, web services, bioinformatics, knowledge-based systems, and information technology for several corporations, startups, and government organizations.

PRE-DOCTORAL APPOINTMENTS

1986-1990	Research Assistant, Computer Science	University of Wisconsin
1984-1986	Teaching Assistant, Electrical and Computer Engineering	University of Wisconsin

RESEARCH INTERESTS

Research Statement

My research interests cut across Computer Science, Information Science, Statistics, Cognitive Science, and Biological Sciences. This research is driven by fundamental scientific questions or important practical problems such as the following:

(a) What are the information requirements and algorithmic basis of learning in specific scenarios?

(b) What are the information requirements and algorithmic basis of inter-agent communication, multi-agent interaction, coordination, and organization?

(c) How can we develop sophisticated machine learning algorithms for knowledge acquisition from richly structured data (sequences, images, graphs, etc.)?

(d) How is information encoded, stored, retrieved, decoded, and used in macromolecular, neural, and cognitive systems?

(e) How can we discover the relationships between macromolecular sequence, structure, expression, interaction and macromolecular function?

(f) How can we construct accurate predictive models of signaling networks involved in cellular development, differentiation, and biological function?

(g) How can we query and use information from heterogeneous, distributed, autonomous data and knowledge sources?

(h) How can we build useful predictive models from large, distributed, semantically heterogeneous, autonomous data sources?

(i) How can we develop software environments for collaborative development, sharing, and use of large, complex, ontologies?

(j) How can we support the design, assembly and execution of complex web services using autonomously developed components?

(k) How can we represent and manipulate scientific knowledge in a form that lends itself to automated processing by the computer and at the same time, is comprehensible by, and communicable to humans?

Current Research Interests

(a) **Artificial Intelligence:** Intelligent agent architectures, Multi-agent organizations, Inter-agent interaction, and Multi-agent coordination, Logical, probabilistic, and decision-theoretic knowledge representation and inference, Neural architectures for knowledge representation and inference, Computational models of perception and action

(b) **Bioinformatics and Computational Molecular Biology:** Data-driven discovery of macromolecular sequence-structure-function-interaction-expression relationships, identification of sequence and structural correlates of protein-protein , protein-RNA, and protein-DNA interactions, protein sub-cellular localization, automated protein structure and function annotation, modeling and inference of genetic regulatory networks from gene expression (micro-array, proteomics) data, modeling and inference of signal transduction and metabolic pathways.

(c) **Data Mining:** Design, analysis, implementation, and evaluation of algorithms and software for data-driven knowledge acquisition, data and knowledge visualization, and collaborative scientific discovery from semantically heterogeneous, distributed data and knowledge sources, Applications to data-driven knowledge acquisition tasks in bioinformatics, medical informatics, geo-informatics, environmental informatics, chemo-informatics, security informatics, social informatics, critical national infrastructure (communication networks, energy networks) e-government, e-commerce, and e-science.

(d) **Machine Learning:** Statistical, information theoretic, linguistic and structural approaches to machine learning, Learning and refinement of bayesian networks, causal networks, decision networks, neural networks, support vector machines, kernel classifiers,, multi-relational models, language models (n-grams, grammars, automata), Learning classifiers from attribute value taxonomies and partially specified data; Learning attribute value taxonomies from data; Learning classifiers from sequential and spatial data; Learning relationships from multi-modal data (e.g., text, images), Learning classifiers from distributed data, multi-relational data, and semantically heterogeneous data; Incremental learning, Ensemble methods, multi-agent learning, selected topics in computational learning theory.

(e) **Semantic Web:** Ontology-based user and query-centric approaches to information integration and acquisition of sufficient statistics for learning from data under different access and resource constraints from heterogeneous, distributed, autonomous, ubiquitous information sources, sensor networks, peer-to peer networks; description logics, ontology design, ontology tools, ontology-extended information sources, ontology-extended workflow components, ontology-extended agents and services, web service composition.

(f) **Other Topics of Interest:** Biological Computation – Evolutionary, Cellular and Neural Computation, Complex Adaptive Systems, Sensory systems and behavior evolution, Language evolution, Mimetic evolution; Computational Semiotics – Origins and use of signs, emergence of semantics; Computational organization theory; Computational Neuroscience; Computational models of creativity, Computational models of discovery.

RESEARCH , TRAINING, AND INFRASTRUCTURE GRANTS

Current Research, Training, and Infrastructure Grants

1. Collaborative Research: Learning Classifiers from Autonomous, Semantically Heterogeneous Distributed Data Sources, National Science Foundation, Vasant Honavar (PI) $449, 000. 2007-2010.
2. Discovering Protein Sequence-Structure-Function Relationships. National Institutes of Health Vasant Honavar (PI), Robert Jernigan and Drena Dobbs (Co-PIs), (2003-2008). $1,022,000.
3. Interactive and Verifiable Composition of Web Services To Satisfy End User Goals. National Science Foundation, Vasant Honavar (Co-PI), with Samik Basu (PI) and Robyn Lutz (Co-PI). (2007-2010), $350,002.
4. IIS: Exploratory Investigation of Modular Ontologies. *National Science Foundation*, Vasant Honavar (PI), Giora Slutzki and Doina Caragea (Co-PIs), (2006-2008). $112,000.
5. NIH-NSF BBSI Summer Institute in Bioinformatics and Computational Biology - Iowa State University. National Science Foundation, Vasant Honavar (Co-PI) with Volker Brendel (PI), Robert Jernigan, Karin Dorman, and Julie Dickerson (Co-PIs) (2006--2009). $499,000.
6. High-Accuracy Protein Models Derived from Lower Resolution Data. National Institutes of Health (2007-2010), Vasant Honavar (Co-PI), with Andrzej Kloczkowski (PI), Robert Jernigan, Mark Gordon, Zhijun Wu, Iowa State University and Janusz Bujnicki, Krzysztof Ginalski and Andrzej Kolinski, Warsaw University (Co-PIs), $744,725.
7. Development of bioinformatics resources to transfer biological information across species. *United States Department of Agriculture.* Vasant Honavar (Co-PI), James Reecy (PI), Anne Kwitek (Co-PI). (2008-2010). $1,000,000.
8. DDAS-TMRP: Auto-Steered Information-Decision Processes for Electric Power Systems Asset Management. *National Science Foundation*, Vasant Honavar (Co-PI) with James McCalley (PI), Sarah Ryan (Co-PI), William Meeker (Co-PI), and Daji Qiao (Co-PI). (2006-2009) $700,000.
9. Integration of Functional Genomics and Quantitative Genetics to Improve Feed Efficiency in Pigs. United States Department of Agriculture (2005-2008), Vasant Honavar (Co-PI) with Jack Dekkers (PI), Chris Tuggle (Co-PI), Dan Nettleton (Co-PI), Lloyd Anderson (Co-PI), Rondhane Rekaya (Co-PI), Richard Barb (Co-PI), and Gary Hausman (Co-PI), $876,000.
10. IGERT: Computational Molecular Biology Training Program. Vasant Honavar (Co-PI) with Dan Voytas (PI), Susan Carpenter (Co-PIs). National Science Foundation, 2005-2010. $2,968,976.
11. Center for Computational Intelligence, Learning, and Discovery. Vasant Honavar (PI). Vice Provost for Research, Iowa State University (2004-2008) $377,500
12. Computational Support Staff for Expanding Animal Functional Genomics Capabilities. Vasant Honavar (PI), Chris Tuggle, Jim Reecy, Diane Spurlock, Jack Dekkers, Susan Lamont, Chad Stahl (Co-PIs), $50,000. (2007-2009).
13. Cyberinnovation Institute. Vasant Honavar (Co-PI), James Oliver (PI), Arun Somani, Krishna Rajan, Doug Jacobson (Co-PIs), $1,000,000. Iowa State University.

Pending Grant Proposals

1. Experimental Study and Computational Modeling of the IGF-1. Vasant Honavar (Co-PI) with Heather Greenlee (PI). Signaling Pathway during Rod Photoreceptor Development, National Institutes of Health, (2008-2011) $1,470,000.

2. Deciphering the Protein-RNA Recognition Code. Vasant Honavar (Co-PI) with Drena Dobbs (PI) and Robert Jernigan (Co-PI). National Institutes of Health. (2008-2011), $1,439,465
3. Structural Aspects of Phosphorylation Regulation. Vasant Honavar (Co-PI), with Andrzej Kloczkowski (PI), Robert Jernigan (Co-PI), and Drena Dobbs (Co-PI). National Institutes of Health (2008-2012), $2,000,000.

Past Research and Training Grants

1. ITR: Algorithms and Software for Knowledge Acquisition from Heterogeneous Distributed Data. *National Science Foundation*, Vasant Honavar (PI) Drena Dobbs (Co-PI), (2002-2007). $223,500.
2. Algorithms and Software for Collaborative Ontology Development. Center for Integrated Animal Genomics, Iowa State University. Vasant Honavar (PI), (2005-2007) $25,000.
3. IGERT: Computational Molecular Biology Training Program. Vasant Honavar (Co-PI) with Dan Voytas (PI), Pat Schnable, Susan Carpenter, Jonathan Wendel (Co-PIs). National Science Foundation, 1999-2004., $2,374,597 (plus $1,161,010 in matching funds).
4. Developmental Proteomics of Retinal Progenitor Cells, National Institutes of Health (2003-2006). Vasant Honavar (Co-PI), (with Heather West-Greenlee and Jan Buss), $438,000.
5. NIH-NSF BBSI Summer Institute in Bioinformatics and Computational Biology - Iowa State University. National Science Foundation, Vasant Honavar (Co-PI) with Volker Brendel (PI), Robert Jernigan, Karin Dorman, and Xun Gu (Co-PIs) (2002-2006). $645,000.
6. Automated Integration of Condition Monitoring with an Optimized Maintenance Scheduler for Circuit Breakers and Power Transformers. Vasant Honavar (Co-PI) with James McCalley (PI) Mladen Kezunovic, and Chanan Singh (Co-PIs), Power Systems Engineering Research Center (a National Science Foundation Industry-University Research Center), 2002-2005. (2002-2005). $255,000.
7. Symposim on Integration of Structural and functional Genomics, Vasant Honavar (co-PI) with Chris Tuggle (PI) and Marit Nielsen-Hamilton (Co-PI) National Science Foundation. (2005). $10,900.
8. Exploring a Novel Proline Switch for Regulation of Protein Recognition. Center for Integrated Animal Genomics, Iowa State University, Vasant Honavar (Co-PI) with Drena Dobbs and Susan Carpenter (Co-PIs) (2004-2005) $30,000.
9. Constructive Neural Network Learning Algorithms for Pattern Classification, National Science Foundation, 1994-1999, Principal Investigator, $111,537 (plus $10,000 in matching funds).
10. SGER: Multidisciplinary Aspects of Computation Theory, National Science Foundation, Vasant Honavar (Co-PI), with Jack Lutz (PI), Pavan Aduri (Co-PI), and Krishna Athreya (Co-PI), (2003-2005). $74,948.
11. Graduate Research Fellowships in Bioinformatics and Computational Biology, Pioneer Hi-Bred, Inc. 2002-2004. Major professor, Adrian Silvesu and Carson Andorf, $80,000.
12. IBM Graduate Research Fellowship in Computer Science, IBM Inc., Major Professor, Doina Caragea, 2003-2004, $30,000 (approx.)
13. Interactive Visual Overviews of Large, Multi-Dimensional Datasets, Vasant Honavar (Co-PI) with Diane Cook (PI) and Les Miller (Co-PI) National Science Foundation, Co-Principal Investigator, 1999-2003, $370,000.
14. Innovative Technologies for Defense Against Catastrophic Failures of Complex, Interactive Power Networks, U.S. Department of Defense (DOD) and Electric Power Research Institute (EPRI), $4,500,000. (1999 - 2004) (Collaborative project involving 9 ISU faculty and faculty from 3 other universities). Co-Principal Investigator.
15. An Agent-Based System for Integration and Analysis of Distributed, Heterogeneous Plant Genome Databases. Pioneer Hi-Bred International, Inc., 2000-2002, Principal Investigator, $40,000.
16. IBM Graduate Research Fellowship in Computer Science, IBM Inc., Vasant Honavar (Major Professor), with Doina Caragea (doctoral student), 2002-2003, $30,000 (approx.)
17. Intelligent Multi-Agent Systems for Intrusion Detection, National Security Agency, 1998-2000, Co-Principal Investigator, $199,769.

18. Data Mining of Electric Power Usage Data to Develop Customer Profiles. Cooperative Research Proposal. Power Domain, Inc. (2001-2002). Vasant Honavar. $43,639.
19. SGER: Distributed Knowledge Networks to Support Security-Economy Decisions in Stressed Electric Power Systems. National Science Foundation, 2000-2001, Co-Principal Investigator, $99,999.
20. Artificial Intelligence Applications to Power System Management and Control, Electric Power Research Institute, Co-principal investigator. 1998-2000, $151,000
21. Distributed Knowledge Networks, John Deere Foundation, 1999-2001, Principal Investigator, $30,000.
22. Development of Algorithmic Approaches to Gene Expression Analysis from Microarray Data, *Carver Foundation*, 2000-2001, Principal Investigator, $25,000.
23. Development of Protein Structure Prediction Algorithms. Carver Foundation, 1999-2000, Co-Principal Investigator, $25,000.
24. Genetic algorithms for protein structure prediction. Ames Laboratory, Co-Principal Investigator, 1999-2000. $35,700.
25. A Gene-Specific DNA Chip for Exploring Molecular Evolutionary Change, *Carver Foundation*, 1998-1999. Co-Principal Investigator, $17,120.
26. Intelligent Diagnosis Systems, John Deere Foundation, 1995-1998, Principal Investigator, $30,000.
27. Graduate Fellowship (Data Mining and Knowledge Discovery), IBM Corporation, 1997-1998, Principal Investigator, $20,800.

SELECTED RESEARCH ACCOMPLISHMENTS

Research Contributions in Machine learning, Data mining and Computational Learning Theory

- Development of scalable algorithms for learning predictive models (e.g., decision trees, bayesian network classifiers, support vector machines) from autonomous, distributed data sources using statistical queries with proven performance guarantees relative to their centralized counterparts (with Ph.D. student Doina Caragea)
- Development of algorithms for learning comprehensible predictive models from data and prior knowledge in the form of attribute value taxonomies (with Ph.D. students Jun Zhang, Adrian Silvescu, and Flavian Vasile)
- Development of algorithms for learning predictive models (e.g., decision tree and Bayesian network classifiers) from partially specified data, i.e., data at varying levels of abstraction (with Ph.D. student Jun Zhang)
- Development of algorithms for learning a family of sequence classifiers with applications in computational biology and computer security (with Ph.D. students Carson Andorf, Adrian Silvescu, and Dae-Ki Kang)
- Development of discriminatively trained probabilistic models for sequence classification (with Ph.D. student Oksana Yakhnenko)
- Development of generalized multiple instance learning algorithms with applications in bioinformatics, text and image analysis (with Ph.D. student Yasser El-Manzalawy)
- Development of multi-relational learning algorithms (with MS student Anna Atramentov)
- Development of independence-based Markov Network learning algorithms (with Facundo Bromberg and Dimitris Margaritis)
- Theoretical characterization of independence and decomposability of functions that take values into an Abelian group including probability distributions, energy functions, value functions, fitness functions, and relations (with Ph.D. student Adrian Silvescu)
- Development and analysis of a machine learning algorithm for inference of temporal boolean network models from multivariate time series data, with applications to inference of genetic networks from gene expression data (with Ph.D. student Adrian Silvescu)

- Development of machine learning methods based on tensor decomposition for discovery of communities, topics, etc. from web data (with Ph.D. student Flavian Vasile)
- Development of algorithms for learning predictive models from multi-relational, semantically heterogeneous data (with Ph.D. student Cornelia Caragea)
- Development of polynomial algorithms for learning regular languages from examples and membership queries (with Ph.D. student Rajesh Parekh and collaborator Codrin Nichitiu)
- Theoretical analysis of the relationship between various models of learning in helpful environments showing that a concept that is learnable under Gold's model for learning from characteristic samples, Goldman and Mathias' polynomial teachability model, and the model for learning from example based queries is also learnable under the PACS model (with Ph.D. student Rajesh Parekh)
- Establishing that simple DFA (i.e., DFA whose canonical representations have logarithmic Kolmogorov complexity) are efficiently PAC learnable under the Solomonoff Levin universal distribution; and that if the examples are sampled at random according to the universal distribution by a teacher that is knowledgeable about the target concept, DFA are efficiently PAC learnable under the universal distribution, thereby answering the open problem posed by L. Pitt in 1989: Are DFA PAC-identifiable if examples are drawn from the uniform distribution, or some other known simple distribution? (with Ph.D. student Rajesh Parekh)
- Development of evolutionary algorithms for feature subset selection for classification problems (with Ph.D. student Jihoon Yang)
- Development of evolutionary approaches to design of sensor systems for adaptive robots (with Ph.D. student Karthik Balakrishnan)
- Development of incremental neural network learning algorithms with applications in nondestructive evaluation (with Ph.D. student Robi Polikar and collaborators Satish Udpa and Lalita Udpa)
- Development of constructive neural network algorithms that take advantage of prior knowledge in the form of classification rules (with Ph.D. students Rajesh Parekh and Jihoon Yang)
- Generalization (with convergence guarantees) of a large family of such algorithms designed for 2-class binary pattern classification problems to handle classification problems involving real-valued patterns and an arbitrary number of classes (with Ph.D. students Rajesh Parekh and Jihoon Yang)
- Development of a simple, inter-pattern distance based provably convergent, polynomial time constructive neural network algorithm which compares very favorably with computationally far more expensive algorithms in terms of generalization accuracy (with Ph.D. students Jihoon Yang and Rajesh Parekh).

Research Contributions in Semantic Web, Information Integration, Knowledge Representation, Ontologies, Service-Oriented Computing

- Development of a family of description-logics based modular ontology languages (P-DL) that support selective sharing of knowledge and establishment of a minimal set of restrictions on the use of imported concepts and roles to support localized semantics, transitive propagation of imported knowledge, and different interpretations from the point of view of different ontology modules (with Ph.D. student Jie Bao and collaborators Giora Slutzki and George Voutsadakis)
- Development of a family of sound and complete tableau-based federated reasoning algorithms for distributed, autonomous, P-DL ontologies that support selective sharing of knowledge across ontology modules that avoid the need to integrate ontologies using message exchanges between modules as needed (with Ph.D. student Jie Bao and collaborators Giora Slutzki and George Voutsadakis)
- Development of a privacy-preserving reasoning framework for answering queries against ontologies using private knowledge without compromising private knowledge (with Ph.D. student Jie Bao and collaborator Giora Slutzki)
- Development and open-source implementation of an ontology-based system for querying multiple semantically disparate data sources from a user's point of view (with Ph.D. students Doina Caragea, Jie Bao, and Neeraj Koul)

- Development of algorithms for specification-driven interactive and verifiable composition of composite web services from autonomous component services (with Ph.D. student Jyotishman Pathak and collaborator Samik Basu)
- Development of algorithms for context-specific substitution of one service by another within a composite service (with Ph.D. student Jyotishman Pathak and collaborator Samik Basu)

Research Contributions in bioinformatics, computational molecular biology, and systems biology

- Application of classifiers trained using machine learning to discover a large set of incorrect Gene Ontology annotations an experimentally well-studied family of proteins - mouse kinases (with Ph.D. student Carson Andorf)
- Development and applications of probabilistic graphical models and related methods for assigning protein sequences to functional families, predicting protein subcellular localization, etc. (with Ph.D. students Carson Andorf and Adrian Silvescu)
- Construction and analysis of PPIDB, a comprehensive database of protein-protein interfaces (with Ph.D. students Feihong Wu, Raphael Osorio and collaborator Drena Dobbs)
- Development of machine learning approaches to prediction of protein-protein interface residues from amino acid sequence, evolutionary and when available, structural information (with Ph.D. student Changhui Yan and collaborator Drena Dobbs and Robert Jernigan)
- Demonstration of the pitfalls of commonly used windows-based cross-validation for sequence-based classification tasks (e.g., phosphorylation site prediction, DNA-binding site prediction) (with Ph.D. student Cornelia Caragea)
- Development of machine learning approaches and implementation of online servers for prediction of protein-RNA interface residues from amino acid sequence and when available, structural information (with Ph.D. students Michael Terribilini, Cornelia Caragea, and collaborator Drena Dobbs)
- Development of machine learning approaches and implementation of online servers for prediction of protein-DNS interface residues from amino acid sequence, and when available, structural information (with Ph.D. student Changhui Yan, Cornelia Caragea, and collaborator Drena Dobbs)
- Structural characterization of protein-protein and protein-RNA interfaces (with Ph.D. students Feihong Wu and Fadi Towfic)
- Development of machine learning methods and online servers for identification of posttranslational modification sites e.g., phosphorylation sites, glycosylation sites in amino-acid sequences (with Ph.D. students Cornelia Caragea, Yasser El-Manzalawy)
- Development of kernel-based methods for predicting B-cell epitopes from amino acid sequences (with Ph.D. student Yasser El-Manzalawy and collaborator Drena Dobbs)
- Demonstrations of the pitfalls of commonly used benchmark datasets for evaluating the performance of machine learning approaches to epitope prediction (with Ph.D. student Yasser El-Manzalawy)
- Prediction of protein and RNA binding sites in recalcitrant (with regard to attempts at structure determination) proteins e.g., HIV-1 and EIAV and experimental confirmation of the predictions (with Ph.D. students Jae-Hyung Lee, Michael Terribilini and collaborators Drena Dobbs and Susan Carpenter).
- Development and application of an approach to combining homology modeling and structure prediction methods with machine learning to predict sequence and structural correlates of functionally important sites of telomerase (RNA, DNA, and protein binding sites) (with Ph.D. students Michael Terribilini, Jae-Hyung Lee, Cornelia Caragea, and collaborator Drena Dobbs)
- Characterization of gene expression changes during the onset of photosynthesis (with collaborator Steve Rodermel)
- Characterization of gene expression changes during differentiation of retinal stem cells into rod photoreceptors (with Ph.D. students Tim Alcon, Alison Barnhill, Laura Hecker and collaborators Heather Greenlee and Don Sakaguchi)

- Characterization differences in the proteome of murine retinal and brain derived progenitor cells (with Ph.D. students Tyra Dunn and collaborators Heather Greenlee and Drena Dobbs)
- Development of a collaborative phenotype ontology development environment (with Ph.D. students Jie Bao, LaRon Hughes, and collaborator James Reecy)
- Development of databases and software tools for capture, analysis, annotation, and integration of gene expression data with other types of 'omics' data (with Ph.D. students Neeraj Koul, Oliver Couture, and collaborator Chris Tuggle).
- Development of a method for interactive querying and analysis of multiple gene expression datasets using an experimentally verified gene network to expand the network, and to prioritize experimental targets (with Ph.D. student Tim Alcon and collaborator Heather Greenlee)
- Development of the Retina Workbench, a software tool for construction, analysis, and comparison of gene and protein networks (with MS student Oksana Kohutyuk, Ph.D. student Fadi Towfic, and collaborator Heather Greenlee)

Research Contributions in Neural and Cognitive Modeling

- Development of algorithms for construction of robust, noise-tolerant neural memories for pattern storage and associative, content-based retrieval (with Ph.D. student Chun-Hsien Chen)
- Development of algorithms for construction of highly parallel neural architectures for syntax analysis (parsing of regular, context-free, and context-sensitive languages) (with Ph.D. student Chun-Hsien Chen).
- Development of a biologically inspired neural architecture and an extended Kalman filter algorithm for place learning and localization in a-priori unknown environments which successfully accounts for a large body of behavioral and neurobiological data from animal experiments and offers several testable predictions (with Ph.D. student Karthik Balakrishnan, and collaborator Olivier Bousquet).

Research Contributions in Distributed Data Driven Applications, Computing and Communication Networks, Critical Infrastructure Monitoring and Protection

- Development of tools for formal specification of intrusions using colored Petri nets and software fault trees and methods for automated generation of multi-agent systems for coordinated intrusion detection in computer and communication networks (with Ph.D. student Guy Helmer, MS student Mark Slagell, and collaborators Johnny Wong, Robyn Lutz, and Les Miller)
- Development of multi-agent system for detection of coordinated or concerted attacks on distributed computing systems in particular by monitoring different processes, resources, users, events, and extract and integrate relevant information from disparate sources over multiple space and time scales (with Ph.D. student Guy Helmer and collaborators Johnny Wong, Robyn Lutz, and Les Miller)
- Development and application of machine learning approaches for learning predictive rules for anomaly and misuse detection (with Ph.D. students Dae-Ki Kang, Guy Helmer, and collaborator Johnny Wong).
- Development of an electronic nose for detection and identification of odorants using machine learning (with Ph.D student Robi Polikar and collaborators R. Shinar, L. Udpa, and M. Porter).
- Development and applications of machine learning methods for non-destructive inspection of nuclear power plant pipes using ultrasound (with Ph.D. student Robi Polikar and collaborators L. Udpa, and S. Udpa)
- Development of a service-oriented distributed software infrastructure for monitoring distributed power systems (with Ph.D. student Jyotishman Pathak and collaborator Jim McCalley)
- Development of statistical methods and software for monitoring and condition assessment of critical components of distributed power systems (with collaborator Jim McCalley)
- Development of distributed multi-agent systems for information integration and decision support in distributed power systems (with MS student Vijay Viswanathan and collaborator Jim McCalley)

- Development of infrastructure for multi-agent negotiation for power systems, e-commerce, and related applications (with collaborator Mokdong Chung)
- Development of a utility-theoretic approach to routing in communication networks that supports a flexible tradeoff between delay for a specific message and the overall network load (and hence expected delay for all routed messages) using a knowledge representation scheme that enables each node in a communication network to maintain and update a small constant-size knowledge base (independent of the network size)
- Theoretical and experimental analysis of utility-theoretic routing that showed the efficacy of the approach in minimizing message delay and load imbalance over the entire network without access to accurate global network state information.

Representative Current Research Projects

Algorithms and Software for Knowledge Acquisition from Semantically Heterogeneous, Distributed Data (funded in part by grants from the National Science Foundation (IIS 0219699 – 2002-2006; IIS 0711356 – 2007-2010)

Recent development of high throughput data acquisition technologies in a number of domains (e.g., biological sciences, atmospheric sciences, commerce) together with advances in digital storage, computing, and communications technologies have resulted in the proliferation of a multitude of physically distributed data repositories created and maintained by autonomous entities (e.g., scientists, organizations). The resulting increasingly data rich domains offer unprecedented opportunities in knowledge acquisition (e.g., discovery of a priori unknown complex relationships, construction of predictive models) from data. However, realizing these opportunities presents several challenges in practice: Data repositories are autonomously designed and operated, large in size, physically distributed, and differ in structure, organization, semantics, and query and processing capabilities. Our research, aimed at addressing some of these challenges, has led to:

(a) The development of a general theoretical framework for learning predictive models (e.g., classifiers) from large, physically distributed data sources where it is neither desirable nor feasible to gather all of the data in a centralized location for analysis. This framework [Caragea et al., 2001; 2003; 2004a] offers a general recipe for the design of algorithms for learning from distributed data that are *provably exact* with respect to their centralized counterparts (in the sense that the model constructed from a collection of physically distributed data sets is provably identical to that obtained in the setting where the learning algorithm has access to the entire data set). A key feature of our approach is the clear separation of concerns between hypothesis construction and extraction and refinement of *sufficient statistics* needed by the learning algorithm from data which reduces the problem of learning from data to a problem of decomposing a query for sufficient statistics across multiple data sources and combining the answers returned by the data sources to obtain the answer for the original query. Our work has resulted in provably exact algorithms (relative to the centralized counterparts) for learning decision trees, neural networks, support vector machines and Bayesian networks from distributed data.

(b) The development of theoretically sound yet practical variants of a large class of algorithms [Caragea et al., 2001; 2003; 2004a] for learning predictive models (classifiers) from distributed data sources under a variety of assumptions (motivated by practical applications) concerning the nature of data fragmentation, and the query capabilities and operations permitted by the data sources (e.g., execution of user supplied procedures), and precise characterization of the complexity (computation, memory, and communication requirements) of the resulting algorithms relative to their centralized counterparts.

(c) The development of a theoretically sound approach to formulation and execution of statistical queries across semantically heterogeneous data sources [Caragea et al., 2004b; Caragea et al., 2005; Caragea et al., 2006; 2007a; 2007b; Bao et al., 2007d]. This work has shown how to use semantic correspondences and *mappings* specified by users from a set of terms and relationships among terms (user ontology) to terms and relations in data source specific ontologies to construct a sound procedure for answering queries for *sufficient statistics* needed for learning classifiers from semantically heterogeneous data. An important component of this work

has to do with the development of statistically sound approaches to handling data specified at different levels of abstraction across different data sources.

(d) The design and implementation of INDUS – a modular, extensible, open source software toolkit for data-driven knowledge acquisition from large, distributed, autonomous, semantically heterogeneous data sources.

Research in progress is aimed at:

(a) Extending algorithms that are being developed by our group as well as others for learning from multiple relational databases to work with semantically heterogeneous data sources, taking advantage of the capability of INDUS to view heterogeneous information sources as though they were a collection of relational databases.

(b) Extending the ontology-based approach to information integration to develop ontology-based frameworks for composition of autonomously developed components and services using emerging frameworks for data source (or more generally resource or service) description, registry services, that are being developed as part of the Semantic Web efforts.

(c) Development of tools for collaborative and modular ontology development, specification of semantic mappings between information sources, ontology merging, learning specific types of ontologies (e.g., attribute value taxonomies) from data.

(d) Extension of approaches used in INDUS to support user and context-specific information integration and knowledge acquisition in peer-to-peer environments and distributed sensor networks.

(e) Further development of the INDUS prototype into a platform to support exploratory data analysis and knowledge acquisition in representative problems in bioinformatics and computational biology e.g., data-driven construction of classifiers of protein function; and predictors of protein-protein interaction, comparative analysis of gene expression data; modeling of interaction networks and pathways.

(f) Applications of INDUS to data integration problems that arise in monitoring and control of complex engineered systems e.g., distributed power grid, distributed information and communication networks.

(g) Dissemination of INDUS and associated software to the broader scientific community.

Modular Ontologies for Semantic Web and e-Science Applications (funded in part by a grant (IIS 0639230) from the National Science Foundation)

The success of the world wide web can be attributed to the *network effect*: The absence of central control on content and organization of the web allows thousands of independent actors to contribute resources (web pages) that are interlinked to constitute the web. Recent efforts to extend the web into a *semantic web* are aimed at enriching the web with machine interpretable content and interoperable resources and services. Realizing the full potential of the semantic web requires the large-scale adoption and use of ontology based approaches to sharing of information and resources. In such a setting, instead of a single, centralized ontology, it is much more natural to have multiple distributed ontologies that cover different, perhaps partially overlapping, domains (e.g., biology, medicine, pharmacology). Such ontologies represent the *local* knowledge of the ontology designers, that is, knowledge that is applicable within a specific *context*. Hence, there is an urgent need for theoretically sound yet practical approaches that support user, context, or application-specific adaptation and reuse of knowledge from multiple autonomously developed ontologies in specific applications. Ontologies on the semantic web need to satisfy apparently conflicting objectives: Selective *sharing* or *r*euse of knowledge across autonomously developed ontologies on the one hand and accommodation of the *local points of view* or *contextuality* of knowledge on the other. Our research on modular ontologies has led to:

- The development of modular variants of description logics that provide mechanisms for *semantic* importing of names (including concept, role and nominal names) across ontology modules and *contextualized interpretation* of reused knowledge [Bao, Caragea, and Honavar, 2006a; 2006b; Bao, Slutzki, and Honavar, 2007]. The resulting family of ontology languages – package-based description logics (P-DL), frees the ontology designer from the burden of ensuring the reusability of an ontology module in contexts that are hard to foresee at the time of construction of the module in question. A

natural consequence of contextualized interpretation is that inferences that are drawn are always *from the point of view* of a *witness* module. Thus, different modules might infer different consequences, based on the knowledge that they import from other modules.

- The development of distributed tableaux-based reasoning algorithms for P-DL [Bao, Caragea, and Honavar, 2006c; 2007b]
- The development of a framework for strategies for protecting private knowledge while truthfully answering queries [Bao, Slutzki, and Honavar, 2007c]
- Implementation of a suite of tools for collaborative development and use of package-based partial order ontologies for applications in biomedical informatics and comparative genomics

A long-term goal of our research on modular ontologies is to transform distributed data and knowledge base applications in the same way that the World-Wide Web has transformed the construction, sharing and use of hyperlinked documents and Wiki has transformed encyclopedia construction.

Algorithms and Software for Interactive Discovery and Composition of Semantic Web Services (Funded in part by a grant (IIS 0702758) from the National Science Foundation).

Recent advances in networks, information and computation grids, and WWW have resulted in the proliferation of a multitude of physically distributed and autonomously developed software components and services in various domains including e-Business and e-Science. Real world applications in these domains call for effective tools for developing composite services using available sets of component services. Existing approaches to web service composition suffer from a very significant limitation in that they require the user (or service developer) to provide a specification of the desired behavior of the composite service (goal) in its entirety. More importantly, the current approaches adopt a single-step request-response paradigm to service composition. That is, if a specified goal service is unrealizable (which would be the case if the goal service specification is incomplete), the process simply fails. It is typically difficult for a developer to provide the complete goal service specification that is needed in the absence of a detailed knowledge of the specifications of the component services available. This argues for an iterative approach to service composition wherein an abstract (and perhaps incomplete) goal service specification can be iteratively reformulated (with guidance from the system) until a composition that realizes the desired goal functionality is found, or the user decides to abort. To address this need, we have introduced a framework for Modeling Service Composition and Execution (MoSCoE) [Pathak et al., 2006a, 2006b]. MoSCoE, given an *abstract* (high-level and possibly incomplete) logical specifications of a goal service T_g, and of available component services T_1 ... T_n, identifies a subset of the component services that when *composed* with a choreographer *C*, realize the goal service T_g. A unique feature of MoSCoE is its ability, in the event of failure to realize a goal service, to identify the specific states and transitions of the goal STS that need to be modified. This information enables the user to *reformulate* the goal specification (iteratively) until a composition that realizes the goal specification is found or the user decides to abort.

Work in progress is aimed at extensions of MoSCoE to exploit relationships among services and among failure causes to reduce the computational complexity of interactive service composition, the use of ontologies and inter-ontology mappings to facilitate composition of composite services from semantically heterogeneous component services, development of efficient approaches to utilizing both functional and non-functional requirements to effectively search for compositions that meet user needs, and systematic evaluation of the performance of MoSCoE on a wide range of benchmark service composition tasks.

Data-Driven Discovery of Macromolecular Sequence-Structure-Function-Interaction-Expression Relationships (in collaboration with Drena Dobbs, Robert Jernigan, Amy Andreotti, funded in part by a National Institutes of Health Grant 5R21GM066387)

Proteins are the principal catalytic agents, structural elements, signal transmitters, transporters and molecular machines in cells. Hence, assigning them putative *functions* from sequences alone remains one of the most challenging problems in functional genomics. Protein function can be understood from multiple perspectives. Although proteins are most commonly classified in terms of their *biochemical* or *enzymatic function*, from

another perspective, their associations with one or more specific small molecules, cofactors, substrates, regulators, and other proteins are also important for functional assemblages. A protein's *cellular functional network* is determined by its interactions with other molecules; the function and the location of such complexes in a cell or in relation to a metabolic or signaling pathway are critical to its functioning. A protein's role at the organismal level (the phenotypic effect of deleting or mutating the corresponding gene) provides another way to define protein function. And, many proteins are multi-functional, and multiple proteins may cooperate to achieve a single function. Improvements in annotating protein sequences can be expected to yield significant improvements in gene annotations.

Protein-protein, protein-DNA, and protein-protein interactions play a pivotal role in protein function. Experimental detection of residues in protein-protein interaction surfaces must come from determination of the structure of protein-protein, protein-DNA and protein-RNA complexes. However, experimental determination of such complexes lags far behind the number of known protein sequences. Hence, there is a need for development of reliable computational methods for identifying protein-protein interface residues. Identification of protein-protein interaction sites and detection of specific amino acid residues that contribute to the specificity and strength of such interactions is an important problem with broad applications ranging from rational drug design to the analysis of metabolic and signal transduction networks.

Against this background, this project aims to develop and systematically evaluate computational methods for discovering sequence and structural correlates of protein function by analyzing large data sets derived from multiple information sources (e.g., protein sequences, protein structures, protein-protein interaction data, gene expression data), from multiple perspectives, based on different views of structure and function.

Specific aims of this research are:

(a) To develop, implement, and evaluate novel data mining algorithms for assigning proteins to structural and functional families that address specific limitations of existing data mining algorithms for computational characterization of protein sequence-structure-function relationships, including in particular, algorithms for exploiting prior knowledge (e.g., hierarchical taxonomies of attributes) and for multi-label classification tasks (e.g., for assigning a protein to several, not necessarily disjoint classes of function (e.g., based on existing models of biological function such as those captured by the Gene Ontology (GO) classifications); probabilistic graphical models and probabilistic language models for sequence classification; and efficient algorithms for learning from multiple tables which provide a natural way to incorporate information (data and knowledge) from multiple sources in analysis of protein structure and function from multiple perspectives.

(b) To develop, implement, and systematically evaluate data mining approaches for characterization and prediction of protein-protein, protein-DNA, and protein-RNA interaction residues primarily from protein sequence data (but utilizing other sources of data when available – including predicted or known structures of the protein *but not the complex*, evolutionary profiles, etc.) As part of this effort, we are assembling comprehensive databases of protein-protein, protein-RNA, and protein-DNA interfaces, and developing software tools for identification of protein-protein, protein-RNA and protein-DNA interface residues.

(c) Experimental validation of predicted protein-protein and protein-RNA interactions through a number of collaborative projects e.g., understanding the significance of proline-mediated *cis-trans* conformational changes in regulating protein function. Our hypothesis is that the "proline switch" may be widely exploited to regulate protein-protein and protein-ligand interactions by controlling the transition between two alternative binding surfaces.

Some results to date include a novel two-stage approach to identification of protein-protein (Yan, Dobbs and Honavar, 2004; Sen et al., 2004), protein-DNA (Yan et al., 2006), protein-RNA (Terribilini et al., 2006; 2007) interface residues from amino acid sequence (and when available protein structure), construction and analysis of a comprehensive database of protein-protein interfaces (Yan et al., 2007; Wu et al., 2007); and new probabilistic graphical models and related methods for assigning protein sequences to functional families (Andorf, Silvescu, and Honavar, 2004; Andorf, Dobbs and Honavar, 2007), several successful applications of machine learning

approaches to identification of functionally important sites (e.g., phosphorylation sites, glycosylation sites (Caragea et al., 2007), MHC-binding sites (El-Manzalawi et al., 2007) of proteins from amino acid sequences.

Data Driven Discovery and Refinement of Genetic Signaling Pathways and Regulatory Networks (in collaboration with Heather Greenlee, and Chris Tuggle funded in part by a National Institutes of Health Grant 5R03EY014931, and by a collaborative USDA grant)

Retinal dystrophies (diseases that involve degeneration of photoreceptors in the retina) are a major cause of blindness. Age related macular degeneration (AMD), which results in degeneration of cone photoreceptors dense macula, affects 10 million people in the United States alone. AMD is an extremely debilitating disease, which robs affected individuals of their high acuity vision. Our research is aimed at understanding the mechanisms that control retinal development and differentiation. Such understanding can lead to the development of therapies involving rehabilitation of retinal cells that can be rescued from degeneration by the application of exogenous survival factors and transplantation of retinal progenitor/stem cells into the degenerate retina.

The widely accepted model of retinal progenitor differentiation asserts that retinal progenitors pass through successive intrinsic "competency states" in which they are capable of responding to extrinsic cues. The extrinsic signal can encompass a whole host of soluble and cell-cell mediated factors. Intrinsic competence to respond to such external factors involves expression of a number of transcription factors as well as expression and plasma membrane localization of appropriate cell surface receptors that make it possible for the progenitors to respond to these cues from the extra-cellular environment. Because the ability of a progenitor cell to respond to cues from its environment is determined to a large extent by the proteins present in its proteome, we are investigating changes in expression among plasma membrane proteins of the retinal progenitor cells from several different ages. This provides us with data on the changing levels of protein expression in the developing retina which can be used to discover coordinated patterns of gene expression changes.

Specific computational aims of this research include development and applications of novel computational approaches to: data-driven inference genetic regulatory networks, and signaling pathways including:

(a) Discovery of co-expressed or co-regulated genes from gene expression patterns
(b) Construction and data-driven refinement of genetic networks from gene expression data
(c) Comparative analysis of gene expression across multiple platforms, multiple tissues, and multiple species using graph-theoretic algorithms
(d) Identification of functional modules from gene expression data
(e) Modeling and analysis of changes in gene expression patterns and the underlying genetic regulatory networks and signaling pathways that control retinal development and differentiation
(f) Development of the Retina Workbench for interactive analysis, visualization, and querying of gene expression (and more generally, interaction) data from multiple sources

This work builds on our recent results on analysis of gene expression in chloroplast biogenesis in maize (Lonosky et al, 2004) and inference of temporal boolean network models of genetic networks from gene expression data (Silvescu and Honavar, 2001) and comparative analysis of mRNA and protein expression data from the developing retina (Barnhill et al., 2007). We have developed a novel approach to prioritizing experimental targets based on querying, analysis, and interpretation of gene expression data in the context of other types of biological information (Hecker et al., 2007).

Topics in Grammatical Inference, Generative Models, and Computational Learning Theory

Grammatical Inference, variously referred to as automata induction, grammar induction, and automatic language acquisition, refers to the process of learning of grammars and languages from data. Machine learning of grammars finds a variety of applications in syntactic pattern recognition, adaptive intelligent agents, diagnosis, computational biology, systems modeling, prediction, natural language acquisition, data mining and knowledge discovery. Regular grammars are the simplest class of formal grammars in the Chomsky hierarchy. An

understanding of the issues and problems encountered in learning regular languages (or equivalently, identification of the corresponding deterministic finite automaton (DFA)) are therefore likely to provide insights into the problem of learning more general classes of languages. Under the standard complexity theoretic assumption P ≠NP, Pitt and Warmuth showed that no polynomial time algorithm can be guaranteed to produce a DFA that has approximately the same number of states as the target DFA from a set of labeled examples corresponding to a DFA. When examples are drawn at random (as in the PAC setting), results proved by Kearns and Valiant suggest that an efficient algorithm for learning DFA would entail efficient algorithms for solving problems such as breaking the RSA cryptosystem, factoring Blum integers, and detecting quadratic residues, all of which are known to be hard under standard cryptographic assumptions.

Against the background of strong negative results we investigated the feasibility of learning regular languages from examples under additional assumptions concerning the distribution from which the examples are drawn, thereby addressing the problem posed by Pitt, in his seminal paper: Are DFA PAC-identifiable if examples are drawn from the uniform distribution, or some other known simple distribution? We showed that:

(a) The class of simple DFA (i.e., DFA whose canonical representations have logarithmic Kolmogorov complexity) is efficiently PAC learnable under the Solomonoff Levin universal distribution (Parekh and Honavar, 1999)

(b) If the examples are sampled at random according to the universal distribution by a teacher that is knowledgeable about the target concept, the entire class of DFA is efficiently PAC learnable under the universal distribution, that is, DFA are efficiently learnable under the PACS Model (Parekh and Honavar, 1999; Parekh and Honavar, 2001)

(c) Any concept that is learnable under Gold's model for learning from characteristic samples, Goldman and Mathias' polynomial teachability model, and the model for learning from example based queries is also learnable under the PACS model (Parekh and Honavar, 2000; 2001).

Work in progress is aimed at extending the theoretical foundations and algorithms for grammatical inference to settings that require stochastic models to handle noise or other sources of uncertainty in the generative models as well as settings which call for learning from multimodal data (e.g., combination of words and pictures).

PUBLICATIONS AND PRESENTATIONS

Books Authored and Edited

1. Honavar, V., and Caragea, D. (2007) Collaborative **Knowledge Acquisition from Semantically Disparate, Distributed, Autonomous Data Sources**. To appear.
2. Patel, M., Honavar, V. & Balakrishnan, K. (Ed.) (2001). **Advances in Evolutionary Synthesis of Intelligent Agents**. Cambridge, MA: MIT Press.
3. Honavar, V. & Uhr, L. (1994) (Ed). **Artificial Intelligence and Neural Networks***:* Steps Toward Principled Integration. New York, NY: Academic Press.

Conference Proceedings Edited

1. Honavar, V. & Slutzki, G. (Ed.) (1998). **Grammatical Inference** Vol. 1433. Lecture Notes in Computer Science. Berlin: Springer-Verlag.
2. Banzaf, W., Daida, J., Eiben, A. Garzon, M., Honavar, V., Jakiela, M., & Smith, R. (Ed.) (1999). Proceedings of the Genetic and Evolutionary Computation Conference. San Mateo, CA: Morgan Kaufmann.
3. W. Langdon, E. Cantu-Paz, K. Mathias, R. Roy, D. Davis, R. Poli, K. Balakrishnan, V. Honavar, G. Rudolph, J. Wegener, L. Bull, M. Potter, A. Schultz, J. Miller, E. Burke, N. Jonoska. (2002). (Ed). Proceedings of the Genetic and Evolutionary Computing Conference. Palo Alto, CA: Morgan Kaufmann.

4. H. J. Caulfield, S.-H. Chen, H.-D. Cheng, R. Duro, V. Honavar, E. E. Kerre, M. Lu, M. G. (2002). Romay, T. K. Shih, D. Ventura, P. P. Wang, and Y. Yang, editors, 2002. Proceedings 6th Joint Conference on Information Sciences, JCIS / Association for Intelligent Machinery.

Refereed Journal and Conference Papers

Note: Acceptance of papers for publication in the proceedings of top Computer Science conferences is typically based on rigorous peer review. Acceptance rates are typically under 30% and conference proceedings are published and distributed by major commercial publishers (e.g., Springer-Verlag) or professional societies (e.g., Association for Computing Machinery). The list includes several papers that are "highly cited", "most viewed" or recipients of "best paper awards." According to Google scholar, the list includes: 1 article (Yang & Honavar, 1998) with over 300 citations, 1 with over 100 citations, 5 articles with between 50 and 99 citations each, 5 articiles with between 30 and 49 citations each, 17 articles with between 20 and 29 citations each and 27 articles with between 10 and 19 citations each.

1. Lee, J-H., Hamilton, M., Gleeson, C., Caragea, C., Zaback, P., Sander, J.D., Li, X., Wu, F., Terribilini, M., Honavar, V., and Dobbs, D. Striking Similarities in Diverse Telomerase Proteins Revealed by Combining Structure Prediction and Machine Learning Approaches. *Pacific Symposium on Bioinformatics*, In press, 2008.
2. Andorf, C., Dobbs, D., and Honavar, V. (2007). Exploring Inconsistencies in Large-Scale Protein Function Annotations: A Machine Learning Approach. BMC Bioinformatics. http://www.biomedcentral.com/1471-2105/8/284/
3. Andorf, C., Silvescu, A., Dobbs, D., and Honavar, V. (2007) Subcellular Localization Prediction using a Hybrid Approach Using Machine Learning and BLAST. Under review.
4. Andorf, C., Dobbs, D., and Honavar, V. (2007) Reduced Alphabet Representations of Amino Acid Sequence for Protein Function Classification. International Journal of Data Mining and Bioinformatics. Under review.
5. Bao, J., Slutzki, G., and Honavar, V. (2007). A Semantic Importing Approach to Reusing Knowledge from Multiple Autonomous Ontology Modules. In: Proceedings of the 22nd Conference on Artificial Intelligence (AAAI-2007). Vancouver, Canada. AAAI Press. pp. 1304-1309
6. Bao, J., Slutzki, G. and Honavar, V. (2007). Privacy-Preserving Reasoning on the Semantic Web. Proceedings of the *ACM/WIC/IEEE International Conference on Web Intelligence*, IEEE Press, In press.
7. Bao, J., Slutzki, G., and Honavar, V. Distributed Reasoning with Modular ALC Ontologies (ALCP). *Web Intelligence and Agent Systems*. In press, 2007.
8. Bao, J., Caragea, D., and Honavar, V (2007). Ontology-Extended Data Sources. International Journal of Semantic Web and Information Systems. Under review.
9. Barnhill, A., Hecker, L., Kohuytuk, O., Alcon, T., Buss, J., Honavar, V., Greenlee, H. (2007). Characterization of the Retinal Proteome During Rod Photoreceptor Genesis. Investigative Ophthalmology and Visual Science. Under revision.
10. Caragea, C., Sinapov, J., Silvescu, A., Dobbs, D. And Honavar, V. (2007). Glycosylation Site Prediction Using Ensembles of Support Vector Machine Classifiers. *BMC Bioinformatics*. In press.
11. Caragea, C., Sinapov, J., Terribilini, M., Dobbs, D. and Honavar, V. (2007) Assessing the Performance of Macromolecular Sequence Classifiers. *Proceedings of the IEEE Conference on Bioinformatics and Bioengineering*, (IEEE BIBE 2007). IEEE Press. pp. 320-326.
12. Dunn-Thomas, T., Dobbs, D.L., Sakaguchi, D. Young, M.J. Honavar, V. Greenlee, H. M. W. (2007). Proteomic Differentiation Between Murine Retinal and Brain Derived Progenitor Cells. *Stem Cells and Development*. In press.
13. El-Manzalawi, Y., Dobbs, D., and Honavar, V. (2007). PepMIL: A Novel Method for Predicting Flexible Length MHC-II binding Peptides. Under review.

14. Hecker, L., Alcon, T., Honavar, V., Greenlee, H. (2007). Querying multiple large-scale gene expression datasets from the developing retina using a seed network to prioritize experimental targets. *Bioinformatics and Biology Insights*. In press.

15. Helmer, G., Wong, J., Slagell, M., Honavar, V., Miller, L., Wang, Y., Wang, X., and Stakhanova, N. (2007). Software Fault Tree and Colored Petri Net Based Specification, Design, and Implementation of Agent-Based Intrusion Detection Systems. *International Journal of Information and Computer Security*. Vol. 1. No. 1/2. pp. 109-142.

16. J. McCalley, V. Honavar, S. Ryan, W. Meeker, D. Qiao, R. Roberts, Y. Li, J. Pathak, M. Ye, Y. Hong (2007). Integrated Decision Algorithms for Auto-steered Electric Transmission System Asset Management. 7th Intl. Conference on Computational Science, Lecture notes in Computer Science. Berlin: Springer-Verlag. Vol. 4487. pp. 1066-1073.

17. Pathak, J., Basu, S., and Honavar, V. (2007). On Context-Specific Substitutability of Web Services. Proceedings of the IEEE International Conference on Web Services, IEEE. 2007.

18. Pathak, J., Basu, S., Lutz, R., and Honavar, V. (2007). MoSCoE: An Approach for Composing Web Services through Iterative Reformulation of Functional Specifications. *International Journal on Artificial Intelligence Tools*. In press.

19. Terribilini, M., Sander, J.D., Lee, J-H., Zaback, P., Jernigan, R.L., Honavar, V. and Dobbs, D. (2007). RNABindR: A Server for Analyzing and Predicting RNA Binding Sites in Proteins. Nucleic Acids Research. Vol. 35. No. 9. pp. doi:10.1093/nar/gkm2, 2007

20. Wu, F., Towfic, F., Dobbs, D. and Honavar, V. (2007) Analysis of Protein Protein Dimeric Interfaces. IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2007). *San Jose*, In press, 2007.

21. Yan, C., Wu, F., Jernigan, R.L., Dobbs, D., Honavar, V. (2007) Analysis of Protein-Protein Interfaces. The Protein Journal. 10.1007/s10930-007-9108-x

22. Yan, C., Lee, J.H., Jernigan, R., Honavar, V. and Dobbs, D. (2007). Identification of DNA binding residues in the S subunit of the *M. jannaschii* type I restriction-modification system. Under review.

23. Bao, J., Caragea, D., and Honavar, V. (2006). A Tableau-based Federated Reasoning Algorithm for Modular Ontologies. ACM / IEEE / WIC Conference on Web Intelligence, Hong Kong, pp. 404-410.

24. Bao, J., Caragea, D., and Honavar, V. (2006). On the Semantics of Linking and Importing in Modular Ontologies. International Semantic Web Conference, Athens, Georgia, USA, Springer-Verlag Lecture Notes in Computer Science, Vol. 4273, pp. 72-86.

25. A Bao, J., Caragea, D., and Honavar, V. (2006). A Distributed Tableau Algorithm for Package-based Description Logics. Proceedings of the Second International Workshop on Context Representation and Reasoning (CRR 2006), Riva del Garda, Italy, CEUR. 2006.

26. Bao. J. and Honavar, V. (2006) Divide and Conquer Semantic Web with Modualr Ontologies - A Brief Review of Modular Ontology Language Proposals. First International Workshop on Modular Ontologies (WoMo2006). International Semantic Web Conference, Athens, GA, 2006.

27. Bao. J. and Honavar, V. (2006) Adapt OWL as a Modular Ontology Language (Position Paper). Accepted by OWL: Experiences and Directions (OWLED 2006). International Semantic Web Conference, Athens, GA. CEUR Report. Vol. 216.

28. Bao. J. and Honavar, V. Representing and Reasoning with Modular Ontologies. In AAAI 2006 Fall Symposium on Semantic Web for Collaborative Knowledge Acquisition (SWeCKa 2006), Arlington, VA, USA, October 2006.

29. Bao. J., Caragea, D., and Honavar, V. Package-based Description Logics - Preliminary Results. International Semantic Web Conference (ISWC) 2006. Springer-Verlag Lecture Notes in Computer Science, Vol. 4273, pp. 967-969.

30. Bao, J., Caragea, D., and Honavar, V. (2006). Towards Collaborative Environments for Ontology Construction and Sharing. Proceedings of the International Symposium on Collaborative Technologies and Systems., Las Vegas, IEEE press. pp. 99-108.

31. Bao, J., Hu, Z., Caragea, D., Reecy, J., and Honavar, V. (2006). A Tool for Collaborative Construction of Large Biological Ontologies. Fourth International Workshop on Biological Data Management (BIDM 2006), Krakov, Poland, IEEE Press. pp. 191-195.

32. Bao, J., Caragea, D., and Honavar, V. (2006). Modular Ontologies - A Formal Investigation of Semantics and Expressivity. In Proceedings of the First Asian Semantic Web Conference, Beijing, China, Springer-Verlag Lecture Notes in Computer Science, Vol. 4185, pp. 616-631. **Best paper award**.

33. Bromberg, F., Margaritis, D., and Honavar, V. (2006). Efficient Markov Network Structure Discovery from Independence Tests. In: SIAM Conference on Data Mining (SDM 06). SIAM Press. pp. 141-152.

34. Caragea, D., Zhang, J., Pathak, J., and Honavar, V. Learning Classifiers from Distributed, Ontology-Extended Data Sources. Proceedings of the 8th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2006), Krakov, Poland, Lecture Notes in Computer Science. Berlin: Springer. Vol. 4081. pp. 363-373.

35. Caragea, D. and Honavar, V. (2006). A General Strategy for Learning Classifiers from Semantically Heterogeneous, Distributed Data. In AAAI 2006 Fall Symposium on Semantic Web for Collaborative Knowledge Acquisition (SWeCKa 2006), Arlington, VA, USA, AAAI. pp. 1-8.

36. Kang, D-K., Silvescu, A. and Honavar, V. (2006). Recursive Naive Bayes Learner for Sequence Classification. In: Proceedings of the Tenth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2006). Lecture Notes in Computer Science. Vol. 3918, Berlin: Springer-Verlag. pp. 45-54

37. Lee, K., Joo, J., Yang, J., and Honavar, V. (2006). Proceedings of the International Conference on Advanced Data Mining and Applications, Second International Conference (ADMA 2006), Springer-Verlag Lecture Notes in Computer Science. Vol. 4093. pp. 465-472.

38. McCalley, J.D., Honavar, V., Ryan, S.M., Meeker, W.Q., Roberts, R.A., Qiao, D., and Li, Y.: Auto-steered Information-Decision Processes for Electric System Asset Management. In: Proceedings of the International Conference on Computational Science, pp. 440-447

39. Pathak, J., Basu, S., Lutz, R., and Honavar, V. (2006). Selecting and Composing Web Services Through Iterative Reformulation of Functional Specifications. Proceedings of the IEEE International Conference on Tools With Artificial Intelligence (ICTAI 2006), Washington, DC, IEEE. **Best Paper Award.**

40. Pathak, J., Basu, S., Lutz, R., and Honavar, V. (2006). Parallel Web Service Composition in MoSCoE: A Choreography Based Approach. Proceedings of the IEEE European Conference on Web Services (ECOWS 2006), Zurich, Switzerland, IEEE. pp. 3-12.

41. Pathak, J., Basu, S., and Honavar, V. (2006). Modeling Web Services by Iterative Reformulation of Functional and Non-Functional Requirements. Proceedings of the International Conference on Service Oriented Computing. Springer-Verlag Lecture Notes in Computer Science. Vol. 4294. pp. 314-326.

42. Pathak, J. Jiang, Y., Honavar, V., McCalley, J. (2006). Condition Data Aggregation for Failure Mode Estimation of Power Transformers. In IEEE 39th Hawaii International Conference on System Sciences (HICSS-2006), IEEE press, Kauai, Hawaii, USA.

43. Silvescu, A. and Honavar, V. (2006) Independence, Decomposability and functions which take values into an Abelian Group. Proceedings of the Ninth International Symposium on Artificial Intelligence and Mathematics, 2006.

44. Terribilini, M., Lee, J-H., Yan, C., Honavar, V. and Dobbs, D. (2006). Identifying Interactions in "Recalcitrant" Proteins. Predicted Protein and RNA binding sites in REV proteins of HIV and EIAV agree with experimental data. In: Proceedings of the Pacific Symposium on Biocomputing, Hawaii, World Scientific. Vol. 11. pp. 415-426.

45. Vasile, F., Silvescu, A., Kang, D-K., and Honavar, V. (2006). TRIPPER: An Attribute Value Taxonomy Guided Rule Learner. In: Proceedings of the Tenth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD). Lecture Notes in Computer Science. Vol. 3918, Berlin: Springer-Verlag. pp. 55-59.

46. Wu, F., Olson, B., Dobbs, D., and Honavar, V. (2006). Using Kernel Methods to Predict Protein-Protein Interaction Sites from Sequence. IEEE Joint Conference on Neural Networks, Vancouver, Canada, IEEE Press. 2006.

47. Terribilini, M., Lee, J.H., Yan, C., Jernigan, R., Honavar, V., and Dobbs, D. (2006). Computational Prediction of Protein-RNA Interfaces. RNA Journal.. Vol. 12. No. 1450. pp. 1462, 2006.

48. Yan, C., Terribilini, M., Wu, F., Jernigan, R.L., Dobbs, D. and Honavar, V. (2006). Identifying amino acid residues involved in protein-DNA interactions from sequence. BMC Bioinformatics. Vol. 7. pp. 262-, 2006. **Highly accessed.**

49. Wang, Y., Behera, S., Wong, J., Helmer, G., Honavar, V., Miller, L. and Lutz, R. (2006). Towards the automatic generation of mobile agents for distributed intrusion detection systems. Journal of Systems and Software. Vol. 79. pp. 1-14.

50. Zhang, J., Silvescu, A., Kang, D-K., and Honavar, V. (2006). Learning Compact and Accurate Naive Bayes Classifiers from Attribute Value Taxonomies and Partially Specified Data. Knowledge and Information Systems. Vol. 9. No. 2. pp. 157-179, 2006.

51. Caragea, D., Zhang, J., Bao, J., Pathak, J., and Honavar, V. (2005). Algorithms and Software for Collaborative Discovery from Autonomous, Semantically Heterogeneous, Distributed, Information Sources. Invited paper. In: Proceedings of the Conference on Algorithmic Learning Theory. Lecture Notes in Computer Science. Vol. 3734. Berlin: Springer-Verlag. pp. 13-44.

52. Caragea, D., Silvescu, A., Pathak, J., Bao, J., Andorf, C., Dobbs, D., and Honavar, V. (2005) Information Integration and Knowledge Acquisition from Semantically Heterogeneous Biological Data Sources. Data Integration in Life Sciences (DILS 2005), San Diego, Berlin: Lecture Notes in Computer Science. Berlin: Springer-Verlag. Vol. 3615 pp. 175-190.

53. Caragea, D., Bao, J., Pathak, J., Andorf, C, Dobbs, D., and Honavar, V. (2005). Information Integration from Semantically Heterogeneous Biological Data Sources. Proceedings of the Third International Workshop on Biological Data Management (BIDM 2005), 16th International Workshop on Database and Expert Systems Applications (DEXA 2005), Copenhagen, Denmark. IEEE Computer Society 2005. pp. 580-584

54. Kang, D-K., Silvescu, A., and Honavar, V. (2005). Multinomial Event Model Based Abstraction for Sequence and Text Classification. Proceedings of the Symposium on Abstraction, Reformulation, and Approximation (SARA 2005), Edinburgh, UK, Lecture Notes in Computer Science, Vol. 3607. Berlin: Springer-Verlag, pp. 134-148

55. Kang, D-K., Fuller, D., and Honavar, V. (2005). Learning Classifiers for Misuse and Anomaly Detection Using a Bag of System Calls Representation. Proceedings of the 6th IEEE Systems, Man, and Cybernetics Workshop (IAW 05), West Point, NY, IEEE Press, pp. 118-125

56. Kang, D-K., Fuller, D., and Honavar, V. Learning Misuse and Anomaly Detectors from System Call Frequency Vector Representation IEEE International Conference on Intelligence and Security Informatics, 2005. Lecture Notes in Computer Science, Vol. 3495. Berlin: Springer-Verlag. pp. 511-516.

57. Pathak, J,, Koul, N., Caragea, D., and Honavar, V. (2005). A Framework for Semantic Web Services Discovery. In: Proceedings of the 7th ACM International Workshop on Web Information and Data Management (WIDM 2005). ACM Press. pp. 45-50.

58. Wu, F., Zhang, J., and Honavar, V. (2005). Learning Classifiers Using Hierarchically Structured Class Taxonomies. In: Proceedings of the Symposium on Abstraction, Reformulation, and Approximation (SARA 2005), Edinburgh, UK, Lecture Notes in Computer Science, Vol. 3607. Berlin: Springer-Verlag, pp. 313-320.

59. Yakhnenko, O., Silvescu, A., and Honavar, V. (2005). Discriminatively Trained Markov Model for Sequence Classification. In: Proceedings of the IEEE Conference on Data Mining (ICDM 2005). IEEE Press. pp. 498-505.

60. Zhang, J., Caragea, D. and Honavar, V. (2005). Learning Ontology-Aware Classifiers. In: Proceedings of the 8th International Conference on Discovery Science. Springer-Verlag Lecture Notes in Computer Science. Singapore. Vol. 3735. pp. 308-321. Berlin: Springer-Verlag.

61. Andorf, C., Silvescu, A., Dobbs, D. and Honavar, V. Learning Classifiers for Assigning Protein Sequences to Gene Ontology Functional Families. In: Proceedings of the Fifth International Conference on Knowledge Based Computer Systems (KBCS 2004), India.

62. Bao, J., Cao, Y., Tavanapong, W., and Honavar, V. (2004). Integration of Domain-Specific and Domain-Independent Ontologies for Colonoscopy Video Database Annotation. In: International Conference on Information and Knowledge Engineering (IKE 04).

63. Caragea, D., Pathak, J., and Honavar, V. (2004). Learning Classifiers from Semantically Heterogeneous Data. In: Proceedings of the International Conference on Ontologies, Databases, and Applications of Semantics (ODBASE 2004), Agia Napa, Cyprus, 2004.

64. Caragea, D., Silvescu, A., and Honavar, V. (2004). A Framework for Learning from Distributed Data Using Sufficient Statistics and its Application to Learning Decision Trees. International Journal of Hybrid Intelligent Systems. **Invited** Paper. Vol 1. pp. 80-89, 2004

65. Cook, D., Caragea, D. and Honavar, V. (2004). Visualization in Classification Problems. In: Proceedings in Computational Statistics (COMPSTAT 2004). pp. 823-830.

66. Kang, D-K., Silvescu, A., Zhang, J., and Honavar, V. (2004). Generation of Attribute Value Taxonomies from Data for Data-Driven Construction of Accurate and Compact Classifiers. In: Proceedings of the IEEE International Conference on Data Mining.

67. D.-K. Kang, A. Silvescu, J. Zhang, V. Honavar: Generation of Attribute Value Taxonomies from Data and Their Use in Data-Driven Construction of Accurate and Compact Naive Bayes Classifiers. In: Proceedings of the ECML/PKDD Workshop on Knowledge Discovery and Ontologies (KDO 2004).

68. Lonosky, P., Zhang, X., Honavar, V., Dobbs, D., Fu, A., and Rodermel, S. (2004) A Proteomic Analysis of Chloroplast Biogenesis in Maize. Journal of Plant Physiology. Vol. 134, pp. 560-574, 2004.

69. Pathak, J., Caragea, D., and Honavar, V. (2004). Ontology-Extended Component-Based Workflows: A Framework for Constructing Complex Workflows from Semantically Heterogeneous Software Components. In: Proceedings of the Workshop on Semantic Web and Databases (SWDB 2004).

70. Polikar, R., Udpa, L., Udpa, S., and Honavar, V. (2004). An Incremental Learning Algorithm with Confidence Estimation for Automated Identification of NDE Signals. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control. 51: 990-1001.

71. Sen,T.Z., Kloczkowski, A., Jernigan, R.L., Yan, C., Honavar, V., Ho, K-M., Wang, C-Z., Ihm, Y., Cao, H., Gu, X., and Dobbs, D. (2004). Predicting Binding Sites of Protease-Inhibitor Complexes by Combining Multiple Methods. BMC Bioinformtics. 5:205- **Highly Accessed**.

72. Yan, C., Dobbs, D. and Honavar, V. (2004). Identifying Protein-Protein Interaction Sites from Surface Residues – A Support Vector Machine Approach. Neural Computing Applications.13: 123-129. 2004.

73. Yan, C., Dobbs, D., and Honavar, V. (2004). A Two-Stage Classifier for Identification of Protein-Protein Interface Residues. Bioinformatics. 20. pp. i371-378, 2004. (also In: Proceedings of the Conference on Intelligent Systems in Molecular Biology (ISMB) 2004.)

74. Zhang, J., and Honavar, V. (2004). Learning Naïve Bayes Classifiers from Attribute Value Taxonomies and Partially Specified Data. In: Proceedings of the Conference on Intelligent Systems Design and Applications (ISDA 2004).

75. Zhang, J. and Honavar, V. (2004). AVT-NBL - An Algorithm for Learning Compact and Accurate Naive Bayes Classifiers from Attribute Value Taxonomies and Data. In: Proceedings of the IEEE International Conference on Data Mining.

76. Atramentov, A., Leiva, H., and Honavar, V. (2003). A Multi-Relational Decision Tree Learning Algorithm - Implementation and Experiments. In: Proceedings of the Thirteenth International Conference on Inductive Logic Programming. (Vol. 2835, Lecture Notes in Computer Science) Berlin: Springer-Verlag. pp. 38-56.

77. Caragea, D., Cook. D., and Honavar, V. Towards Simple, Easy-to-Understand, yet Accurate Classifiers. In: Proceedings of the IEEE International Conference on Data Mining (ICDM 2003). Melbourne, Florida, 2003.

78. Caragea, D., Cook, D., and Honavar. V. Visualizing Ensemble of Hyperplane Classifiers. In: Proceedings of Visual Data Mining Workshop. IEEE International Conference on Data Mining (ICDM 2003). Melbourne, Florida, 2003. Springer Verlag.

79. Caragea, D., Silvescu, A., and Honavar, V. (2003) Decision Tree Induction from Distributed, Heterogeneous, Autonomous Data Sources. In: Proceedings of the Conference on Intelligent Systems Design and Applications (ISDA 2003). Springer Verlag.

80. Helmer, G., Wong, J., Honavar, V., and Miller, L. (2003). Lightweight Agents for Intrusion Detection. Journal of Systems and Software. Vol. 67, No. 2. pp. 109-122.

81. Reinoso-Castillo, J., Silvescu, A., Caragea, D., Pathak, J., and Honavar, V. (2003). An Ontology-Based Query-Centric Approach to Information Extraction and Integration from Heterogeneous, Distributed, Autonomous Data Sources. IEEE International Conference on Information Reuse and Integration.

82. Wang, X., Schroeder, D., Dobbs, D., and Honavar, V. (2003). Data-Driven Discovery of Rules for Protein Function Classification Based on Sequence Motifs. Information Sciences. Vol. 155. No. 1., pp. 1-18

83. Yan, C., Dobbs, D., and Honavar, V. (2003) Identification of Residues Involved in Protein-Protein Interaction from Amino Acid Sequence - A Support Vector Machine Approach. In: Proceedings of the Conference on Intelligent Systems Design and Applications (ISDA 2003). Springer-Verlag.

84. Zhang, J., and Honavar, V. (2003). Learning from Attribute-Value Taxonomies and Partially Specified Instances. In: Proceedings of the International Conference on Machine Learning. Washington, DC. AAAI Press. pp. 880-887.

85. Z. Zhong, V. Vishwanathan, J. McCalley, V. Honavar, Multiagent System Solutions for Distributed Computing, Communications, and Data Integration Needs in the Power Industry. In. Proc. of the 2003 IEEE PES Summer Meeting, Toronto, July, 2003.

86. Zhang, J., Silvescu, A., and Honavar, V. Ontology-Driven Induction of Decision Trees at Multiple Levels of Abstraction. In: Proceedings of the Symposium on Abstraction, Reformulation, and Approximation (SARA-2002). Kananaskis, Alberta, Canada. Lecture Notes in Computer Science. Berlin: Springer-Verlag. pp. 316-323

87. Andorf, C., Dobbs, D., and Honavar, V. (2002). Discovering Protein Function Classification Rules from Reduced Alphabet Representations of Protein Sequences. In: Proceedings of the Conference on Computational Biology and Genome Informatics. Durham, North Carolina.

88. Helmer, G., Wong, J., Slagell, M., Honavar, V., Miller, L., and Lutz, R. (2002) A Software Fault Tree Approach to Requirements Specification of an Intrusion Detection System. Requirements Engineering. Vol 7 (4) (2002) pp. 207-220.

89. Helmer, G., Wong, J., Honavar, V., and Miller, L. (2002). Automated Discovery of Concise Predictive Rules for Intrusion Detection. Journal of Systems and Software. Vol. 60. No. 3. pp. 165-175.

90. Wang, X., Schroeder, D., Dobbs, D., and Honavar, V. (2002). Data-Driven Discovery of Protein Function Classifiers: Decision Trees Based on MEME Motifs Outperform Those Based on PROSITE Patterns and Profiles on Peptidase Families.

91. Caragea, D., Cook, D., and Honavar, V. (2001). Gaining Insights into Support Vector Machine Classifiers Using Projection-Based Tour Methods. In: Proceedings of the Conference on Knowledge Discovery and Data Mining.

92. Helmer, G., Wong, J., Slagell, M., Honavar, V., Miller, L. and Lutz, R. (2001). A Software Fault Tree Approach to Requirements Analysis of an Intrusion Detection System. In: Proceedings of the Symposium on Requirements Engineering for Information Security, Indianapolis, IN, USA.

93. Mikler, A., Honavar, V. and Wong, J. (2001). Autonomous Agents for Coordinated Distributed Parameterized Heuristic Routing in Large Communication Networks. Journal of Systems and Software. Vol. 56. pp. 231-246.

94. Parekh, R. and Honavar, V. (2001). Learning DFA from Simple Examples. Machine Learning. Vol. 44. pp. 9-35.

95. Polikar, R., Udpa, L., Udpa, S., and Honavar, V. (2001) Learn++: An Incremental Learning Algorithm for Multi-Layer Perceptron Networks. IEEE Transactions on Systems, Man, and Cybernetics. Vol. 31. No. 4. pp. 491-508.

96. Polikar, R., Shinar, R., Honavar, V., Udpa, L., and Porter, M. (2001). Detection and Identification of Odorants Using an Electronic Nose. In: Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing.

97. Silvescu, A., and Honavar, V. (2001). Temporal Boolean Network Models of Genetic Networks and Their Inference from Gene Expression Time Series. Complex Systems. Vol. 13. pp. 54-75. (preliminary version in In: Proceedings of the Atlantic Symposium on Computational Biology, Genome Information Systems & Technology.)

98. Wang, D., Wang, X., Honavar, V., and Dobbs, D. (2001). Data-Driven Generation of Decision Trees for Motif-Based Assignment of Protein Sequences to Functional Families. In: Proceedings of the Atlantic Symposium on Computational Biology, Genome Information Systems & Technology.

99. Wong, J., Helmer, G., Naganathan, V. Polavarapu, S., Honavar, V., and Miller, L. (2001) SMART Mobile Agent Facility. Journal of Systems and Software. Vol. 56. pp. 9-22.

100. Balakrishnan, K., Bousquet, O. & Honavar, V. (2000). Spatial Learning and Localization in Animals: A Computational Model and its Applications for Mobile Robots. Adaptive Behavior. Vol. 7. no. 2. pp. 173-216.

101. Pai, P., Miller, L., Nilakanta, S., Honavar, V., and Wong, J. (2000). In: Challenges of Information Technology Management in the 21st Century. Proceedings of the Eleventh International Conference of the Information Resources Management Association, Anchorage, Alaska. Challenges of Information Technology Management in the 21st Century M. Khosrowpour (Ed.) pp. 325-329.

102. Parekh, R., Yang, J. & Honavar, V. (2000). Constructive Neural Network Learning Algorithms for Multi-Category Pattern Classification. IEEE Transactions on Neural Networks. Vol. 11. No. 2. pp. 436-451.

103. Parekh, R. and Honavar, V. (2000). On the Relationships between Models of Learning in Helpful Environments. In: Proceedings of the Fifth International Conference on Grammatical Inference. Lecture Notes in Artificial Intelligence Vol. 1891. Berlin: Springer-Verlag. pp. 207-220.

104. Polikar, R., Udpa, L., Udpa, S., and Honavar, V. (2000). Learn++: An Incremental Learning Algorithm for Multilayer Perceptron Networks. In: Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2000. Istanbul, Turkey.

105. Yang, J., Parekh, R. and Honavar, V. (2000). Comparison of Performance of Variants of Single-Layer Perceptron Algorithms on Non-Separable Data. Neural, Parallel, and Scientific Computations. Vol. 8. pp. 415-438.

106. Bhatt, R., Balakrishnan, K., and Honavar, V. (1999). A Hybrid Model for Rodent Spatial Learning and Localization. In: Proceedings of the International Joint Conference on Neural Networks. Washington, D.C.

107. Chen, C-H., & Honavar, V. (1999). A Neural Network Architecture for Syntax Analysis. IEEE Transactions on Neural Networks. Vol. 10, No. 1. pp. 94-114.

108. Dandu, R., L. Miller, S. Nilakanta, and V. Honavar. (1999). Populating a data warehouse with mobile agents. In: Proceedings of the The Tenth International Conference of the Information Resources Management Association. Hershey, PA. pp. 468-473.

109. Janakiraman, J. & Honavar, V. (1999). Adaptive Learning Rate for Speeding up Gradient Descent Learning. Microcomputer Applications. vol. 18. pp. 89-95.

110. Parekh, R. and Honavar, V. (1999). Simple DFA are Polynomially Probably Exactly Learnable from Simple Examples. In: Proceedings of the International Conference on Machine Learning. Bled, Slovenia. San Mateo, CA: Morgan Kaufmann. pp. 298-306.

111. Yang, J., Parekh, R., Honavar, V., and Dobbs, D. (1999). Data-Driven Theory Refinement Using KBDistAl. In: Proceedings of the Conference on Intelligent Data Analysis (IDA 99). Berlin: Springer-Verlag. pp. 331-342.

112. Yang, J., Parekh, R., Honavar, V., and Dobbs, D. (1999). Data-Driven Theory Refienement Algorithms for Bioinformatics. In: Proceedings of the International Joint Conference on Neural Networks. Washington, D.C. pp. 4064-4068.

113. Yang, J., Parekh, R., and Honavar, V. (1999). DistAl: An Inter-pattern Distance Based Constructive Learning Algorithm. Intelligent Data Analysis. vol. 3. pp. 55-73.

114. Balakrishnan, K., Bhatt, R., and Honavar, V. (1998). A Computational Model of Rodent Spatial Learning and Some Behavioral Experiments. In: Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society. Madison, WI. Mahwah, NJ: Lawrence Erlbaum. pp. 102-107.

115. Balakrishnan, K., Bhatt, R., and Honavar, V. (1998). Spatial Learning and Localization in Animals: A Computational Model and Behavioral Experiments. In: Proceedings of the Second European Conference on Cognitive Modelling. pp. 112-119. Nottingham University Press.

116. Balakrishnan, K. & Honavar, V. (1998). Intelligent Diagnosis Systems. Journal of Intelligent Systems. Vol 8. pp. 239-290.

117. Bousquet, O., Balakrishnan, K. & Honavar, V. (1998). Is the Hippocampus a Kalman Filter? In: Proceedings of the Pacific Symposium on Biocomputing. Singapore: World Scientific. pp. 655-666.

118. Helmer, G., Wong, J., Honavar, V., and Miller, L. (1998) Intelligent Agents for Intrusion Detection and Countermeasures. In: Proceedings of the IEEE Information Technology Conference. pp. 121-124.

119. Honavar, V., Miller, L. & Wong, J. (1998). Distributed Knowledge Networks: Design, Implementation and Applications. In: Proceedings of the IEEE Information Technology Conference. pp. 87-90.

120. Honavar, V., Miller, L., Yang, J. & Wong, J. (1998). Intelligent Mobile Agents for Information Retrieval and Knowledge Discovery from Distributed Data and Knowledge Sources. In: Proceedings of the IEEE Information Technology Conference. pp. 99-104.

121. Leavens, G., Baker, A., Honavar, V., Lavalle, S., and Prabhu, G. (1998). Programming is Writing: Why Student Programs Must be Carefully Evaluated. Mathematics and Computer Education. Vol. 32, pp. 284-295.

122. Miller, L., Honavar, V., and Wong, J. (1998). An Object-Oriented Data Warehouse for Seamless Information Access from Heterogeneous Distributed Data and Knowledge Sources. In: Proceedings of the IEEE Information Technology Conference. pp. 27-32.

123. Parekh, R. & Honavar, V. (1998). Constructive Theory Refinement in Knowledge-Based Neural Networks. In: Proceedings of the International Joint Conference on Neural Networks. Anchorage, Alaska. pp. 2318-2323.

124. Parekh, R., Nichitiu, C., & Honavar, V. (1998). A Polynomial Time Incremental Algorithm for Regular Grammar Inference. In: Proceedings of the Fourth International Colloquium on Grammatical Inference. Lecture Notes in Artificial Intelligence Vol. 1433. Berlin: Springer-Verlag. pp. 37-49.

125. Spartz, R. & Honavar, V. (1998). An Empirical Analysis of the Expected Source Values Rule. Microcomputer Applications Vol. 17, pp. 29-34.

126. Yang, J. & Honavar, V. (1998). DistAl: An Inter-Pattern Distance-Based Constructive Neural Network Learning Algorithm. In: In: Proceedings of the International Joint Conference on Neural Networks. Anchorage, Alaska. pp. 2208-2213.

127. Yang, J., Pai, P., Miller, L., & Honavar, V. (1998). Mobile Intelligent Agents for Document Classification and Retrieval: A Machine Learning Approach. In: Proceedings of the European Symposium on Cybernetics and Systems Research. Berlin: Springer-Verlag. pp. 707-712.

128. Yang, J., Havaldar, R., Honavar, V., Miller, L., & Wong, J. (1998). Coordination of Distributed Knowledge Networks Using the Contract Net Protocol. In: Proceedings of the IEEE Information Technology Conference. pp. 71-74.

129. Yang, J. & Honavar, V. (1998). Feature Subset Selection Using a Genetic Algorithm. **Invited** paper. IEEE Intelligent Systems vol. 13, no. 2., pp. 44-49. **IEEE Computer Society's Top Ten Articles list, Nov 2007**

130. Yang, J. & Honavar, V. (1998). Experiments with the Cascade-Correlation Algorithm. Microcomputer Applications. Vol. 17, pp. 40-46.

131. Balakrishnan, K. & Honavar, V. (1997). Spatial Learning for Robot Localization. In: Proceedings of International Conference on Genetic Programming. Stanford, CA. Morgan Kaufmann. pp. 389-397.

132. Mikler, A., Wong, J. & Honavar, V. (1997). An Object-Oriented Approach to Modelling and Simulation of Routing in Large Communication Networks. Journal of Systems and Software Vol. 40, pp 151-164.

133. Mikler, A., Wong, J. & Honavar, V. (1997). Quo Vadis - A Framework for Intelligent Routing in Large High Speed Communication Networks. Journal of Systems and Software. 37 61-73.

134. Miller, L., Honavar, V. & Barta, T.A. (1997). Warehousing Structured and Unstructured Data for Data Mining. In: Proceedings of the American Society for Information Science Annual Meeting (ASIS 97). Washington, D.C.

135. Parekh, R.G. & Honavar, V. (1997) Learning DFA from Simple Examples. In Proceedings of the International Workshop on Algorithmic Learning Theory. Berlin: Springer-Verlag. Lecture notes in Computer Science vol. 1316. pp. 116-131.

136. Parekh, R.G., Yang, J. & Honavar, V. (1997). MUPStart - A Constructive Neural Network Learning Algorithm for Multi-Category Pattern Classification. In: Proceedings of IEEE International Conference on Neural Networks. Houston, TX. pp 1924-1929.

137. Parekh, R.G., Yang, J. & Honavar, V. (1997). Pruning Strategies for Constructive Neural Network Learning Algorithms. In: Proceedings of IEEE International Conference on Neural Networks. Houston, TX. pp. 1960-1965.

138. Yang, J. & Honavar, V. (1997). Feature Subset Selection Using a Genetic Algorithm. In: Proceedings of International Conference on Genetic Programming. Stanford, CA. Morgan Kaufmann. pp. 380-385.

139. Zhou, G., McCalley, J., & Honavar, V. (1997). Power System Security Margin Prediction Using Radial Basis Function Networks. In: Proceedings of the North American Power Symposium, Laramee, WY.

140. Balakrishnan, K. & Honavar, V. (1996). Analysis of Neurocontrollers Designed by Simulated Evolution. In: Proceedings of the IEEE International Conference on Neural Networks. Washington, D.C.

141. Balakrishnan, K. & Honavar, V. (1996). On Sensor Evolution in Robotics. In: Proceedings of the Conference on Genetic Programming. Koza, Goldberg, Fogel, Riolo (Ed.) Cambridge, MA: MIT Press. pp. 455-460.

142. Balakrishnan, K. & Honavar, V. (1996). Experiments in Evolutionary Neuro-robotics. In: Proceedings of the World Congress on Neural Networks. San Diego, CA. pp. 1025-1031.

143. Chen, C. & Honavar, V. (1996). A Neural Network Architecture for High-Speed Database Query Processing. Microcomputer Applications 15 7-13.

144. Mikler, A., Honavar, V. & Wong, J. (1996). Utility-Theoretic Heuristics for Intelligent Adaptive Routing in Large Communication Networks. In: Proceedings of the Fourth International Conference on Telecommunication Systems. Nashville, TN. pp. 660-676.

145. Mikler, A., Honavar, V. & Wong, J. (1996). Analysis of Utility-Theoretic Heuristics for Intelligent Adaptive Routing. In: Proceedings of the Thirteenth National Conference on Artificial Intelligence. Portland, Oregon. vol. 1. pp. 96-102.

146. Parekh, R. & Honavar, V. (1996). An Incremental Interactive Algorithm for Regular Grammar Inference. In: Proceedings of the Third International Colloquium on Grammar Inference (ICGI-96). Berlin: Springer-Verlag. pp. 238-249.

147. Yang, J., Parekh, R. & Honavar, V. (1996). MTiling: A Constructive Neural Network Learning Algorithm for Multi-Category Pattern Classification. In: Proceedings of the World Congress on Neural Networks. San Diego, CA. pp. 182-188.

148. Yang, J. & Honavar, V. (1996). A Simple Randomized Quantization Algorithm for Network Pattern Classifiers. In: Proceedings of the World Congress on Neural Networks. San Diego, CA. pp. 223-229.

149. Balakrishnan, K. & Honavar, V. (1995). Properties of Genetic Representations of Neural Architectures. Proceedings of the World Congress on Neural Networks. Washington, D.C. pp. 807-813.

150. Chen, C., Yang, J., Balakrishnan, K., Parekh, R. & Honavar, V. (1995). Analysis of Decision Boundaries Generated by Constructive Neural Network Learning Algorithms. In: Proceedings of the World Congress on Neural Networks. Washington, D.C. pp. 628-635.

151. Chen, C. & Honavar, V. (1995). A Neural Memory Architecture for Content as Well as Address-Based Storage and Recall: Theory and Applications. Connection Science. 7 293-312.

152. Mikler, A., Wong, J. & Honavar, V. (1995). Adaptive Heursistic Routing in Very Large High Speed Communication Networks Using Quo Vadis: Experimental Results. In: Proceedings of the Third International Conference on Telecommunication Systems, Modelling, and Analysis. Nashville, TN. pp. 66-75.

153. Chen, C. & Honavar, V. (1994). Neural Network Automata. In: Proceedings of the World Congress on Neural Networks. San Diego, CA. pp. 470-477.

154. Mikler, A., Wong, J. & Honavar, V. (1994). Quo Vadis - A Framework for Intelligent Traffic Management. In: Proceedings of the International Conference on Intelligent Information Management Systems. Washington, D. C. pp. 25-28.

155. Honavar, V. & Uhr, L. (1993). Generative Learning Structures and Processes for Generalized Connectionist Networks. Information Sciences 70 75-108.

156. Janakiraman, J. & Honavar, V. (1993). Adaptive Learning Rate for Speeding up Learning in Backpropagation Networks. In: SPIE Conference on Artificial Neural Networks, Proceedings of the SPIE. vol. 1966. pp. 225-235. Bellingham, MA: SPIE.

157. Janakiraman, J. & Honavar, V. (1993). Adaptive Learning Rate for Increasing Learning in Backpropagation Networks. In: Proceedings of the World Congress on Neural Networks. vol. IV. pp. 378-381. Portland, Oregon.

158. Mikler, A., Wong, J. & Honavar, V. (1993). Quo Vadis - A Framework for Adaptive Routing in Very Large High-Speed Communication Networks. In: Alspector, J., Goodman, R. & Brown, T. X. (Ed.) Proceedings of the International Workshop on Applications of Neural Networks to Telecommunications. pp. 196-202. Princeton, New Jersey. Lawrence Erlbaum.

159. Parekh, R. & Honavar, V. (1993). Efficient Learning of Regular Languages Using Teacher-Generated Positive Samples and Learner-Generated Queries. In: Proceedings of the Fifth UNB AI Symposium. Goldfarb, L. (Ed.) New Brunswick, Canada. pp. 195-203.

160. Thambu, P., Honavar, V. & Barta, T. (1993). Knowledge-base Consistency Maintenance in an Evolving Intelligent Advisory System. In: Proceedings of FLAIRS-93. pp. 116-120. Fort Lauderdale, Florida.

161. Yang, J. & Honavar, V. (1993). A Hierarchical Representation Scheme for Three-Dimensional Object Recognition and Description. In: Proceedings of FLAIRS-93. pp. 194-197. Fort Lauderdale, Florida.

162. Balakrishnan, K. & Honavar, V. (1992). Faster Learning in Multi-Layer Networks by handling output-layer flat-spots. In: Proceedings of International Joint Conference on Neural Networks. Beijing, China. vol. 2. pp. 486-492.

163. Balakrishnan, K. & Honavar, V. (1992). Improving Convergence of Back-propagation by Handling Flat-Spots in the Output Layer. In: Proceedings of the Second International Conference on Artificial Neural Networks. pp. 1003-1009. Brighton, UK.

164. Honavar, V. (1992). Some Biases for Efficient Learning of Spatial, Temporal, and Spatio-Temporal Patterns. In: Proceedings of International Joint Conference on Neural Networks. Beijing, China. vol. 2. pp. 419-27.

165. Honavar, V. (1992). Inductive Learning Using Generalized Distance Measures. SPIE Conference on Adaptive and Learning Systems. Proceedings of SPIE. vol. 1706. pp. 50-58. Bellingham, MA: SPIE.

166. Mikler, A., Honavar, V., & Wong, J. (1992). Simulating a Traveller: A Heuristic Approach to Routing in Large Communication Networks. In: Proceedings of the European Simulation Symposium. pp. 297-301. Dresden, Germany.

167. Parekh R., Balakrishnan, K. & Honavar, V. (1992). Empirical Comparison of Flat-Spot Elimination Techniques in Back-propagation Networks. In: Proceedings of Simtec/WNN92. Houston, Texas. pp. 463-468.
168. Spartz, R. & Honavar, V. (1992). An Empirical Analysis of the Expected Source Values Rule. In: Proceedings of International Joint Conference on Neural Networks. Beijing, China. vol. 2. pp. 480-485.
169. Yang, J. & Honavar, V. (1991). Experiments with the Cascade Correlation Algorithm. In: Proceedings of the Fourth UNB Artificial Intelligence Symposium. Fredericton, Canada. pp. 369-380.
170. Honavar, V. & Uhr, L. (1990). Successive Refinement of Multi-Resolution Internal Representations of the Environment in Connectionist Networks. In: Proceedings of the Second Conference on Neural Networks and Parallel-Distributed Processing. Indiana University-Purdue University. pp. 90-99.
171. Honavar, V. & Uhr, L. (1990). Coordination and Control Structures and Processes: Possibilities for Connectionist Networks. Journal of Experimental and Theoretical Artificial Intelligence 2 277-302.
172. Honavar, V. & Uhr, L. (1989). Brain-Structured Connectionist Networks that Perceive and Learn. Connection Science 1 139-159.
173. Honavar, V. & Uhr, L. (1989). Generation, Local Receptive Fields, and Global Convergence Improve Perceptual Learning in Connectionist Networks. In: Proceedings of the 1989 International Joint Conference on Artificial Intelligence, San Mateo, CA: Morgan Kaufmann. pp. 180-185.
174. Honavar, V. & Uhr, L. (1988). A Network of Neuron-Like Units That Learns To Perceive By Generation As Well As Reweighting Of Its Links. In: Proceedings of the 1988 Connectionist Models Summer School, D. S. Touretzky, G. E. Hinton & T. J. Sejnowski (ed). pp. 472-484. San Mateo, CA: Morgan Kaufmann.

Invited or Refereed Book Chapters

1. Bao, J., Voutsadakis, G., Slutzki, G., and Honavar, V. (2007). P-DL: A Semantic Importing Approach to Selective Knowledge Reuse In Modular Ontologies. Ontology Modularization. Parent, C., Spaccapietra, S., and Stuckenschmidt, H. Berlin: Springer. To appear.
2. Caragea, D. and Honavar, V. (2007). Learning Classifiers from Distributed Data. In: Encyclopedia of Database Technologies and Applications, Ferraggine, V.E., Doorn, J.H., and Rivero, L.C. (Ed). New York: Idea Group. In press.
3. Caragea, D. and Honavar, V. (2007). Learning Classifiers from Semantically Heterogeneous Data. In: Encyclopedia of Data Warehousing and Mining. Wang, J. (Ed). In press.
4. Pathak, J., Basu, S., and Honavar, V. (2007). Assembling Composite Web Services from Autonomous Components. In: Emerging Artificial Intelligence Applications in Computer Engineering. Soldatos, J. et al. (ed). IOS Press. In press.
5. Honavar, V. Miller, L., and Wong, J. (2004). Distributed Knowledge Networks. In: Unifying Themes in Complex Systems. Bar-Yam, Y. and Minai, A. (Ed). New York: Perseus Books.
6. McCalley, J., Honavar, V., Zhang, Z., and Vishwanathan, V. (2004). Multiagent negotiation models for power system applications**, Invited Chapter.** In: Autonomous Systems and Intelligent Agents in Power System Control and Operation. Christian Rehtanz (Ed). 2004.
7. Pai, P., L.L. Miller, V. Honavar, J. Wong, and S. Nilakanta. (2002). Supporting Organizational Knowledge Management with Agents. in D. White (eds) Knowledge Mapping and Management. Pages 266-280.
8. Caragea, D., Silvescu, A., and Honavar, V. (2001). **Invited** Chapter. Towards a Theoretical Framework for Analysis and Synthesis of Agents That Learn from Distributed Dynamic Data Sources. In: Emerging Neural Architectures Based on Neuroscience. Berlin: Springer-Verlag. pp. 547-559.
9. Balakrishnan, K. & Honavar, V. (2001). Evolving Neurocontrollers and Sensors for Artificial Agents. In: Evolutionary Synthesis of Intelligent Agents. Patel, M., Honavar, V. and Balakrishnan, K. (Ed). Cambridge, MA: MIT Press. pp. 109-152. In press.

10. Chen, C. & Honavar, V. (2000). A Neural Network Architecture for Information Retrieval and Query Processing. **Invited** chapter In: Handbook of Natural Language Processing. Dale, Moisl & Somers (Ed). New York: Marcel Dekker. pp. 873-888.
11. Honavar, V. and Balakrishnan, K. (2001). Evolutionary Synthesis of Neural Architectures for Intelligent Agents. In: Evolutionary Synthesis of Intelligent Agents. Patel, M., Honavar, V. & Balakrishnan, K. (Ed). Cambridge, MA: MIT Press. In press. pp. 1-28.
12. Balakrishnan, K., and Honavar, V. (2000). Some Experiments in the Evolution of Robot Sensors. **Invited** chapter. In: Evolution of Engineering and Information Systems and Their Applications. Jain, L. (Ed). New York: CRC Press. pp. 191-228.
13. Parekh, R. & Honavar, V. (2000). Automata Induction, Grammar Inference, and Language Acquisition. **Invited** chapter. In: Handbook of Natural Language Processing. Dale, Moisl & Somers (Ed). New York: Marcel Dekker. pp. 727-764. In press.
14. Honavar, V., Parekh, R., and Yang, J. (1998). Constructive Learning and Structural Learning. **Invited** article In: Encyclopedia of Electrical and Electronics Engineering, Webster, J. (Ed.), New York: Wiley. Vol. 4. pp. 226-231.
15. Honavar, V. (1998). Machine Learning. **Invited** article In: *Encyclopedia of Electrical and Electronics Engineering*, Webster, J. (Ed.), New York: Wiley. Vol. 11. pp. 656-659.
16. Yang, J. & Honavar, V. (1998). Feature Subset Selection Using A Genetic Algorithm. **Invited** chapter. In: Feature Extraction, Selection, and Construction: A Data Mining Perspective. Liu, H. & Motoda, H. (Ed.) New York: Kluwer. pp. 117-136.
17. Honavar, V., & Uhr, L. (1995). Integrating Symbol Processing and Connectionist Networks. **Invited** chapter. In: Intelligent Hybrid Systems. Goonatilake, S. & Khebbal, S. (Ed). London: Wiley. pp. 177-208.
18. Honavar, V. (1994). Symbolic Artificial Intelligence and Numeric Artificial Neural Networks: Toward a Resolution of the Dichotomy. **Invited** chapter In: Computational Architectures for Integrating Symbolic and Neural Processes. Sun, R. & Bookman, L. (Ed). New York: Kluwer. pp. 351-388.
19. Honavar, V. (1994). Toward Learning Systems That Integrate Multiple Strategies and Representations. In: Artificial Intelligence and Neural Networks: Steps Toward Principled Integration. Honavar, V. & Uhr, L. (Ed). pp. 615-644. New York: Academic Press.
20. Uhr, L. & Honavar, V. (1994). Artificial Intelligence and Neural Networks: Steps Toward Principled Integration. In: Artificial Intelligence and Neural Networks: Steps Toward Principled Integration. Honavar, V. & Uhr, L. (Ed). pp. xvii-xxxii. New York: Academic Press.

Juried Papers, Extended Abstracts, and Posters in Conferences and Workshops

1. Towfic, F., Gemperline, D., Caragea, C., Wu, F., Dobbs, D., and Honavar, V. Structural Characterization of RNA-Binding Sites of Proteins: Preliminary Results. *IEEE BIBM Computational Structural Bioinformatics Workshop*, In press, 2007.
2. Andorf, C., Dobbs, D., and Honavar, V. Potential Errors in Mouse Protein Gene Ontology Annotations Returned by AmiGO. Oral Presentation in: Gene Ontology Users Workshop, MGED, Seattle, Washington, September, 2006.
3. Caragea, D. and Honavar, V. (2006). Knowledge Discovery from Disparate Earth Data Sources. Second NASA Data Mining Workshop: Issues and Applications in Earth Sciences. Poster Session. Pasadena, CA, May 23-24, 2006.
4. Pathak, J., Basu, S., and Honavar, V. (2006). Modeling Web Service Composition Using Symbolic Transition Systems. AAAI '06 Workshop on AI-Driven Technologies for Services-Oriented Computing (AI-SOC), Boston, MA, 2006.
5. J. Pathak, S. Basu, R. Lutz, and V. Honavar. (2006). MoSCoE: A Framework for Modeling Web Service Composition and Execution. IEEE Conference on Data Engineering Ph.D. Workshop, Atlanta, GA, 2006.
6. Sander, J., Fu, F. Terribilini, M., Townsend, J., Winfrey, R., Wright, D., Lee, J.J., Zaback, P., F. Wu, F., Honavar, V., Voytas, D. and Dobbs, D. (2006) Designing C2H2 Zinc Finger Proteins to Target Specific

Sites in Genomic DNA. 10th Annual Pacific Symposium on Biocomputing (PSB 2006), Maui, Hawaii. Poster Presentation.

7. Terribilini, M., Lee, J.H., Yan, C., Jernigan, R., Carpenter, S., Honavar, V., and Dobbs, D. (2006) Identifying interaction sites in "recalcitrant" proteins: Predicted protein and RNA binding sites in Rev proteins of HIV and EIAV agree with experimental data. 10th Pacific Symposium on Biocomputing. Maui, Hawaii (PSB 2006). Oral Presentation.
8. Andorf, C., Dobbs, D., and Honavar, V. (2006) Learning classifiers for assigning sequences to subcellular localization families. Intelligent Systems in Molecular Biology (ISMB 2006), Fortaleza, Brazil. Poster Presentation.
9. EL-Manzalawy, Y., Caragea, C., Dobbs, D., Honavar, V. (2006) On the quality of motifs for protein phosphorylation site prediction. Intelligent Systems in Molecular Biology (ISMB 2006), Fortaleza, Brazil. Poster Presentation.
10. Terribilini, M., Sander, J., Olson, B., Lee, J.-H., Jernigan, R., Honavar, V., and Dobbs, D. (2006) A computational method to identify amino acid residues involved in protein-RNA interactions. Intelligent Systems in Molecular Biology (ISMB 2006), Fortaleza, Brazil. Poster Presentation.
11. Andorf, C., Silvescu, A., Dobbs, D. and Honavar, V. (2005). Learning Classifiers for Assigning Proteins to Gene Ontology Functional Families. Poster Presentation. Intelligent Systems in Molecular Biology (ISMB 2005).
12. Bao, J. and Honavar, V. (2005). Collaborative Package-Based Ontology Building and Usage. In: IEEE Workshop on Knowledge Acquisition from Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources. Held in conjunction with the IEEE International Conference on Data Mining (ICDM 2005), Houston, Tx.
13. Caragea, C., Caragea, D., and Honavar, V. (2005). Learning Support Vector Machine Classifiers from Distributed Data. Proceedings of the 22nd National Conference on Artificial Intelligence (AAAI 2005).
14. Caragea, D., Silvescu, A., Bao, J., Pathak, J., Andorf, C., Yan, C., Dobbs, D., and Honavar, V. (2005) Poster presentation. Knowledge Acquisition from Semantically Heterogeneous, Autonomous, Distributed Data Sources. Intelligent Systems in Molecular Biology (ISMB 2005).
15. Terribilini, M., Yan, C., Lee, J-H, Honavar, V. and Dobbs, D. (2005). Computational Prediction of RNA binding sites in Proteins based on Amino Acid Sequence. Poster Presentation. Intelligent Systems in Molecular Biology (ISMB 2005).
16. Terribilini, M., Lee, J-H., Sen, T., Yan, C., Andorf, C., Sparks, W., Carpenter, S., Jernigan, R., Honavar, V., and Dobbs, D, (2005). Computational Identification of RNA binding sites in proteins. Poster presentation. Pacific Symposium on Biocomputing (PSB 2005).
17. Vasile, F., Silvescu, A., Kang, D-K., and Honavar, V. (2005): TRIPPER: Rule Learning Using Attribute Value Taxonomies. In: AAAI-05 Workshop on Human-Comprehensible Machine Learning.
18. Yan, C., Terribilini, M., Wu, F., Dobbs, D. and Honavar, V. (2005). A Computational Method for Identifying Amino Acid Residues Involved in Protein-DNA interactions. Poster Presentation. Intelligent Systems in Molecular Biology (ISMB 2005).
19. Bao, J. and Honavar, V. Collaborative Ontology Building with Wiki@nt. In: Proceedings of the Third International Workshop on Evaluation of Ontology Building Tools, http://km.aifb.uni-karlsruhe.de/ws/eon2004/EON2004_Proceedings.pdf. The Third International Semantic Web Conference (ISWC-0004), Hiroshima, 2004.
20. D.-K. Kang, A. Silvescu, J. Zhang, and V. Honavar, "Generation of Attribute Value Taxonomies from Data and Their Use in Data-Driven Construction of Accurate and Compact Naive Bayes Classifiers," Proceedings of ECML/PKDD-2004 Knowledge Discovery and Ontologies Workshop (KDO-2004), Pisa, Italy, September 24, 2004.
21. Silvescu A., and Honavar V. (2004). A Graphical Model for Shallow Parsing Sequences. In: The AAAI-04 Workshop on Adaptive Text Extraction and Mining (ATEM-2004). July 2004, San Jose CA.

22. Yan, C., Honavar, V. and Dobbs, D. (2004). Application of a Two-Stage Method for Identification of Protein-Protein Interface Residues. Poster Presentation. Eighth Annual International Conference on Research in Computational Molecular Biology (RECOMB 2004).
23. Caragea, D., Reinoso-Castillo, J., Silvescu, A., and Honavar, V. Statistics Gathering for Learning from Heterogeneous, Distributed, Autonomous Data Sources. In: Proceedings of the Workshop on Information Integration on the Web. International Joint Conference on Artificial Intelligence, Acapulco, Mexico, 2003.
24. Atramentov, A., and Honavar, V. Speeding up Multi-Relational Data Mining. In: Proceedings of the Workshop on Learning Statistical Models from Relational Data. International Joint Conference on Artificial Intelligence, Acapulco, Mexico, 2003.
25. Honavar, V., Dobbs, D., Jernigan, R., Caragea. D., Reinoso-Castillo, J., Silvescu, A., Pathak, J., Andorf, C., Yan, C., and Zhang, J. (2003). Algorithms and Software for Information Extraction, Integration, and Data-Driven Knowledge Acquisition from Heterogeneous, Distributed, Autonomous, Biological Information Sources. Poster Presentation. Biomedical Information Science and Technology Initiative (BISTI) Symposium; Digital Biology: The Emerging Paradigm. National Institutes of Health.
26. Silvescu A., and Honavar V. (2003) Ontology Elicitation: Structural Abstraction = Structuring + Abstraction + Multiple Ontologies. Poster presentation. Learning Workshop, Snowbird, Utah, 2003.
27. Leiva, H., Atramentov, A., and Honavar, V. (2002). Experiments with MRDTL -- A Multirelational Decision Tree Learning Algorithm. In: Proceedings of the Workshop on Multi-Relational Decision Tree Learning. http://www-ai.ijs.si/SasoDzeroski/MRDM2002/
28. Silvescu, A., Reinoso-Castillo, J., Andorf, C., Honavar, V., and Dobbs, D. (2001). Ontology-Driven Information Extraction and Knowledge Acquisition from Heterogeneous, Distributed Biological Data Sources. In: Proceedings of the IJCAI-2001 Workshop on Knowledge Discovery from Heterogeneous, Distributed, Autonomous, Dynamic Data and Knowledge Sources.
29. Bhatt, R., Balakrishnan, K., and Honavar, V. (2000). Representation and Learning of Spatial Maps. In: Workshop on Machine Learning of Spatial Knowledge, International Conference on Machine Learning (ICML-2000), Stanford University.
30. Caragea, D., Silvescu, A., and Honavar, V. (2000). Distributed and Incremental Learning Using Extended Support Vector Machines. In: Proceedings of the 17th National Conference on Artificial Intelligence. Austin, TX.
31. Caragea, D., Silvescu, A., and Honavar, V. (2000). Multi-Agent Learning from Distributed Data Sources. In: Workshop on Multi-Agent learning: Theory and Practice International Conference on Machine Learning (ICML-2000), Stanford University.
32. Caragea, D., Silvescu, A., and Honavar, V. (2000). Agents that Learn from Distributed Dynamic Data Sources. In: Proceedings of the Workshop on Learning Agents, Agents 2000/ECML 2000. Stone, P. and Sen, S. (Ed.) ECML. Barcelona, Spain. pp. 53-61.
33. Caragea, D., Silvescu, A., and Honavar, V. (2000). Distributed, Parallel, and Incremental Learning: A Theoretical Model. In: KDD 2000 Workshop on Distributed and Parallel Knowledge Discovery. Boston, MA.
34. Silvescu, A., and Honavar, V. (2000). Genetic Network Inference from Gene Expression Data. In: Workshop on Learning from Sequential and Temporal Data. International Conference on Machine Learning (ICML 2000), Stanford University.
35. Helmer, G., Wong, J., Honavar, V., and Miller, L. (1999). Data-Driven Induction of Compact Predictive Rules for Intrusion Detection from System Log Data. In: Proceedings of the Conference on Genetic and Evolutionary Computation (GECCO 99). San Mateo, CA: Morgan Kaufmann. pp. 1781.
36. Tiyyagura, A., Chen, F., Yang, J., and Honavar, V. (1999). Feature Subset Selection in Rule Induction. In: Proceedings of the Conference on Genetic and Evolutionary Computation (GECCO 99). San Mateo, CA: Morgan Kaufmann. pp. 1800.

37. Honavar, V. (1997). The Design Process: A Computational Perspective. In: Proceedings of the NSF Workshop on Decision-based Design, Sacramento, CA.
38. Balakrishnan, K. & Honavar, V. (1996). Experiments in Evolutionary Synthesis of Neurocontrollers. In: Proceedings of the Thirteenth National Conference on Artificial Intelligence. AAAI Press. vol. 2. pp. 1378.
39. Parekh, R. & Honavar, V. (1996). An Incremental Interactive Algorithm for Grammar Inference. In: Proceedings of the Thirteenth National Conference on Artificial Intelligence. AAAI Press. vol. 2. pp. 1397.
40. Parekh, R., Yang, J. & Honavar, V. (1996). In: Proceedings of the Thirteenth National Conference on Artificial Intelligence. AAAI Press. vol. 2. pp. 1398.
41. Honavar, V. (1993). Learning with Symbolic and Subsymbolic Representations: Some Possibilities for Vision. In: Proceedings of the AAAI Fall Symposium on Machine Learning in Computer Vision. Raleigh, North Carolina. (Also published as AAAI Tech. Rep. FS 93-04). pp. 162-166.
42. Honavar, V. (1992). Generalized Distance Measures - A Basis for the Integration of Symbolic and Connectionist Learning. In: Workshop on Integrating Neural and Symbolic Processes - The Cognitive Dimension. AAAI-92, San Jose, California.
43. Honavar, V. (1992). Symbolic and Sub-symbolic Computation in Biological Neural Circuits and Systems. In: Neural Information Processing Systems Post-Conference Workshop on Symbolic and Sub-symbolic Computation in Biological Neural Circuits and Systems. Vail, Colorado.
44. Honavar, V. (1991). Generative Learning in Generalized Connectionist Networks. In: Constructive Induction Session - Eighth International Workshop on Machine Learning. Evanston, IL.
45. Honavar, V. (1991). Language and Knowledge: Communication, Acquisition, and Evolution. Invited presentation in: Second International Workshop on Human and Machine Cognition. Perdido Key, Florida.
46. Mikler, A., Honavar, V. & Wong, J. (1992). A Knowledge-Based Approach to Dealing With Uncertain and Incomplete Information in Communication Network Management. In: Proceedings of the First Canadian Workshop on Uncertainty Management: Theory and Practice. Vancouver, B. C., Canada. pp. 30-38.
47. Honavar, V. (1991). Toward Integrated Models of Natural Language Evolution, Development, Acquisition, and Communication in Multi-Agent Environments. In: Powers, D. and Reeker, L. (Ed.) Proceedings of the AAAI Spring Symposium on Machine Learning of Natural Language and Ontogeny. (MLNLO '91) pp. 82-86. Kaiserslautern, Germany: German AI Centre (DFKI).
48. Honavar, V. (1990). Toward Generalized Connectionist Networks: An Integration of Symbolic and Sub-Symbolic Approaches to the Design of Intelligent Systems. In: AAAI-90 Workshop on the Integration of Symbolic and Neural Processes. Boston, MA.
49. Honavar, V. (1990). Generative Learning Algorithms for Connectionist Networks. In: NIPS-90 Post-Conference Workshop on Constructive and Destructive Learning Algorithms. Keystone, CO.

Invited Book Reviews

1. Honavar, V. (1990). Parallel Distributed Processing: Implications for Psychology and Neurobiology. Invited review. Connection Science.
2. Honavar, V. (1992). Neural Network Design and the Complexity of Learning. Invited review. Machine Learning 9 95-98.

Theses and Dissertations

1. Honavar, V. (1990). Generative Learning Structures and Processes for Generalized Connectionist Networks. Doctoral Dissertation. Madison, WI: Computer Sciences Dept. University of Wisconsin-Madison. Advisor: Professor Leonard Uhr.
2. Honavar, V. (1984). Automated Analysis of Dark-Field Autoradiographs. Masters Thesis. Philadelphia, PA: Center for Image Processing and Pattern Recognition. Department of Electrical and Computer Engineering. Drexel University. Advisor: Professor Oleh Tretiak.

INVITED LECTURES AND TUTORIALS

Plenary Lectures and Invited Talks at Conferences

1. **Keynote Talk,** Computational Structural Bioinformatics Workshop, IEEE Conference on Bioinformatics and Biomedicine, Silicon Valley, 2007.
2. **Invited Talk,** Making Biology and Medicine a Predictive Science. NSF Workshop on Biomedical Informatics. Oregon, 2007.
3. **Invited Talk,** Knowledge Acquisition from Semantically Disparate Distributed Data. NSF Workshop on Next Generation Data Mining and Cyber-Enabled Discovery, Baltimore, Maryland, 2007.
4. **Invited Talk,** On Selective Sharing and Reuse of Ontologies, Semantic Technology Conference, San Jose, CA, USA, May **2007**
5. **Keynote Talk,** Semantic Web for Collaborative e-Science, International Conference on Intelligent Sensing and Information Processing, Bangalore, India, 2006.
6. **Invited Lecture:** Querying Semantically Heterogeneous Data Sources from a User's Point of View, Semantic Technology Conference, San Jose, CA, USA, March 2006.
7. **Invited Plenary Talk,** Algorithms and Software for Collaborative Discovery from Semantically Heterogeneous, Distributed, Autonomous Information Sources Sixteenth International Conference on Algorithmic Learning Theory (ALT 2005) and Eighth International Conference on Discovery Science (DS 2005). Singapore.
8. **Plenary Talk,** Data-Driven Discovery of Macromolecular Sequence-Structure-Function Relationships. International Conference on Intelligent System Design and Applications, 2003.
9. **Invited Talk,** Agent-Based Distributed Intelligent Information Networks for Computational Inference and Knowledge Discovery in Bioinformatics. In: Workshop on Agents in Bioinformatics, Italy, 2002.
10. **Plenary Talk,** Computational Discovery of Protein Sequence-Structure-Function Relationships, Diversity in Information Science and Technology, Nebraska EPSCOR Conference, 2002
11. **Keynote Address**, Learning from Large, Distributed, Heterogeneous Data Sets. International Symposium on Artificial Intelligence (ISAI 2001), Kolhapur, India.
12. **Invited** Talk, Distributed Intelligent Information Networks. Midwestern Conference on Artificial Intelligence and Cognitive Science, 2000.
13. **Invited** Talk, Cumulative Learning in Open Environments. International Workshop on Current Computational Architectures Integrating Neural Networks and Neuroscience. Durham Castle, United Kingdom. 2000.
14. **Invited** Talk, Distributed Knowledge Networks. Artificial Intelligence for Distributed Information Networks (AiDIN '99) Workshop held during the 1999 National Conference on Artificial Intelligence (AAAI 99), Orlando, Florida. July 1999.

Selected Invited Colloquia

1. **Invited Colloquium,** Algorithms and Software for Knowledge Acquisition from Semantically Heterogeneous, Distributed Data Sources. Dept. of Electrical and Computer Engineering. University of Iowa. 2006.
2. **Invited Colloquium,** Algorithms and Software for Collaborative Discovery in Systems Biology. Dept. Biostatistics, Bioinformatics & Epidemiology. Medical University of South Carolina, 2006.
3. **Invited Talk,** Algorithms and Software for Knowledge Acquisition from Semantically Heterogeneous, Distributed, Autonomous Information Sources. Google Research, 2005.
4. **Invited Talk,** All Science is Computer Science. Iowa Undergraduate Consortium. Drake University, 2004.

5. **Invited Colloquium,** Computational Discovery of Protein Sequence-Structure-Function Relationships: Bioinformatics Infrastructure and Sample Applications. University of Wisconsin-Madison Biostatics and Medical Informatics Department. 2002.
6. **Invited Talk,** Algorithmic and Systems Approaches to Computer Assisted Knowledge Discovery from Biological Data. Iowa State University - University of Iowa Joint Workshop on Bioinformatics. November 3-4, 2000.
7. **Invited Talk,** Neuromimetic Adaptive Autonomous Intelligent Systems. Institute for Computer Applications in Science and Engineering. NASA-Langley Research Center. Hampton, VA. September 28, 1999.
8. **Invited** Colloquium, Kolmogorov Complexity and Computational Learning Theory: Some Emerging Connections and Recent Results. Center for Neural Basis of Cognition, Carnegie Mellon University and University of Pittsburgh, Pittsburgh, PA. 1998.
9. **Invited Colloquium,** Experiments in Evolutionary Robotics. Department of Mathematics and Computer Science, Grinnell College, Iowa. October 1996.
10. **Invited Lecture,** Data Mining and Knowledge Discovery. Irish Life, Des Moines, Iowa. September 1996.
11. **Invited Lecture,** Knowledge Acquisition through Machine Learning. Principal Mutual, Des Moines, Iowa. January 1994.
12. **Invited Colloquium,** Generalized Connectionist Networks and Processes for Intelligent Systems. International Computer Science Institute, Berkeley, *CA.* (1990).
13. **Invited Colloquium,** Generative Learning Structures and Processes for Generalized Connectionist Networks. Cognitive and Learning Systems Laboratory, Siemens Research, Princeton, NJ. (1990).

Invited Tutorials

1. Honavar, V. Tutorial: Semantics-Enabled e-Science Cyberinfrastructure for Data Mining, 15th Italian Symposium on Advanced Database Systems. Torre Canne, Italy.
2. Honavar, V. Tutorial: Machine Learning Approaches in Computational and Systems Biology. International Conference on Intelligent Sensing and Information Processing, Bangalore, India, 2006
3. Honavar, V. and Caragea, D. Tutorial: Semantic Web for Collaborative Knowledge Acquisition, IEEE International Conference on Digital Information Management, Bangalore, India, 2006.
4. Honavar, V. and Caragea, D. Tutorial: Collaborative Knowledge Acquisition from Semantically Disparate, Distributed Data Sources, 2006 International Symposium on Collaborative Technologies and Systems, Las Vegas, Nevada, USA, May 2006.
5. Honavar, V. and Caragea, D. Semantic Web Technologies for Collaborative Knowledge Acquisition, International Conference on Digital Information Management, Bangalore, India, December 2006.
6. Honavar, V. Intelligent Agents and Multi-Agent Systems IEEE Conference on Evolutionary Computation (CEC), Washington, DC. 1999.
7. Honavar, V. Computational Learning Theory, Genetic Programming Conference, Stanford, 1997.
8. Honavar, V. Intelligent Agents, Genetic Programming Conference, Madison, WI, 1998.

Invited Panel Presentations

1. Panel on Semantic Data Integration. NSF Workshop on Biomedical Informatics, 2007.
2. Panel on Learning in Knowledge-Based Systems. Second World Congress on Expert Systems. Lisbon, Portugal (1994).
3. Panel on Hybrid Architectures for Intelligent Systems. Second World Congress on Expert Systems. Lisbon, Portugal (1994).
4. Panel on Hybrid Intelligent Systems (SIGHI meeting) World Congress on Neural Networks. San Diego, U.S.A. (1994).

STUDENT MENTORING STATEMENT

I find working with graduate and undergraduate students to be an extremely rewarding experience. I work with exceptional Ph.D. students with diverse backgrounds – ranging from very theoretical to very experimental. I also enjoy working with M.S. students and undergraduates interested in research. Students in my group benefit from strong mentoring and close interaction on a daily basis within a collaborative research environment. Research-based training in my research group emphasizes development of skills and expertise necessary for the pursuit of a successful independent research career:

(a) Ability to identify and formulate fundamental research problems;
(b) Ability to critically review scientific work;
(c) Ability to conceive, plan, and propose research projects;
(d) Ability to develop creative and innovative solutions;
(e) Strong experimental and/or theoretical expertise in relevant areas;
(f) Effective writing and presentation skills;
(g) Ability to develop and sustain productive research collaborations;
(h) Strong sense of ethics and responsibility in conduct of research.

Fundamental scientific questions (e.g., what is the algorithmic basis of cumulative multi-task learning? how is information encoded, stored, retrieved, decoded, and used in biological systems? how can we precisely characterize the syntax and semantics of the language of macromolecular sequences?); or important practical problems (how do we extract, assimilate, and use information from heterogeneous, distributed, autonomous data and knowledge sources to facilitate collaborative scientific discovery in biology?) drive our research.

My group takes a problem-centered approach to research. In addition to all the usual requirements for successful research, this requires a willingness to acquire, adapt, develop, and apply techniques and tools from areas that lie outside the traditional boundaries of the discipline (e.g., Computer Science) or a sub-discipline (e.g., Machine Learning) when necessary to solve a research problem.

Graduate students who join my lab typically have a broad-based training in Computer Science or a closely related discipline. Many have a strong interest in developing algorithmic or computational models of intelligent behavior (including learning and multi-agent interaction). Some have an interest in developing and applying algorithmic tools for scientific discovery in computational biology and bioinformatics. Some have an interest in building scalable, flexible, extensible, robust, and open-ended distributed information systems. I encourage and nurture interaction among members of my group through research seminars and collaborative research projects.

All of my former Ph.D. students have taken up academic careers or research-oriented careers in the industry. M.S. graduates typically end up in industry. Undergraduates who have worked in my lab often pursue graduate study at one of the other universities with strong programs in Artificial Intelligence or a related area (e.g., Computational Biology).

GRADUATE STUDENT AND POSTDOCTORAL RESEARCH SUPERVISION

Postdoctoral Fellows

1. **Jie Bao** (2007-2008). Research Associate, Center for Computational Intelligence, Learning, and Discovery (CCILD), Iowa State University. Topic: Representing and Reasoning with Federated Ontologies: Selective Knowledge Reuse, Privacy-preserving reasoning. Jie is supported in part by a grant from the National Science Foundation and in part by CCILD.
2. **Doina Caragea** (2004-2006) Research Associate, Center for Computational Intelligence, Learning, and Discovery (CCILD), Iowa State University. Topic: Knowledge Acquisition from Semantically Heterogeneous, Distributed Information Sources. 2004-2006. Doina is supported in part by a grant from the National

Institutes of Health and in part by CCILD. Current Position: Assistant Professor of Computer Science, Kansas State University.

3. **Byron Olson** (2005-2006). Research Associate, Center for Computational Intelligence, Learning, and Discovery (CCILD), Iowa State University. Topic: Discovery of Macromolecular Sequence-Structure-Function Relationships, Transcriptomics, Proteomics, and Interactomics. Byron was supported in part by a grant from the National Institutes of Health and in part by CCILD. Current Position: Research Assistant Professor of Electrical and Computer Engineering, Arizona State University.

Major Professor, Ph.D.

Ph.D. Graduates

1. **Jie Bao** (Computer Science). Ph.D., 2007. Thesis: Representing and Reasoning with Modular Ontologies. Jie Bao was supported by a research assistantship in Computer Science funded in part by the National Science Foundation, the Iowa State University Center for Integrative Animal Genomics, and the Center for Computational Intelligence, Learning, and Discovery. Current Position: Research Associate, Center for Computational Intelligence, Learning, and Discovery, Iowa State University.
2. **Tyra Dunn** (Bioinformatics and Computational Biology, with M. Heather West Greenlee,). Ph.D., 2007. Thesis: Characterizing and Influencing Differentiation of Retinal Progenitor Cells. Tyra was supported in part by an Integrative Graduate Education and Research Training (IGERT) fellowship from the National Science Foundation.
3. **Jyotishman Pathak** (Computer Science). Ph.D., 2007. Thesis: Interactive and Verifiable Web Service Composition, Reformulation, and Adaptation. Jyotish was supported by a graduate research assistantship in Computer Science funded in part by grants from the National Science Foundation. Initial Employment: Research Scientist, Division of Biomedical Informatics, Mayo College of Medicine, Rochester, Minnesota.
4. **Dae-Ki Kang** (Computer Science). Abstraction, Aggregation, and Refinement Strategies for Machine Learning. Dae-Ki was funded by a Teaching assistantship from the Department of Computer Science and a Research assistantship funded by the National Science Foundation. Current Employment: Assistant Professor, Department of Computer Engineering, Dongseo University, Pusan, Korea.
5. **Jun Zhang** (Computer Science). Ontology Aware Learning Algorithms. 2005. Jun was supported in part by research assistantship funded by a grant from the National Science Foundation and a teaching assistantship from the ISU Department of Computer Science. Current Employment: Research Scientist, Fair Isaac, San Diego.
6. **Changhui Yan** (Bioinformatics and Computational Biology, with Drena Dobbs), Computational Approaches to Prediction of Protein-Protein, Protein-DNA and Protein-RNA Interfaces. 2005. Changhui was supported by a Plant Sciences Fellowship and a research assistantships funded by the ISU Graduate College and a grant from the National Institutes of Health. Current Employment: Assistant Professor of Computer Science, Utah State University.
7. **Doina Caragea** (Computer Science). Learning Classifiers from Semantically Heterogeneous, Distributed, Autonomous Data Sources. 2004. Initial Employment: Research Associate, Center for Computational Intelligence, Learning, and Discovery, Iowa State University. Current position: Assistant Professor of Computer Science, Kansas State University.
8. **Jihoon Yang** (Computer Science). Learning Agents for Information Retrieval and Knowledge Discovery, 1999. Initial Employment: Research Scientist, Information Sciences Laboratory, Hughes Research Laboratory, Malibu, CA. Current Employment: Assistant Professor of Computer Science, Sogang National University, Korea.
9. **Karthik Balakrishnan** (Computer Science). Biologically Inspired Computational Structures and Processes for Autonomous Agents and Robots, 1998. Initial Employment: Senior Research Scientist, Datamining Group, Allstate Research and Planning Center, Menlo Park, CA. USA. Current Employment: Vice President of Analytics, ISO Analytics, CA.

10. **Rajesh Parekh** (Computer Science), Machine Learning of Automata and Neural Network Pattern Classifiers, 1998. Initial Employment: Senior Research Scientist, Data mining Group, Allstate Research and Planning Center, Menlo Park, CA. USA. Current Employment: Director of Data Mining and Research, Yahoo!
11. **Chun-Hsien Chen** (Computer Science). Neural Architectures for Knowledge Representation and Inference, 1997. Initial Employment: Research Scientist, Advanced Technology Center, Computer and Communication Research Laboratories, Industrial Technology Research Institute, Taiwan. Current Employment: Associate Professor, Department of Information Management, Chang Gung University, Taiwan.
12. **Armin Mikler** (Computer Science, with Johnny Wong), Quo Vadis - A Framework for Intelligent Routing in Large Communication Networks, 1995. Initial Employment: Research Associate, Scalable Computing Laboratory, DOE Ames Lab, Ames, Iowa. Current Employment: Associate Professor of Computer Science, University of North Texas, Denton, TX, USA.

Current Ph.D. Students

1. **Tim Alcon** (Bioinformatics & Computational Biology; **with Heather Greenlee**)**.** In progress. Area: Modeling and Discovery of Genetic Regulatory Networks. Tim is supported by a Multidisciplinary Graduate Education and Training (MGET) fellowship funded by the USDA.
2. **Carson Andorf** (Bioinformatics & Computational Biology; **with Drena Dobbs**), In progress. Areas: Bioinformatics and Computational Biology; Ontology Assisted Data Mining and Data Visualization for Characterization of Macromolecular Structure-Function Relationships. Carson is supported by an IGERT fellowship funded by the National Science Foundation and a research assistantship funded by the National Institutes of Health. Expected Graduation: Fall 2007. Current employment: Research Scientist, Newlink Genetics, Ames, Iowa.
3. **Oliver Couture** (Bioinformatics & Computational Biology; **with Chris Tuggle**), In progress. Areas: Bioinformatics, Systems Biology, Gene Expression Analysis. Oliver is supported by an MGET fellowship funded by the United States Department of Agriculture. Expected Graduation: Spring 2008.
4. **Neeraj Koul** (Computer Science). In progress. Interests: Information Integration; Semantic Web; Data Mining.
5. **Yasser El-Manzalawy** (Computer Science), In progress. Areas: Machine Learning, Data Mining, Probabilistic Relational Models, Semantic Web. Yasser is supported by a fellowship from the Egyptian Government. Expected graduation: Spring 2009.
6. **Laron Hughes** (Bioinformatics and Computational Biology; **with Jim Reecy**). Interests: Phenotype ontologies, Comparative Genomics. Expected Graduation: Fall 2007.
7. **Raphael J. Osorio** (Computer Science). In progress. Interests: Information Integration, Bioinformatics, Data Mining.
8. **Jivko Sinapov** (Computer Science, with Alex Stoytchev). In progress. Interests: Machine Learning, Developmental Robotics.
9. **Adrian Silvescu** (Computer Science), In progress. Areas: Intelligent Agents, Multi-Agent Systems, Machine Learning, Complex Adaptive Systems, Bioinformatics. Adrian has been supported by graduate assistantships funded by grants from Pioneer Hi-Bred and by the National Institutes of Health. Expected Graduation: Spring 2008.
10. **Kewei Tu** (Computer Science), In progress. Areas : Machine Learning, Data Mining, Probabilistic Relational Models. Kewei has been supported by a teaching assistantship in Computer Science. Expected Graduation: Spring 2010.
11. **Cornelia Caragea** (Computer Science), In progress. Areas: Machine Learning, Data Mining, Probabilistic Relational Models. Cornelia is being supported by a teaching assistantship in Computer Science. Expected Graduation: Spring 2009.
12. **Michael Terribilini**, (Bioinformatics & Computational Biology; **with Drena Dobbs**), In progress. Area: Bioinformatics and Computational Molecular Biology, protein-RNA interactions. Michael is supported by a Multidisciplinary Graduate Education and Training (MGET) fellowship funded by the USDA. Expected Graduation: Spring 2007.

13. **Fadi Towfic** (Bioinformatics and Computational Biology; with **M. Heather West Greenlee**). In progress. Interests: Computational Systems Biology; Genomics; Gene and Protein Networks. Machine Learning. Fadi is supported by a fellowship funded by an Integrative Graduate Education and Research Training (IGERT) grant from the National Science Foundation. Expected Graduation: Spring 2010.
14. **Flavian Vasile** (Computer Science), In progress. Areas: Interests: Machine Learning, Human-Computer Interaction, and the Semantic Web. Flavian is being supported by a teaching assistantship in Computer Science. Expected Graduation: Spring 2009.
15. **Kent Vander Velden** (Bioinformatics & Computational Biology; **with Peter Reilly**), In progress. Area: Systems Biology, Metabolic Networks, Computational Biology. Kent is an NSF IGERT fellow. He is at present a research scientist at Pioneer Hi-Bred and his doctoral work is supported by Pioneer Hi-Bred. Expected Graduation: Fall 2006.
16. **Oksana Yakhnenko** (Computer Science), in progress. Areas: Machine Learning, Multi-relational learning, Learning from Semantically Heterogeneous data, Probabilistic Relational Models, Data Visualization, Learning from Sequence Data. Oksana is funded by a graduate teaching assistantship in Computer Science and a research assistantship funded by the National Science Foundation. Expected Graduation: Spring 2009.
17. **Feihong Wu** (Bioinformatics & Computational Biology; **with Robert Jernigan**), In progress. Areas: Bioinformatics and Computational Biology; Data Mining and Data Visualization Algorithms for Characterization of Macromolecular Structure-Function Relationships. Feihong is supported by research assistantships funded by the Iowa State University Graduate College and a grant from the National Institutes of Health. Expected Graduation: Summer 2008.
18. **Peter Zaback** (Bioinformatics & Computational Biology; **with Drena Dobbs**), In progress. Area: Bioinformatics and Computational Biology. Expected Graduation: Spring 2009.

Major Professor, M.S.

Current M.S. Students

1. Ganesh Ram Santhanam (Computer Science), in progress. Topic: Service Oriented Computing.

M.S. Graduates

1. **Oksana Kohutyuk** (Computer Science), in progress. Areas: Machine Learning, Data Mining, Probabilistic Graphical Models, Causal Models, Bioinformatics, Gene Expression Analysis. Oksana is supported by a research assistantship funded by the National Institutes of Health and a teaching assistantship in Computer Science. 2007.
2. **Charles Giesler** A Java Reinforcement Learning Module for the Recursive Porous Agent Simulation Toolkit: Facilitating study and experimentation with reinforcement learning in social science multi-agent simulations. 2003. Current Position: Lawrence Livermore Labs.
3. **Anna Atramentov**, A Multi-Relational Decision Tree Learning Algorithm – Implementation and Experiments. 2003. Current position: Ph.D. Student, University of Illinois at Urbana-Champaign
4. **Zhong Gao**, Genome wide recognition of tumor necrosis factor (TNF) like ligands in human and Arabidopsis genomes: A structural genomics approach. 2003 Initial Employment: Post-doctoral fellow, The Center for Cardiovascular Bioinformatics and Modeling, Johns Hopkins University.
5. **Jaime Reinoso-Castillo**, Ontology-Driven Query-Centric Information Integration from Heterogeneous, Distributed, Autonmous Data Sources for Computer Assisted Scientific Discovery. 2002. Initial Employment: Universidad Javeriana, Colombia.
6. **Hector Leiva**, Learning Classifiers from Relational Data. 2002. Initial Employment: Research Scientist, Research Scientist, Universidad Nacional de San Luis. Argentina.
7. **Xiaosi Zhang**, Identification of Functionally Related Genes from Gene Expression Data. 2002.Initial Employment: Papajohn Center for Entreprenuership, Ames, Iowa.

8. **Xiangyun Wang**, Data Mining Approach to Discovery of Protein Sequence-Structure-Function Relationships. 2002. Initial Employment: Astra-Zeneca Inc.
9. **Kent Vander Velden** (joint supervision with Gavin Naylor), Spatial Clustering of Differences in Measured Homoplasy with Respect to Protein Structure. 2002. Initial Employment: Pioneer Hi-Bred, Inc.
10. **Neeraj Koul**, Clustering With Semi-Metrics, 2001. Initial Employment: Motorola.
11. **Dake Wang**, Data-Driven Generation of Decision Trees for Motif-Based Assignment of Protein Sequences to Functional Families., 2001. Initial Employment: Lumicyte, Inc. Current Employment: Genentech, Inc.
12. **Rushi Bhatt**, Spatial Learning and Localization: A Computational Model and Behavioral Simulations, 2001. Ph.D. Program, Boston University.
13. **Fajun Chen**, Learning Information Extraction Patterns from Text, 2000. Initial Employment: Ericsson.
14. **Tarkeshwari Sharma** Agent Toolkit for Distributed Knowledge Networks, 2000. Initial Employment: Motorola, Inc.
15. **Asok Tiyyagura**, Alternative Criteria for Association Rule Mining, 2000. Cisco Systems, Inc.
16. **Di Wang**, 1997. Mobile Agents for Information Retrieval.
17. **Shane Konsella**, 1996. Trie Compaction Using Genetic Algorithms. Initial Employment: Hewlett-Packard.
18. **Karthik Balakrishnan**, 1993. Faster Learning Approximations of Backpropagation by Handling Flat-Spots. Continued as a Ph.D. student.
19. **Jayathi Janakiraman**, 1993. Adaptive Learning Rate for Increasing Learning Speed in Backpropagation Networks. Initial Empoyment: Motorola.
20. **Priyamvadha Thambu**, 1993. Automated Knowledge-Base Consistency Maintenance in an Evolving Intelligent Advisory System. Initial Employment: Inference Corporation.
21. **Rajesh Parekh**, 1993. Efficient Learning of Regular Languages Using Teacher-Supplied Positive Examples and Learner-Generated Queries. Continued as a Ph.D. student.
22. **Richard Spartz**, 1992. Speeding Up Backpropagation Using Expected Source Values. Initial Employment: IBM.

Member of Graduate Program of Study (Thesis) Committees

Ph.D. Committees

1.	Alison Barnhill	Biomedical Sciences	In progress
2.	Facundo Bromberg	Computer Science	2007
3.	Jae-Hyung Lee	Bioinformatics and Computational Biology	2007
4.	Alexei Kroujiline	Economics	In progress
5.	Tu-Liang Lin	Computer Science	In progress
6.	Changsung Kang	Computer Science	2007
7.	Haitao Cheng	Bioinformatics and Computational Biology	In progress
8.	Brian Patterson	Computer Science	In progress
9.	Jeff Sander	Bioinformatics and Computational Biology	In progress
10.	Aimin Yan	Bioinformatics and Computational Biology	In progress
11.	Lei Yang	Bioinformatics and Computational Biology	In progress
12.	Yu Cao	Computer Science	2007
13.	Laura Hecker	Neuroscience	2007
14.	Oleksiy Atramentov	Physics	2006
15.	Xiaonan Li	Industrial and Manufacturing Systems Engg.	2006
16.	Di Wu	Bioinformatics and Computational Biology	2006
17.	Yungok Ihm	Bioinformatics and Computational Biology	2004
18.	Cizhiong Zhang,	Bioinformatics and Computational Biology	2004

19. Zhong Zhang,	Electrical and Computer Engineering	2004
20. Haibo Cao	Physics	2003
21. Brooke Peterson	Genetics	2003
22. Marybeth Gurski	Computer Science	2001
23. Guy Helmer	Computer Science	2001
24. Robi Polikar	Electrical and Computer Engineering	2001
25. Vincent Van Acker	Electrical and Computer Engineering	2000
26. Chun-Fu Chen	Economics	1999
27. Victoria Bascunana	Chemical Engineering	1999
28. Guozhong Zhou	Electrical and Computer Engineering	1998
29. Cheng-Chi Tai	Electrical and Computer Engineering	1998
30. James Lathrop	Computer Science	1997
31. Krishna Dhara	Computer Science	1997
32. Babak Fourouraghi	Computer Science	1995
33. Timothy Wahls	Computer Science	1995
34. Chang-Chun Tsai	Industrial and Manufacturing Engg.	1995
35. Sonmez Rifat	Civil and Constructional Engg.	1995
36. Richa Agrawala	Computer Science	1994
37. Bamshad Mobasher	Computer Science	1994
38. Hun Kang	Electrical and Computer Engg.	1993

M.S. Committees

1. Georgi Batinov	Economics	2007
2. Inya Nlenanya	Agricultural Engineering	2005
3. Jing Xu	Psychology	2005
4. Suxing Cheng	Computer Science	2005
5. Kyongryun Lee	Computer Science	2005
6. Haitao Cheng	Computer Science	2004
7. Patricia Lonosky	Genetics	2002
8. Mallika Bachan	Statistics	2002
9. Melinda Vander Velden	Electrical and Computer Engineering	2002
10. Jeremy Patterson	Computer Science	2001
11. Sa Lin	Computer Science	2001
12. Vijay Viswanathan	Electrical and Computer Engineering	2001
13. Fengmei Liu	Computer Science	2001
14. Xinhua Dong	Computer Science	2001
15. Mark Slagell	Computer Science	2001
16. Hao Dong	Computer Science	2001
17. Jun Li	Computer Science	2001
18. Thai-Tin Huang	Computer Science	2000
19. Ran Liu	Computer Science	2000
20. Xumei Lu	Computer Science	2000
21. Nanchang Yang	Computer Science	2000
22. Peng Han	Botany	1999
23. Jeffrey Yakey	Computer Science	1999

24. Sunitha Kothapalli	Electrical Engineering	2000
25. Raghunandan Havaldar	Computer Science	1998
26. Guy Helmer	Computer Science	1998
27. Ngee Jenn Lee	Mechanical Engineering	1998
28. Laura Nelson	Computer Science	1998
29. Jibin Xiang	Computer Science	1998
30. Venkat Naganathan	Computer Science	1998
31. Dean Stevens	Computer Science	1998
32. Prashant Pai	Computer Science	1998
33. Abhinav Rawat	Nuclear Engineering	1998
34. Rishi Nayar	Computer Science	1997
35. Marcie Goodman	Computer Science	1997
36. Jon Schultze-Hewett	Computer Science	1997
37. Chin Khor	Mechanical Engineering	1997
38. Qiang-lin Zhao	Computer Science	1996
39. Chi-Chuan Chen	Agricultural Engineering	1996
40. Mahesh Subramaniam	Computer Science	1996
41. Glen Holt	Computer Science	1996
42. Niranjan Vaidya	Computer Science	1996
43. Thirumalai Anandapillai	Industrial and Manufacturing Engineering	1995
44. Thomas DeWulf	Electrical Engineering	1994
45. Raghav Trivedi	Computer Science	1994
46. Arun Barboza	Computer Science	1994
47. Brian Schmidt	Electrical and Computer Engineering	1993
48. Brian Peterson	Computer Science	1992
49. Salim Chandani	Industrial and Manufacturing Engineering	1992
50. Prerana Vaidya	Computer Science	1991
51. James Wittry	Computer Science	1991
52. Srinivas Boddu	Electrical Engineering	1991
53. Simanta Mitra	Computer Science	1991

UNDERGRADUATE STUDENT RESEARCH SUPERVISION

Supervisor, Undergraduate Honors Project

1. Oksana Yakhneko (2003-2004). Topics in Machine Learning.
2. Eric Barsness (1993), An Object-Oriented Implementation of a Genetic Algorithms Testbed.
3. Daniel Graves (1992), Parallel Architectures for Artificial Intelligence.

Supervisor, Undergraduate Research

1. Tyson Williams, Undergraduate Research Assistant, 2007-2008
2. David Gemperline, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2007.
3. Keith Callenberg, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2007.
4. Remy Younes, Undergraduate Research Assistant. Topics in Data Integration. 2007.
5. Matt Miller, Topics in Learning Classifiers from Distributed Data, 2006-2007.
6. John Leacox, Topics in Data Integration, 2006.

7. Peter Wong, Topics in Collaborative Ontology Development, 2006.
8. Ryan Bruce (2004), Topics in Bioinformatics
9. Cody Pfau (2003). Topics in Data Mining.
10. Amy Nienaber (2003). Computational Discovery of Protein-Protein Interactions
11. Matthew Beard (2003). Computational Discovery of Protein-Protein Interactions
12. Diane Schroeder (2001-2002) Data Mining Approaches to Discovery of Protein Sequence Function *Relationships* (Graduate School: Stanford University)
13. Kent Vander Velden (1998-1999), Protein Structure Prediction. (Graduate School: Bioinformatics, Iowa State University)
14. Jeremy Ludwig (1996-1997), Topics in Neural Computing. (Graduate School: Intelligent Systems, University of Pittsburgh).
15. David DeYoe (1996-1997), Topics in Cognitive Modelling.
16. Carl Pecinovsky (1996-1997), Constructive Neural Network Learning Algorithms.
17. Brian Walenz, Topics in Genetic Algorithms (Graduate School: Computer Science, University of New Mexico)
18. Gabriel Ki (1996-), Situated Robotics.
19. Todd Lindsey (1995-96), Constructive Neural Network Learning Algorithms.
20. Jouko Rytilahti (1994), Explorations in Evolutionary Algorithms.
21. Leigh McMullen (1993), Adaptive Game-Playing Programs.

Mentor, Freshman Honors Study

1. Jeffrey Schroeder (1997) Topics in Artificial Intelligence.
2. Matthew Potter (1997) Topics in Artificial Intelligence.
3. Brian George (1994), Topics in Neural Computing.
4. Adam Johnson (1994), Topics in Neural Computing.
5. Marcus Ryan (1993), Topics in Artificial Intelligence.

PRE-COLLEGE STUDENT RESEARCH SUPERVISION

Mentor (for pre-college students)

1. Eric Solan, Nic Dayton, Luke Rolfes, and Julian Sheldahl. Animus Facticius. Adventures in Supercomputing (1998) project. First Place.
2. Sara Karbeling et al., A Computational Model of Animal Spatial Learning Adventures in Supercomputing (1998) project.
3. Sara Karbeling, Kellan Brumback, Anna Keyte, and Angel Sherif (1997), Lateral Inhibition and Sensory Processing in the Limulus Polyphemus Eye, Adventures in Supercomputing (AIS-97) project. First place in Iowa, and Second Place in the National Competition.
4. Stephen Lee (1993), Topics in Neural Computing.
5. John Farragher (1992), Topics in Neural Computing.
6. Peter Luka (1991), Topics in Neural Computing.

SPONSORSHIP OF VISITING RESEARCHERS

1. Oswaldo Velesz-Langs, Madrid Polytechnic University, Madrid, Spain. 2003.
2. Professor Yigon Kim (on Sabbatical from Yosu National University, Yosu, Korea 2000-2001) Data Mining and Knowledge Discovery.
3. Professor Mok Dong Chung (on Sabbatical from Pukyong National University, Korea, 1999-2000) Agent-based systems and knowledge-based systems.
4. Olivier Bousquet (from Ecole Polytechnique, France, 1997) Topics in Cognitive Modeling and Robotics.

5. Codrin Nichitiu (from ENS Lyon, France, Summer 1996), Topics in Machine Learning.
6. Dimitri Kotchetkov (Visitor from Ukraine, Summer 1996), Topics in Robotics.
7. Vadim Kirillov (Fullbright Scholar from Ukraine, 1995), Constraint-based Reasoning under uncertainty.

TEACHING STATEMENT

Teaching Philosophy

My teaching philosophy is perhaps best summed up by a quote from Joseph Chesterton: "The Foundation of teaching is research; and the object of research is teaching, that is, the dissemination of knowledge".

In my view, teaching is an integral and extremely rewarding part of academic life. I like to teach what I like to learn. It is hard to top the joy of grasping a new idea or solving an unsolved problem for the very first time. I subscribe to John Dewey's view of education: "Education is not preparation for life; education is life itself". For me, teaching is not just about communicating what is thought to be known, but cultivating the ability in students to challenge the current state of knowledge, and to venture beyond the current frontiers of knowledge into the unknown, on a life-long journey of learning and discovery. As Richard Dann remarked, "He who dares to teach must never cease to learn". Most of my curriculum development and teaching activities in computer science strongly complement, and are sustained by an active research program.

In my view, teaching computer science is not *just* about teaching students to be competent programmers or creators and users of sophisticated software tools; it is about introducing computer science as a profoundly interesting scientific discipline. Computer science, as a discipline provides us with the best language we have so far – the language of algorithms – precise recipes for describing processes that manipulate information – for modeling biological, cognitive, and social phenomena – just as calculus provided Newton and his successors with a language for modeling physical phenomena.

I was drawn to Computer Science because of the possibility of understanding biological, cognitive, and social phenomena in terms of processes that acquire, store, retrieve, manipulate, and use *information*. I am passionate about introducing students to algorithmic approaches to exploring fundamental questions in biological, cognitive, and social sciences and to the challenges of developing information processing artifacts and software that can dramatically improve our quality of life.

Teaching Style

My personal teaching style involves:

(a) Engaging students so that they become active participants in the learning process rather than passive observers ("What I hear, I forget. What I see, I remember. What I do, I understand" – Confucius).

(b) Setting the stage for the topic of study – We learn what we *almost* already know. Hence I find it useful to introduce a complex idea or a new topic or through a succession of smaller steps, establishing their connection with familiar topics, each building on what my students already know.

(c) Setting challenging yet realistic goals – In my experience, students respond best, and learn most, from learning goals that are both challenging and achievable. Assignments that are trivial are boring. Assignments that are excessively difficult can be frustrating and intimidating.

(d) Letting students make mistakes and learn from them. Exploration and experimentation are essential to learning and discovery and learning what does not work (and why) is as important as learning what works (and why).

(e) Accommodating different learning styles – Kolb identifies several learning styles. Some learn best from observations and examples; others by acquiring general principles and knowledge of how to apply them in specific situations; others by deducing and discovering general principles or theories; and yet others by doing – that is, by trying things, making mistakes, and learning from them. I find it useful to develop course materials that exercise multiple learning styles.

Teaching and Curriculum Development Experience

Over the past fourteen years, I have designed, developed, and taught undergraduate as well as graduate courses and seminars in artificial intelligence, intelligent agents and multi-agent systems, machine learning, data mining and knowledge discovery, neural and evolutionary computation, computational learning theory, bioinformatics and computational molecular biology. The material covered in the courses is chosen with an emphasis on concepts that are likely to have a lasting impact on the discipline in the years to come. In addition to introducing students to a core body of knowledge in the areas of study, these courses present such knowledge in the broader context of computer science as an intellectual discipline and to develop the students into creative thinkers and problem-solvers, be it in academic research or advanced technology development.

The undergraduate and graduate courses that I have developed and taught over the years introduce students to some of the most challenging topics in computer science - involving the application of concepts and tools from the theory of computation, design and analysis of algorithms, and design of software systems in the construction of *intelligent* artifacts: computer programs that represent and reason with and about knowledge, acquire knowledge from interaction with their environment, and discover and use regularities from data.

I have developed and taught a combined upper level undergraduate and introductory graduate course in Artificial Intelligence with emphasis on fundamental problems and approaches in the design of intelligent agents. Because no suitable textbook was available at the time I first offered this course (in 1991), I developed all of the necessary course material in the form of lecture notes. Some of these notes were adapted for use in similar courses at other universities. In 1995, when Russell and Norvig's textbook became available, I adopted it for use in this course with my lecture notes serving as supplementary material. This course is aimed at introducing students to the foundations of artificial intelligence, including intelligent agents, problem solving, knowledge representation and reasoning (including representing and reasoning with uncertainty, decision theory, planning, and machine learning. This course has been fairly popular with graduate and senior undergraduate students in Computer Science, Engineering, and related disciplines.

I have developed and taught a graduate course in Machine Learning which presents a unified framework for formulation and solution of a broad class of machine learning problems using mathematical tools drawn from probability theory, statistics, information theory, decision theory, and algorithm design and analysis. This course is designed to help students gain a broad understanding of the current state of the art in machine learning, adapt and apply machine learning approaches to real-world applications (e.g., in computational biology, semantic web), and begin to conduct original research in machine learning.

I have developed an undergraduate course on neural computation. Because no textbook on this topic suitable for undergraduate students in Computer Science was available in 1992 when I first offered the course, I developed most of the material for this course in the form of lecture notes some of which were adapted for use in graduate and undergraduate courses in other universities. The focus of the course was on computational models of neurons and networks of neurons (neural networks) and neural network learning algorithms. This course was popular with undergraduates in Computer Science, Engineering, as well as graduate students from other disciplines until the course was eventually superseded by a broad-based introduction to machine learning which I began to teach in 2000.

I have developed and taught a graduate course on intelligent agents and multi-agent systems which draws on material from game theory, knowledge representation and inference, decision theory, contract theory, bargaining theory, and related areas to develop an understanding of fundamental problems in the design of open-ended systems consisting of loosely coupled systems consisting of interacting autonomous entities (information sources, intelligent agents) e.g., the semantic web.

I have also developed and taught advanced graduate "topics in artificial intelligence" courses that have covered computational learning theory, probabilistic graphical models, information retrieval, computational molecular biology, knowledge representation and inference, neural computation, evolutionary algorithms, reinforcement learning, and intelligent agents and multi-agent systems, and semantic web.

I am currently developing a course on Functional Genomics and Computational Systems Biology which is the fourth in a sequence of 4-course core curriculum that I helped develop for the Bioinformatics and Computational Biology graduate program. This course will be taught for the first time in 2007.

In addition to the regular courses, current research topics are explored in depth in research seminars which I have organized or co-organized (with other faculty in Computer Science and Bioinformatics and Computational Biology) with help and active participation by my graduate students.

The nature of the material taught in my courses requires a delicate balance between theory and experimentation. In a fast-paced field like computer science in general and artificial intelligence and bioinformatics in particular, the courses have to anticipate key developments in the field that are likely to have a long-term impact and provide students with a solid understanding of the fundamentals as well the insight that comes with hands-on experience. Hence, I have invested in efforts to develop the laboratory facilities that are essential to support experiments, exercises, and projects that enhance the students' understanding the material covered in the courses.

To help develop the written and oral communication skills of students, most of my courses require individual or team research projects culminating in a short paper. It has been my experience that team projects promote collaborative learning and problem-solving. The projects often serve as vehicles for integrating latest research results into the graduate and undergraduate curriculum. They also provide an opportunity for students to exercise their creativity and explore new solutions to open problems in artificial intelligence. In many instances, such class projects have evolved into thesis research topics or produced results that were eventually published in refereed national and international conferences.

In addition to developing and teaching courses, I have had substantial experience developing new curricula and programs. As a founding member of the interdepartmental graduate program in Bioinformatics and Computational Biology (BCB), I worked with an interdisciplinary team of biologists to secure an Integrative Graduate Education and Research Training (IGERT) award which helped establish one of the first (and currently one of the largest, and perhaps one of the strongest Bioinformatics Ph.D. programs in the United States. I have led the development of a set of 4 core courses in Bioinformatics and Computational Biology covering Genome Informatics, Structural Genome Informatics, Functional Genomics, and Systems Biology. I have developed and taught short course modules as part of an NSF-supported Bioinformatics Summer Institute in Bioinformatics and Computational Systems Biology. I have also contributed to the design of a new undergraduate curriculum in Bioinformatics and Computational Biology, to be offered beginning in fall 2007.

Teaching Interests

I have a strong interest in teaching undergraduate and graduate courses in artificial intelligence, machine learning, bioinformatics and computational biology, intelligent agents and multi-agent systems, knowledge representation and inference, semantic web and related topics. I am very interested in developing and teaching courses that incorporate significant research advances in the relevant disciplines and prepare students to address new research challenges (e.g. information integration, e-science) in computer science, bioinformatics, and related areas.

I am also interested in offering focused special topics or seminar courses aimed at students in the undergraduate honors program, and graduate students on topics of current interest.

Given an opportunity, I would especially enjoy developing and teaching a course that introduces computer science as an intellectual discipline to a broad audience of undergraduates (and perhaps graduate students from other disciplines) drawn from across all areas of science, engineering, and the humanities.

I would also enjoy contributing to the development of new graduate and undergraduate programs designed to train a new generation of scientists well-versed in computer and information sciences and the creative and skillful application of information processing approaches to address fundamental scientific problems in biological, agricultural, cognitive, environmental, health, physical and social sciences.

OTHER PROFESSIONAL ACTIVITIES

(Not including memberships in editorships of journals and service on standing review panels and study sections e.g., NIH)

ADVISORY BOARDS AND WORKING GROUPS

2005-	External Advisory Committee, NSF HBCU-UP Program, Tuskegee University.
1999-03	Member of Advisory Board, Emergent Computational Neural Network Architectures, Universities of York, Edinburgh, and Sunderland, United Kingdom
2004	Member of Bioinformatics Group, National Science Foundation Intelligent Data Management PI Workshop, Boston, MA
2003	Member of Information Integration Working Group, National Science Foundation Intelligent Data Management PI Workshop, Seattle, WA.
1999	Member of Intelligent Agents Working Group, Information Institute, Information Directorate, Air Force Research Labs, Rome, New York
1999	Member of Search Committee, Senior Science and Technology Position in Defensive Information Warfare, Air Force Research Laboratory, Rome, New York.
2000	Reviewer, University of Georgia Board of Regents, Proposal for a Ph.D. Program in Artificial Intelligence.
1999	Member, Intelligent Agents Working Group, Information Institute, Information Directorate, Air Force Research Labs
1998	Member, Intelligent Knowledge-Based Systems Working Group, Information Institute, Information Directorate, Air Force Research Labs
1997	Invited Participant, NSF Workshop on Decision Based Design, Sacramento, CA

CONFERENCE OR WORKSHOP PROGRAM CHAIR

2008	Track Chair	SIAM Conference on Data Mining
2007	Organizer and Co-Chair	IJCAI Workshop on Semantic Web for Collaborative Knowledge Acquisition (SWeCKa 2007), International Joint Conference on Artificial Intelligence, Hyderabad, India
2006	Organizer and Co-Chair	First International Workshop on Modular Ontologies, International Semantic Web Conference, Athens, GA
2006	Organizer and Co-Chair	AAAI Fall Symposium on Semantic Web for Collaborative Knowledge Acquisition (SWeCKa 2006), Washington, DC.
2005	Organizer and Chair	IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Information Sources, IEEE Conference on Data Mining, Houston, Texas
2004	Program Co-Chair	International Conference on Intelligent Information Processing, Beijing, China
2003	Organizer and Chair	Computational Intelligence Workshop John Vincent Atanasoff Symposium on Advanced Computing, Iowa State University, Ames, Iowa
2002	Program Chair	Conference on Computational Biology and Genome Informatics, Durham, North Carolina
2002	Track Chair	Artificial Life, Agents, and Adaptive Behavior Genetic and Evolutionary Computing Conference, New York
2001	Organizer and Chair	Workshop on Knowledge Discovery from Heterogeneous, Distributed, Dynamic, Autonomous Data and Knowledge Sources. International Joint Conference on Artificial Intelligence, Seattle.
2001	Organizer and Co-Chair	Workshop on Cognitive Agents and Inter-agent Interaction, International Conference on Cognitive Science, Beijing

2000	Organizer and Co-Chair	Workshop on Learning from Sequential and Temporal Data, International Conference on Machine Learning, Palo Alto, CA
1999	Organizer and Co-Chair	Workshop on Computation with Neural Systems National Conference on Artificial Intelligence (AAAI), Orlando
1999	Track Chair	Artificial Life, Agents, and Adaptive Behavior Genetic and Evolutionary Computing Conference, Orlando
1998	Program Chair	International Colloquium on Grammatical Inference Ames, Iowa
1997	Organizer and Co-Chair	Workshop on Automata Induction, Grammatical Inference, and Language Acquisition International Conference on Machine Learning, Nashville
1992	Organizer and Chair	Workshop on Symbolic and Subsymbolic Information Processing, Neural Circuits and Systems Conference on Neural Information Processing, Vail, CO

Conference Advisory Committee Membership

2007	Advisory Board Member	Computational Structural Bioinformatics Workshop, BIBE 2007
2004	Advisory Board Member	International Conference on Intelligent Knowledge Systems, Turkey
1997	Advisory Board Member	International Conference on Evolutionary Computation
1997	Advisory Committee Member	International Conference on Artificial Intelligence Applications (ICAIA '97), Cairo, Egypt. 1997.

Conference or Workshop Program Committee Membership

2008	International Conference on Machine Learning (ICML 2008)
2008	ACM/WIC/IEEE Conference on Intelligent Agent Technology (IAT 2008)
2008	SIAM Conference on Data Mining (SDM 2008)
2008	International Colloquium on Grammatical Inference (ICGI 2008)
2008	International Symposium on Bioinformatics Research and Applications (ISBRA 2008)
2007	ACM SIGKDD Conference on Data Mining and Knowledge Discovery (KDD 2007)
2007	Intelligent Systems in Molecular Biology (ISMB 2007)
2007	International Colloquium on Grammatical Inference (ICGI 2007)
2007	ACM/WIC/IEEE Conference on Intelligent Agent Technology (IAT 2007)
2007	IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2007)
2007	AAAI Workshop on Semantic e-Science (SeS 2007)
2007	2nd International Workshop on Modular Ontologies (WoMO 2007)
2007	ICWS Workshop on Service Composition and Adaptation (WSCA 2007)
2007	IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2007)
2007	IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2007)
2006	International Conference on Machine Learning (ICML 2006)
2006	AAAI Fall Symposium on Semantic Web for Collaborative Knowledge Acquisition (SWeCKa 2006)
2006	First International Workshop on Modular Ontologies (WoMO 2006)
2006	IEEE Conference on Tools with Artificial Intelligence (ICTAI 2006)
2006	International Colloquium on Grammatical Inference (ICGI 2006)
2006	ACM / IEEE / WIC Conference on Web Intelligence (WI 2006)
2006	International Workshop on Algorithms in Bioinformatics (WABI 2006)
2006	8th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 06)
2006	IASTED International Conference on Computational Intelligence (CI 2006)
2006	IEEE Conference on Granular Computing (IEEE-GrC 2006)
2005	IEEE International Conference on Data Mining (ICDM 2005)
2005	International Conference on Algorithmic Learning Theory (ALT 2005)

2005	IEEE Conference on Tools with Artificial Intelligence (ICTAI 2005)
2005	International Conference on Machine Learning (ICML 2005)
2005	IEEE/ACM Conference on Intelligent Agent Technology (IAT 2005)
2005	ACM SIGKDD Workshop on Data Mining in Bioinformatics (BIOKDD 2005)
2005	Indian International Conference on Artificial Intelligence (IICAI 2005)
2004	IEEE International Conference on Data Mining (ICDM 2004)
2004	International Conference on Machine Learning (ICML 2004)
2004	IEEE/WIC/ACM Conference on Intelligent Agent Technology (IAT 2004)
2004	AAAI Workshop on Semantic Web Personalization
2004	IEEE Conference on Bioinformatics and Bioengineering (BIBE 2004)
2004	International Colloquium on Grammatical Inference (ICGI 2004)
2004	SIAM Bioinformatics Workshop, SIAM International Conference on Data Mining (SDM 04)
2004	IEEE Conference on Tools with Artificial Intelligence (ICTAI 2004)
2004	Midwestern Conference on Artificial Intelligence and Cognitive Science (MAICS 2004)
2003	IEEE International Conference on Data Mining (ICDM 2003)
2003	IEEE Conference on Tools with Artificial Intelligence (ICTAI 2003)
2003	International Conference on Intelligent Systems Design and Applications (ISDA 2003)
2003	Midwestern Conference on Artificial Intelligence and Cognitive Science (MAICS 2003)
2002	International Colloquium on Grammatical Inference (ICGI 2002)
2002	International Conference on Hybrid Intelligent Systems (HIS 2002)
2002	International Conference on Intelligent Systems Design and Applications (ISDA 2002)
2002	Midwestern Conference on Artificial Intelligence and Cognitive Science (MAICS 2002)
2002	Network Applications in Bioinformatics (NETTAB) Workshop at Bologna, Italy, 2002.
2002	Fourth International Workshop in Frontiers of Evolutionary Computation (FEA-2002)
2001	International Conference on Machine Learning (ICML 2001)
2001	International Symposium on Artificial Intelligence, India (ISAI 2001)
2001	International Joint Conference on Neural Networks (IJCNN 2001)
2001	International Conference on Hybrid Intelligent Systems (HIS 2001)
2001	SIAM Workshop on Mining Scientific Data Sets (at SDM 2001)
2001	Midwestern Conference on Artificial Intelligence and Cognitive Science (MAICS 2001)
2001	Genetic and Evolutionary Computing Conference (GECCO 2001)
2001	Workshop on Gene Expression. Genetic and Evolutionary Computing Conference, 2001
2000	International Colloquium on Grammatical Inference (ICGI 2000)
2000	Genetic and Evolutionary Computing Conference (GECCO 2000)
2000	Midwestern Conference on Artificial Intelligence and Cognitive Science (MAICS 2000)
1999	National Conference on Artificial Intelligence (AAAI 1999)
1999	Midwestern Conference on Artificial Intelligence and Cognitive Science (MAICS 1999)
1998	International Colloquium on Grammatical Inference (ICGI 1998)
1998	Genetic Programming Conference (GP 1998)
1997	International Conference on Machine Learning (ICML 1997)
1997	Genetic Programming Conference (GP 1997)
1996	Genetic Programming Conference (GP 1996)
1997	Midwest Artificial Intelligence and Cognitive Science Conference (MAICS 1997)
1996	World Congress on Neural Networks (WCNN 1996)
1996	Midwest Artificial Intelligence and Cognitive Science Conference (MAICS 1996)
1995	World Congress on Neural Networks (WCNN 1995)
1995	Midwest Artificial Intelligence and Cognitive Science Conference (MAICS 1995)
1993	International Simulation Technology Conference (SIMTEC 1993)
1993	University of New Brunswick Artificial Intelligence Symposium
1992	International Simulation Technology Conference (SIMTEC 1992)

Journal Referee

Applied Intelligence
Bioinformatics
BMC Bioinformatics
BMC Cancer
Connection Science
Genetic Programming
IEEE Computer
IEEE Expert
IEEE Intelligent Systems
IEEE Transactions on Data and Knowledge Engineering
IEEE Transactions on Evolutionary Computation
IEEE Transactions on Neural Networks
IEEE Transactions on Pattern Analysis and Machine Intelligence
IEEE Transactions on System, Man, and Cybernetics
International Journal of Data Mining and Bioinformatics
Information and Computation
Information Fusion
Information Sciences
Journal of Computational Biology
Journal of Machine Learning Research
Nucleic Acids Research
Neural Computation
Neural Networks
Machine Learning
Pattern Recognition
Proteins Structure, Function and Bioinformatics

Proposal Reviewer

National Science Foundation, USA
National Institutes of Health, USA
US Civilian Research Development Foundation, USA
United States Department of Agriculture, USA
Dutch National Science Foundation, Netherlands
European Physical Science and Engineering Research Council, United Kingdom
Irish National Science Foundation, Ireland
National Science and Engineering Research Council, Canada

Other Professional Service

Book and/or journal proposals for MIT Press, Academic Press, CRC Press, and Springer Verlag.
External Referee, Promotion and Tenure.

Departmental, College, and University Service

2007-2008	Member, Research Computing Council		ISU
2007-2008	Member, Systems Biology Steering Committee		ISU
2007-2008	Member, Graduate Committee	Computer Science	ISU
2007-2008	Member, Promotion & Tenure Steering Committee	Computer Science	ISU
2007-2008	Member, Graduate Admissions Committee	Computer Science	ISU
2007-2008	Member, Faculty Search Committee	Computer Science	ISU

2007-2008	Member, Supervisory Committee, Bioinformatics and Computational Biology Graduate Program		ISU
2007-2008	Member, Professional and Scientific Staff Hiring Committee (ad hoc)	Computer Science	ISU
2006-2007	Member, Planning Committee, Information Science and Technology Institute		ISU
2006-2007	Member, Graduate Committee	Computer Science	ISU
2006-2007	Chair, Research Infrastructure Committee	Computer Science	ISU
2005-2006	Member, Graduate Committee	Computer Science	ISU
2005-2006	Chair, Graduate Admissions Committee	Computer Science	ISU
2005-2006	Member, Departmental Web Committee	Computer Science	ISU
2005-2007	Ex Officio Member, Supervisory Committee, Bioinformatics & Computational Biology Graduate Program		ISU
2005-2007	Member, Planning Committee, Bioinformatics Undergraduate Program		ISU
2005-2006	Member, Promotion and Tenure Steering Committee	Computer Science	ISU
2004-2005	Member, Graduate Committee	Computer Science	ISU
2004-2005	Committee for review of the Office of the Vice President of Business & Finance		ISU
2003-2004	Member, ISU Information Technology Working Group: Research		ISU
2003-2004	Member, Graduate Admissions Committee	Computer Science	ISU
2002-2003	Member, Graduate Admissions Committee	Computer Science	ISU
2003-2004	Member, Provost's Information Technology Advisory Group		ISU
2002-2003	Member, Human Computer Interaction Graduate Program Steering Committee		ISU
2000-2001	Member, Department Chair Search Committee	Computer Science	ISU
1999-	Member, Supervisory Committee, Interdepartmental Bioinformatics & Computational Biology Graduate Program		ISU
1999	Coordinator, Complex Adaptive Systems Workshop		ISU
1999-2002	Member, Supervisory Committee, Complex Adaptive Systems Graduate Minor		ISU
1998-1999	Member, Graduate Committee	Computer Science	ISU
1998	Departmental Delegate, Annual Faculty Conference	Computer Science	ISU
1998	Member, Advisory Committee to Department Chair, Presidential Taskforce on Information Technology	Computer Science	ISU
1997-1998	Member, Graduate Committee	Computer Science	ISU
1996-1997	Member, Carver Trust Grant Proposal Review Committee		ISU
1997-1999	Member, Bioinformatics & Computational Biology Program Steering Committee		ISU
1996-	Member, Promotion and Tenure Committee	Computer Science	ISU
1996-1997	Member, Graduate Committee	Computer Science	ISU
1996-1997	Coordinator, Graduate Student Orientation	Computer Science	ISU
1996	Member, Industry Day Organization Committee	Computer Science	ISU
1996	Advisor, Iowa State University Student Team, AAAI Robot Competition		ISU
1996-1997	Member, Carver Trust Grant Proposal Review Committee		ISU
1995-1996	Faculty Secretary	Computer Science	ISU
1995-1999	Member, Interdepartmental Neuroscience Graduate Program Supervisory Committee		ISU
1994-1996	Member, Departmental Strategic Planning Committee	Computer Science	ISU
1990-1996	Member, Graduate Admissions Committee	Computer Science	ISU
1992-1994	Member, Liberal Arts and Sciences Honors Program Committee		ISU

REFERENCES AVAILABLE ON REQUEST