VASANT G. HONAVAR

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

honavar@cs.iastate.edu
515-294-1098
http://www.cs.iastate.edu/~honavar/

EDUCATION

University of Wisconsin-Madison, USA	Computer Science and Cognitive 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 Biology, Discovery Informatics, Big Data Analytics, Data Mining, Health Informatics, Social Informatics, Knowledge Representation and Inference, Semantic Web.

POSITIONS HELD

2010-present	Program Director, Information and Intelligent Systems Division	National Science Foundation
2001-present	Professor of Computer Science	Iowa State University
2005-present	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
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 1990-2010 have been Research (50%), Teaching (50%), and Service. During 2010-2013, I have been on an IPA assignment as a Program Director at NSF while maintaining my research activities at ISU.

- **National Science Foundation: Program Director in the Information and Intelligent Systems Division of the Computer and Information Sciences and Engineering Directorate** (2010-2012)
    - Lead Program Director for the Big Data Science and Engineering Program (Midscale Competition)
    - Program Director, Information Integration and Informatics Program (especially Data Mining and Machine Learning, Social Media and Social Informatics, and Semantic Technologies Portfolios).
    - Program Director for two Expeditions in Computing Awards
    - Additional Responsibilities for CAREER, Smart Health and Wellbeing, Software for Sustained Innovation, Computing Research Infrastructure and Expeditions in Computing, Integrative Graduate Education and Research Training (IGERT) and Science of Learning Centers programs.
    - Research Planning Workshops in Discovery Informatics (organized by Yolanda Gil and Haym Hirsh), Population Health Measurement and Analysis (organized by Bruce Schatz), Finance Informatics (organized by Louiqa Raschid), Research Challenges and Opportunities in Knowledge Representation (organized by Natasha Noy and Deborah McGuinness).

- **Iowa State University: Professor of Computer Science and of Bioinformatics and Computational Biology** (1990-Present)
    - Research contributions in Machine learning, Data Mining, Knowledge Representation and Semantic Web and interdisciplinary contributions in Bioinformatics and Computational Biology reported in 240 refereed journal and conference papers including two receiving best paper awards, 20 book chapters, and 1 research monograph during 1990-2012.
    - Founded (in 2005) and served as the director of 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.
    - Secured approximately $20 million in research and training grants as PI or co-PI during 1990-2012; Currently serving as PI or Co-PI on approximately $4 million in research and training grants.
    - Mentored 28 Ph.D. graduates (8 employed as tenure-track or tenured faculty in academia, 6 employed in academic research, and 10 employed in industrial research and development), 25 MS graduates (24 employed in industry, one in academic research); Currently mentoring or jointly mentoring 12 Ph.D. students.
    - Developed the "Discovery Informatics" initiative at Iowa State University, aimed at the creation of several faculty positions in Computer Science and several departments in mathematical, physical, and biological sciences aimed at building up a critical mass of research expertise in Discovery Informatics.
    - Established, 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 served as associate chair of the program during 2001-2003 and chair during 2003-2005.
    - Lead 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.
    - Established, 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.
    - Served 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.

- Established, and over the years, refined and broadened graduate and undergraduate curricula in Artificial Intelligence and Machine Learning at Iowa State University.
- Served on editorial boards of several journals, organized several scientific conferences and workshops, and served on the scientific program committees of major conferences in artificial intelligence, machine learning, data mining, semantic web, and bioinformatics.

BIOGRAPHICAL SUMMARY

Biography: Dr. Vasant Honavar received his Ph.D. in Computer Science and Cognitive Science in 1990 from the University of Wisconsin Madison, specializing in Artificial Intelligence. Since 1990, he has been on the faculty of Iowa State University (ISU) where he is currently a Professor of Computer Science. He directs the Artificial Intelligence Research Laboratory (which he founded in 1990) and the Center for Computational Intelligence, Learning & Discovery (which he founded in 2005). Honavar is also a member 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 ISU Bioinformatics and Computational Biology Graduate Program, which he helped establish in 1999 with support from an Integrative Graduate Education and Research Training (IGERT) award. Honavar is currently on an Intergovernmental Personnel Act (IPA) Assignment as a Program Director in the Information and Intelligent Systems Division at the National Science Foundation where his programmatic responsibilities have included the Big Data Science and Engineering, Information Integration and Informatics, Smart Health and Wellbeing programs, among others.

Honavar has over 20 years of research experience in Artificial Intelligence. His research has resulted in foundational contributions in machine learning, bioinformatics and computational biology, and knowledge representation and inference. He has published over 200 research articles in peer-reviewed journals and conferences (including several that have received best paper awards), and several invited book chapters (Based on Google Scholar citations data, as of December 2012, Honavar's *h*-index[1], g-index and *i*10-index are 38, 69, and 153 respectively). This research has been supported in part by grants from the National Science Foundation, the National Institutes of Health, the US Department of Agriculture, and the US Department of Defense. Honavar has extensive curriculum development and teaching experience in Artificial Intelligence, Machine Learning, and Bioinformatics. He also has substantial industrial consulting experience in Data Mining, Bioinformatics, and related topics.

Honavar has served on, or currently serves on the editorial boards of Cognitive Systems Research, Machine Learning, 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, and Webmed Central. Honavar has served on the program committees of major research conferences in artificial intelligence, data mining, and bioinformatics including the Conference on Artificial Intelligence (AAAI), 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), Intelligent Systems in Molecular Biology (ISMB), ACM Conference on Bioinformatics and Computational Biology (ACM-BCB), among others. Honavar has served as a charter member of the National Institutes of Health study section on Biological Data Management and Analysis (2002-2007). Honavar is a senior member of the Association for Computing Machinery (ACM), and of the Institute of Electrical and Electronic Engineers (IEEE) and a member of the Association for Advancement of Artificial Intelligence (AAAI), International Society for Computational Biology (ISCB) Society for Industrial and Applied Mathematics, and the American Association for the Advancement of Science. He currently serves on the Board of Directors for ACM Special Interest Group on Bioinformatics.

Honavar has received several awards during his career including the National Science Foundation Director's Award for Collaborative Integration in 2012, the Margaret Ellen White Graduate Faculty Award for Excellence in Student Mentoring in 2011, the Iowa Board of Regents Award for Faculty Excellence in 2007 and the Iowa State University College of Liberal Arts and Sciences Award for Research Excellence in 2008, and two best paper awards at major conferences. However, his proudest accomplishments are the 28 PhD students, 25 MS Students and several undergraduate researchers that he has worked with and mentored during his career.

---

[1] The *h*-index of an author is *n* if the author has *n* publications with at least *n* citations each. The *g*-index of an author is *m* if

2012	National Science Foundation Director's Award for Collaborative Integration
2011	Margaret Ellen White Graduate Faculty Award, Iowa State University
2011	Best poster award, ACM Conference on Bioinformatics and Computational Biology (ACM-BCB 2011)
2008	Liberal Arts and Sciences Award for Excellence in Research, Iowa State University
2010	Elected Fellow, International Society for Intelligent Biological Medicine
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

EDITORIAL BOARD MEMBERSHIP

2012-	Associate Editor	IEEE/ACM Transactions on Bioinformatics and Computational Biology
2013	Guest Associate Editor	PLOS Computational Biology
2012-	Specialty Editor	Bioinformatics, Webmedcentral
2013-	Member of Editorial Board	Journal of Proteomics & Computational Biology
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 Editorial Board	International Journal of Functional Informatics and Personalized Medicine
2007-2008	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 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
2005-2008	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

2010	Member of Special Emphasis Study Section, National Center for Research Resources	National Institutes of Health
2010	Reviewer, National Centers for Biomedical Computing	National Institutes of Health
2010	Member of Study Section, Biological Data Management and Analysis	National Institutes of Health
2010	Reviewer	Israeli Science Foundation
2010	CISE Review Panels	National Science Foundation
2009	Member of Study Section, Bioinformatics Resource Centers for	National Institutes of Health

	Infectious Diseases	
2009	Member of Genomics, Bioinformatics and Systems Biology Review Panel	European Research Council Genetics, and Systems Biology
2009	Member of Review Panel	Chaires de Recherché du Canada
2009	CISE Review Panel	National Science Foundation
2008	CISE Review Panel	National Science Foundation
2008	Chair of Study Section, Data Ontologies and Sharing Data and Tools	National Institutes of Health
2004-2007	Member of Study Section, Biological Data Management & Analysis	National Institutes of Health
2006	CISE Review Panel	National Science Foundation
2005	CISE 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

PROFESSIONAL AFFILIATIONS

2010-	Board of Directors, ACM Special Interest Group on Bioinformatics and Computational Biology (ACM-BCB)
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

2012	Program Director	Big Data Science and Engineering Program		NSF
2011-2012	Program Director	Expeditions in Computing: Understanding Climate Change: A Data Driven Approach		NSF
2012	Program Director	Expeditions in Computing: Making Sense at Scale with Algorithms, Machines and People		NSF
2010-2012	Program Director	Information Integration and Informatics Program		NSF
2011-2012	Program Director	Smart Health and Wellbeing Program		NSF
2010	Program Director	Computing Research Infrastructure Program		NSF
2011-2012	Technical Coordinator	Pittsburgh Science of Learning Center		NSF
2005-2012	Director	Center for Computational Intelligence, Learning, and Discovery		Iowa State Univ.
2009-2010	Chair	Promotion and Tenure	Computer Science	Iowa State Univ.

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

CONSULTING EXPERIENCE

Scientific consulting on data mining, artificial intelligence, bioinformatics, semantic technologies, knowledge-based systems, information integration, 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 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) How can we build useful predictive models from large, distributed, autonomous data sources (Big Data Analytics)?

(b) How can we build predictive models from semantically disparate data?

(c) How can we integrate data, hypothesis, and knowledge-based inference, predictive modeling, experimentation, simulation, and hypothesis testing into an exploratory apparatus for scientific discovery (Discovery Informatics)?

(d) How can we extract useful knowledge from richly structured data (sequences, images, graphs, etc.)?

(e) How can we efficiently represent and reason about preferences?

(f) How can we query and reason with federated data and knowledge bases?

(g) How can we answer queries from knowledge bases that contain secrets without revealing secrets?

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

(i) How can we construct, compare, analyze multi-scale, predictive models of molecular networks involved in cellular development, differentiation, immune response, and biological function?

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

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

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

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

(n) How can we build robust intelligent agents that incorporate multiple facets of intelligence?

My research contributions have spanned Computer Science (especially on the topics of Machine Learning and Data Mining, Knowledge Representation and Inference, Service-oriented computing) and in Bioinformatics and Computational Biology (especially on the topic of analysis and prediction of biomolecular (protein-protein, protein-DNA, and protein-RNA) interfaces and comparative analysis of biomolecular interaction networks). Some of my most recent work has focused on (1) Development of novel statistical query based approach to building predictive models from large distributed data sets in settings where centralized, in memory access to data is not feasible because of memory, access or bandwidth constraints; (2) Secrecy-preserving approaches to answering queries against knowledge bases that contain secrets using inferences that make use of secrets whenever it is possible to do so without revealing secrets; (3) Techniques for representing and reasoning about qualitative, possibly conditional preferences over a set of alternatives; (4) Algorithms for comparative analysis of large bio-molecular networks and (5) Machine learning approaches to analysis and prediction of macromolecular interactions and interfaces, including in particular, the first algorithm for partner-specific prediction of protein-protein interface sites.

Over the next 5 years, I plan to focus my research in Computer Science on (1) Discovery Informatics, especially novel approaches to integrating data, hypothesis, and knowledge-based inference, predictive modeling, experimentation, simulation, and hypothesis testing in large-scale scientific discovery; (2) Scalable machine learning approaches to predictive modeling from very large, richly structured data sets (including sequences, graphs, relational and RDF data) (3) Analysis and prediction of macromolecular interactions, elucidation of complex biological pathways e.g., those involved in immune response, development, and disease and (4) Applications in health informatics.

Current Research Interests

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

(b) **Bioinformatics and Computational Molecular and Systems 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, comparative analysis of biological networks (network alignment), integrative analysis of molecular interaction networks and macro-molecular interfaces.

(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), grammars; 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, multiple instance data, multiple instance, multiple class

data; networked data; multi-relational data, linked open data (RDF), and semantically heterogeneous data; Incremental learning, Ensemble methods, multi-agent learning, curriculum-based learning; selected topics in computational learning theory.

(e) **Discovery Informatics:** Computational models of scientific discovery; Discovery informatics infrastructure to integrate data, hypothesis, and knowledge-based inference, predictive modeling, experimentation, simulation, and hypothesis testing to provide an orderly formal framework and exploratory apparatus for science; Applications in computational systems biology.

(f) **Knowledge Representation and Semantic Web:** Ontology-based and probabilistic approaches to information integration; description logics, federated data bases – statistical queries against federated databases, knowledge bases – federated reasoning, selective knowledge sharing, services – service composition, substitution, and adaptation; epistemic description logics; secrecy-preserving query answering, representing and reasoning about qualitative preferences.

(g) **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[2]

1. Collaborative Research: Learning Classifiers from Autonomous, Semantically Heterogeneous Distributed Data Sources, National Science Foundation, Vasant Honavar (PI) $527,000 (including supplements). 2007-2013.

2. Intergovernmental Personnel Act (IPA) Appointment, Program Director, Information Integration and Informatics Program, Information and Intelligent Systems Division, Computer and Information Science and Engineering Directorate, National Science Foundation, $725,479. 2010-2013

3. Secrecy-Preserving Reasoning: Foundations, Algorithms, and Software**,** Giora Slutzki (PI), Vasant Honavar (Collaborator) National Science Foundation**,** $450,895. 2011-2014.

4. Decision Support System for Reasoning About Preferences. Samik Basu (PI), Vasant Honavar (Collaborator) National Science Foundation, $111,393, 2011-2013.

5. Identifying porcine genes and gene networks involved in effective response to PRRS virus using functional genomics and systems biology, Joan Lunney (PI), Vasant Honavar (Co-PI), Zhihua Jiang (Co-PI), Roman Pogranichniy (Co-PI), Juan Pedro Steibel (Co-PI), Chris Tuggle (Co-PI), United States Department of Agriculture, 2010-2013, $750,000.

Pending Grant Proposals

1. Learning Predictive Models from Linked Open Data. Vasant Honavar (PI), National Science Foundation, $500,000. Under review.

2. SHF: Small: Preference Reasoning for Software Design, Testing, and Cyberdefense. Samik Basu (PI), Vasant Honavar (Co-PIs), National Science Foundation, $500,000. Under review.

3. Integrative Analysis and Prediction of Protein-Protein Interaction Networks and Protein-Protein Interfaces. Vasant Honavar (PI), Robert Jernigan, Drena Dobbs (Co-PIs), National Institutes of Health, $1,500,000.

4. Aligning Macromolecular Interaction Networks: Efficient Algorithms, Implementation, and Applications, Vasant Honavar (PI), Robert Jernigan, Giora Slutzki (Co-PIs), Department of Energy, $1,000,000. Under revision.

5. Sequence and Structural Correlates of Protein-RNA Interactions. Vasant Honavar (PI) with Drena Dobbs (PI) and Robert Jernigan (Co-PI). National Institutes of Health. (2010-2013), $1,500,000. Under revision.

---

[2] All the NSF grants that were active when I began my IPA appointment at NSF have been assigned substitute PIs to manage them during my tenure at NSF.

Past Research and Training Grants

1. IGERT: Computational Molecular Biology Training Program. Vasant Honavar (Co-PI) with Dan Voytas (PI), Drena Dobbs (Co-PIs). National Science Foundation, 2005-2012. $2,968,976.

2. 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.

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). (2007-2011), $350,002.

4. Machine Learning Algorithms and Software for Collaborative Medical Decision Support. Grow Iowa Values Fund, and Collaborative Health Solutions, LLC 2010-2012, $238,586

5. Developing Predictive Models for Identifying Pigs with Superior Immune Response and Improved Food Safety, Chris Tuggle (PI), Bearson, S., Honavar, V., Nettleton, D. Wannemuehler, M., Lunney, J. and Nettleton, D. (Co-PIs). United States Department of Agriculture (2009-2012), $1,000,000.

6. 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--2010). $499,000.

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). (2006-2011) $700,000.

9. 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.

10. IIS: Exploratory Investigation of Modular Ontologies. *National Science Foundation,* Vasant Honavar (PI), Giora Slutzki and Doina Caragea (Co-PIs), (2006-2008). $112,000.

11. 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.

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. Center for Computational Intelligence, Learning, and Discovery. Vasant Honavar (PI). Vice Provost for Research, Iowa State University (2004-2009) $477,500

14. 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.

15. Algorithms and Software for Collaborative Ontology Development. Center for Integrated Animal Genomics, Iowa State University. Vasant Honavar (PI), (2005-2007) $25,000.

16. 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).

17. 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.

18. 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.

19. 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.

20. Symposium 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.

21. 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.

22. Constructive Neural Network Learning Algorithms for Pattern Classification, National Science Foundation, 1994-1999, Principal Investigator, $111,537 (plus $10,000 in matching funds).

23. 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.

24. Graduate Research Fellowships in Bioinformatics and Computational Biology, Pioneer Hi-Bred, Inc. 2002-2004. Major professor, Adrian Silvesu and Carson Andorf, $80,000.

25. IBM Graduate Research Fellowship in Computer Science, IBM Inc., Major Professor, Doina Caragea, 2003-2004, $30,000 (approx.)

26. 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.

27. 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.

28. 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.

29. IBM Graduate Research Fellowship in Computer Science, IBM Inc., Vasant Honavar (Major Professor), with Doina Caragea (doctoral student), 2002-2003, $30,000 (approx.)

30. Intelligent Multi-Agent Systems for Intrusion Detection, National Security Agency, 1998-2000, Co-Principal Investigator, $199,769.

31. Data Mining of Electric Power Usage Data to Develop Customer Profiles. Cooperative Research Proposal. Power Domain, Inc. (2001-2002). Vasant Honavar. $43,639.

32. SGER: Distributed Knowledge Networks to Support Security-Economy Decisions in Stressed Electric Power Systems. National Science Foundation, 2000-2001, Co-Principal Investigator, $99,999.

33. Artificial Intelligence Applications to Power System Management and Control, Electric Power Research Institute, Co-principal investigator. 1998-2000, $151,000

34. Distributed Knowledge Networks, John Deere Foundation, 1999-2001, Principal Investigator, $30,000.

35. Development of Algorithmic Approaches to Gene Expression Analysis from Microarray Data, *Carver Foundation*, 2000-2001, Principal Investigator, $25,000.

36. Development of Protein Structure Prediction Algorithms. Carver Foundation, 1999-2000, Co-Principal Investigator, $25,000.

37. Genetic algorithms for protein structure prediction. Ames Laboratory, Co-Principal Investigator, 1999-2000. $35,700.

38. A Gene-Specific DNA Chip for Exploring Molecular Evolutionary Change, *Carver Foundation*, 1998-1999. Co-Principal Investigator, $17,120.

39. Intelligent Diagnosis Systems, John Deere Foundation, 1995-1998, Principal Investigator, $30,000.

40. Graduate Fellowship (Data Mining and Knowledge Discovery), IBM Corporation, 1997-1998, Principal Investigator, $20,800.

Algorithms and Software for Knowledge Acquisition from Semantically Heterogeneous, Distributed Data (funded in part by grants from the National Science Foundation)

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) Development of a scalable statistical query based approach to learning and updating sequence classifiers from very large sequence data sets (Koul et al., 2010).

(d) 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.

(e) *Abstraction-Driven Algorithms for Building Compact yet Accurate Classifiers.* We have developed a general approach for exploiting attribute value hierarchies (AVH) that group the values of attributes to learn compact yet accurate predictive models from data specified at different levels of abstraction. Instantiations of this approach in the case of Naïve Bayes (Zhang et al., 2004; 2006) and Decision Trees (Zhang et al., 2003) show that the resulting algorithms yield

predictive models that are more compact than those produced by their counterparts that do not have access to AVH without sacrificing the quality of the predictors.

(f) Development of a statistical query based approach to learning classifiers from semantically disparate multi-relational data (Caragea et al., 2010).

(g) Demonstration of the theoretical equivalence of a certain class of inter-ontology mapping errors and noise models, and hence the reduction of the problem of learning in the presence of mapping errors from semantically disparate data to the problem of learning from noisy data (Koul et al., 2010, 2012).

(h) 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 (http://code.google.com/p/induslearningframework/; http://code.google.com/p/indusintegrationframework/).

Research in progress is aimed at:

(a) Extension of the statistical query based learning framework to learning predictive models from Linked Open (RDF) Data e.g., algorithms for learning Relational Bayesian Classifiers from RDF data in settings where the learner can access the RDF data only through a restricted set of queries against an access interface (Lin et al., 2011).

(b) Extension of the statistical query based learning framework to learning predictive models from network data

(c) Applications of the resulting algorithms to social network and social media analytics and analysis and prediction of biomolecular interactions

Learning Predictive Models from Richly Structured Data

Many applications e.g., biomolecular sequence analysis, image classification, text classification, social network analysis, among others require methods for classification of structured data. Of particular interest are topologically structured data, i.e. data whose topology reflects intrinsic dependencies between the constituent elements that make up the data. Learning from topologically or relationally structured data presents several challenges:

- Data representation: The representation of the data that is presented to a learner has to be rich enough to capture distinctions that are relevant from the standpoint of learning, but not so rich as to make the task of learning harder due to over-fitting.

- Sparsity of Labeled Data: In many applications, e.g., image annotation, sequence annotation, social network analysis, the available data (sequences, images, etc.) are only partially labeled. Hence there is a need for methods that can take advantage of vast amounts of unlabeled or partially labeled data, together with limited amounts of labeled data.

To cope with these challenges, we have developed a novel approach to learning *compact* yet *accurate* predictive models from topologically structured data. Our approach exploits the complementary strengths of *super-structuring* (constructing complex features by combining existing features) and *abstraction* (grouping similar features to generate more abstract features). Super-structuring provides a way to increase the predictive accuracy of the learned models by enriching the data representation (hence, super-structuring increases the complexity of the learned models), whereas abstraction helps reduce the number of model parameters by simplifying the data representation (Silvescu et al., 2011).

Some results of this work to date include:

(i). *Abstraction Augmented Markov models* (AAMMs). AAMMs are generalizations of Markov Models (MM). AAMMs simplify the data representation used by the standard MMs by grouping similar subsequences to organize them in an abstraction hierarchy (Caragea et al., 2010a, 2010b, 2010c). Experimental results on text document classification and protein subcellular localization show that adapting data representation by combining super-structuring and abstraction makes it possible to construct predictive models that use substantially smaller number of features (by one to three orders of magnitude) than those obtained using super-structuring alone (whose size grows exponentially with the length of direct dependencies). Super-structuring and abstraction-based models are competitive with and, in some cases, outperform, models that use only super-structuring. Our experiments have also demonstrated the promise of abstraction-augmented Markov Models in learning sequence classifiers in a semi-supervised setting where only some of the sequences are labeled.

(ii).   Development of *Abstraction-super-structuring Normal Forms* (Silvescu and Honavar, 2011) that offer a general theoretical framework for structural (as opposed to parametric) aspects of induction using abstraction (grouping of similar entities) and super-structuring (combining topologically close entities) and exploration of its relation to ideas e.g., radical positivism in the philosophy of science (with PhD student Adrian Silvescu).

(iii). Development of discriminatively trained probabilistic models for sequence classification (Yakhnenko et al., 2005), generalized multiple instance learning algorithms with applications in bioinformatics, text and image analysis (El-Manzalawy et al., 2009).

(iv).   Development multi-relational learning algorithms (Atramentov et al., 2003), relational Bayesian classifiers from RDF data, independence-based Markov Network learning algorithms (Bromberg et al., 2009), recursive Naïve Bayes learning algorithms with applications to sequence classification (Kang et al., 2006).

(v).   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 (Silvescu and Honavar, 2006).

Work in progress is aimed at extending this approach to learning predictive models from richly structured data at multiple levels of abstractions (images and text (multi-modal data), social networks and social media, linked open data, biomolecular interaction network data).

Learning Predictive Models from Multi-Modal Data

Recent years have witnessed rapid advances in our ability to acquire and store massive amounts of data across different modalities (such as text, speech, images, etc.). The availability of such data presents us with challenges of information organization, retrieval of images based on user text, identification of regions of interest in the image based on user text and so on. Scene understanding requires determining the objects that appear in the scene, their locations relative to each other, etc. Of particular interest is image annotation, the task of assigning keywords to an image based on its contents. Traditional machine learning approaches to image annotation require large amounts of labeled data. This requirement is often unrealistic, as obtaining labeled data is, in general, expensive and time consuming. In contrast, weakly or partially labeled data (e.g., tagged images) are plentiful. This presents us with the challenge of learning predictive models from partially labeled data. We have formulated the problem of image annotation as multiple instance, multiple label learning problem. Multiple instance, multiple label learning is a generalization of supervised learning in which the training examples are bags of instances and each bag is labeled with a set of labels. We have explored two learning frameworks: generative and discriminative, and proposed models within each framework to address the problem of assigning text keywords to images:

- Multimodal Hierarchical Dirichlet Process, which is a non-parametric generalization of the hierarchical mixture models. Our experimental evaluation shows that the performance of this model does not depend on the number of mixture components, unlike the standard mixture model which suffers from over-fitting (Yakhnenko and Honavar, 2009).

- A discriminative model, which given the input bag of instances what is the most likely assignment of labels to the bag. We learn as many classifiers as there are possible labels and force the classifiers to share weights using trace-norm regularization. Our experimental results on standard benchmark datasets show that the performance of this model is comparable to the state-of-the-art multiple instance multiple label classifiers and that unlike some state-of-the-art models, it is scalable and practical for datasets with a large number of training instances and possible labels (Yakhnenko and Honavar, 2010).

- A generalization of the discriminative model to a semi-supervised setting to allow the model take advantage of labeled and unlabeled data. We assume that the data lies in a low-dimensional manifold and introduce a penalty that ensures that the classifiers assign similar labels to similar instances (i.e. instances that are near-by in the manifold induced by the samples). Our experimental results show the effectiveness of this approach in learning to annotate images from partially labeled data (Yakhnenko and Honavar, 2010).

Topics in Grammatical Inference 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.

Our work on learning Regular Grammars demonstrates 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 (against the background of strong negative results regarding the feasibility of learning regular grammars within the standard PAC learning framework): "Are DFA PAC-identifiable if examples are drawn from the uniform distribution, or some other known simple distribution?"

(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).

Related work has led to the development of polynomial algorithms for learning regular languages from examples and membership queries (Nichitiu et al., 2000).

Our work on learning of grammars used to model natural languages (in particular, dependency grammars, stochastic context free grammars) has led to:

(a) The development of a novel regularization scheme, namely, *unambiguity regularization* that favors grammars that yield unambiguous parses, which includes as special cases and improves upon, standard expectation maximization (EM), Viterbi EM, and Softmax EM algorithms for unsupervised learning of grammars (Tu and Honavar, 2012).

(b) Demonstration (and explanation) of the benefits of *curricula* (e.g., a means of ordering training samples presented to the learner in an inductive learning setting) using an incremental construction hypothesis which asserts (loosely speaking) that when the target of learning is a structure e.g., a grammar that can be decomposed into a set of sub-structures e.g., grammar rules, an ideal curriculum is one that gradually emphasizes data samples that help the learner to successively discover new substructures (Tu and Honavar, 2011).

(c) An iterative bi-clustering approach to learning probabilistic context free grammars (Tu et al., 2008, 2011).

Work in progress is aimed at extending the theoretical foundations and algorithms for grammatical inference to settings that call for learning from multimodal data (e.g., combination of words and pictures). Some results to date include an algorithm for learning a multi-modal hierarchical Dirichlet process model for annotating images from partially labeled data (Yakhnenko and Honavar, 2009).

Federated Ontologies Knowledge Representation and Inference (funded in part by a grant 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 a federation of 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:

(a) 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 [Bao et al., 2006a; 2006b; 2007; 2009].

(b) Decidability and undecidability results for modular ontology languages establishing that (a) variants of distrtibuted description logics (DDL) that allow negated roles or cardinality restrictions in bridge rules or inverse bridge rules that connect *ALC* ontologies are undecidable; (b) a variant of P-DL *ALCHIO*(¬)*P* that supports role mappings between ontology modules in *ALCHIO*(¬) (an extension of *ALC* that allows general role inclusions, inverse roles, nominals and negated roles) is decidable [Bao et al., 2008].

(c) The development of distributed tableau-based reasoning algorithms for P-DL [Bao, Caragea, and Honavar, 2006c; 2007b]

(d) Tools for collaborative development of package-based partial order ontologies for applications in biomedical informatics and comparative genomics [Hughes et al., 2008].

(e) Development of a sound and complete PSPACE tableau reasoner for the Description logic ALCK, i.e., Description Logic ALC augmented with Epistemic operators to support modeling of, and inference using, knowledge of multiple agents [Tao et al., 2012].

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.

Secrecy-Preserving Inference and Query Answering (Funded in part by a grant from the National Science Foundation)
Productive interaction and collaboration among business partners, different governmental agencies (e.g., intelligence, law enforcement, public policy), or independent nations acting on matters of global concern (e.g., counter-terrorism, international finance) requires the need to share information to be balanced against the need to protect sensitive or confidential information from unintended disclosure. We aim to address this need by developing the theoretical foundations of, and algorithms and software for secrecy-preserving reasoning, that is, the process of answering queries against knowledge bases that include secret knowledge, based on inference that may use secret knowledge without revealing it. The privacy-preserving reasoning framework introduced by us in Bao et al. (2007) offered one of the first semantics-based approaches to secrecy-preserving query answering in the simple case of hierarchical knowledge bases. Work in progress is focused on:

• The development of theoretical foundations of secrecy-preserving reasoning leading to general strategies for transforming sound and complete reasoners for knowledge bases into provably secrecy-preserving reasoners that differentially tradeoff the informativeness of the reasoning against the computational overhead of secrecy preservation (relative to reasoners that are oblivious to secrecy).

• The development of secrecy-preserving reasoners for a broad class of knowledge bases of practical interest in networked information systems including hierarchical, propositional, RDF (resource description framework) and description logic (DL) knowledge bases.

• Extensions of the secrecy-preserving reasoning algorithms to settings with multiple querying agents, under various restrictions on communication between the agents.

• Modular, open source implementations and experimental evaluation of secrecy-preserving reasoners.

Representing and Reasoning About Qualitative Preferences (Funded in part by a grant from the National Science Foundation)

Many applications, e.g., composing web services, assembling complex pieces of legislation or legal arguments, designing embedded systems, etc. rely on methods for identifying collections of objects satisfying some criteria (functional specifications). Among the collections that satisfy the functional specifications, it is often necessary to identify one or more collections that are the most preferred with respect to user preferences over non-functional attributes, e.g., reliability, cost, etc. To address the needs of such applications, we have developed:

- A first practical solution to the problem of determining whether one outcome dominates another with respect to a given set of qualitative preferences over the attributes of the outcomes (Santhanam et al., 2010).
- A dominance relation that allows comparison of collections of objects in terms of preferences over attributes of the objects that make up the collection the set of most preferred collections, and algorithms for dominance testing that are guaranteed to return only (sound), all (complete), or at least one (weakly complete) of the most preferred collections (Santhanam et al., 2011).
- Applications of preference reasoning algorithms in service composition and substitution based on qualitative preferences over non-functional attributes of the services (Santhanam et al., 2008, 2009), and in minimizing credential disclosure based on qualitative preferences over sensitivity attributes of the credentials (Oster et al., 2012).

Work in progress is focused on:
- Extensions of the framework to handle reasoning with the preferences of multiple stake holders.
- Applications of preference reasoning in software design, testing, and cyberdefense.
- Applications of the resulting algorithms to problems in service composition, substitution, and adaptation.

Algorithms and Software for Interactive Discovery and Composition of Web Services (Funded in part by a grant 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. Moreover, most current approaches to service composition do not consider user preferences with regard to non-functional attributes of services (e.g., cost, performance). Against this background, our work focuses on the development of powerful interactive methods for service composition with provable guarantees with respect to user-specified functional and non-functional requirements. Some results of this research include

- Algorithms for interactive specification-driven functional assembly of composite services from a repository of available component services (Pathak et al., 2007a; 2007b; 2008)
- Algorithms for efficient identification of feasible replacement of one or more component services of a composite service while maintaining its functionality (Pathak et al., 2007)
- Algorithms for incorporating user preferences over non-functional attributes of a service (e.g., cost, security, reliability) in assembling a most preferred composition that achieves the user-specified functionality (Santhanam et al., 2008).

Comparative Analysis of Biomolecular Networks (funded in part by a grant from the USDA and in part by a National Science Foundation IGERT fellowship)

Network models are playing an increasingly important role in the interpretation of complex interactions among genes, proteins, regulatory RNAs, small ligands and other signaling agents. In particular, comparative analysis of network models of biomolecular interactions across different species or tissues has emerged as an important tool for identifying conserved

modules, predicting functions of specific genes or proteins and studying the evolution of biological processes, among other applications. Hence, there is a growing need for scalable modular, and extensible algorithms software for construction, querying, and comparative analysis of diverse types of biomolecular networks. Our work has led to:

Development of a suite of modular graph kernel based scalable and customizable algorithms and their open source implementations (as part of BiNA, http://www.cs.iastate.edu/~ftowfic/twiki/bin/view/Projects/BinaToolkit) for aligning protein-protein interaction networks and gene co-expression networks (Towfic et al., 2009). Graph kernels allow the computation of global alignment of networks to be broken down into a set of local graph kernel computations, thus contributing to the scalability of the technique.

- Application of comparative protein-protein interaction network analyses to reliably distinguish orthologs from paralogs (Towfic et al., 2010) and of comparative gene co-expression network analyses to identify B-cell ligand processing pathways (Towfic et al., 2012).

- Characterization of gene expression changes during the onset of photosynthesis (Lonosky et al., 2004), differentiation of retinal stem cells into rod photoreceptors (Hecker et al., 2010).

- Characterization differences in the proteome of murine retinal and brain derived progenitor cells (Dunn-Thomas et al., 2008).

- Development of databases and software tools for capture, analysis, annotation, and integration of gene expression data with other types of 'omics' data (Couture et al., 2009).

- Development of BioNetwork Bench, an open source, user-friendly suite of database and software tools for constructing, querying, and analyzing gene and protein network models. BioNetworkbench currently supports a broad class of gene and protein network models (eg, weighted and un-weighted, undirected graphs, multi-graphs). Bionetworkbench enables biologists to analyze public as well as private gene expression, macromolecular interaction and annotation data; interactively query gene expression datasets; integrate data from multiple networks; query multiple networks for interactions of inter¬est; store and selectively share the data as well as results of analyses. BioNetworkbench is fully interoper¬able with, Cytoscape, a popular open-source software suite for visualizing macromolecular interaction networks (Kohutyuk et al., 2012; Hecker et al., 2008).

- 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 (Silvescu and Honavar, 2001).

Work in progress is aimed at the:

- Further development of BiNA, our modular, extensible, and scalable suite of graph kernel based algorithms to enable alignment of richer networks including (a) Undirected graphs that contain multiple types of links (e.g., interaction, co-localization, etc. in the case of protein-protein interaction networks), or multiple types of nodes (e.g., in the case of macromolecular interaction networks that simultaneously model the interactions among proteins, RNA, DNA, etc.), or both; (b) Directed graphs with one or more types of links (up or down regulation of one gene by another in the case of transcriptional networks), or one or more types of nodes, or both as in the case of richly annotated signaling networks and metabolic networks; (c) The weighted counterparts of undirected (e.g., gene expression correlation networks) as well as directed graphs and (d) Undirected or directed multi-graphs with multiple links between nodes as well as variants that accommodate sets of labels on nodes (e.g., Gene ontology functional annotation, subcellular localization, etc.) and links, as well as their weighted counterparts.

- Systematic evaluation of the graph kernel based network alignment algorithms (including comparisons them with existing algorithms) on several representative applications including: identifying differences in patterns of biomolecular interactions across different species or tissues; identifying conserved modules or subnetworks; predicting functions of specific genes or proteins (and identification of functional orthologs); compensating for limited experimental data concerning biomolecular interactions within one species through transfer of information from another species on several benchmark datasets.

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

Assigning putative *functions* from sequences remains one of the most challenging problems in functional genomics. 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. Some 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, probabilistic graphical models and probabilistic language models for sequence classification; and new algorithms for learning from network data 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 and other functionally important sites (e.g., B-cell and T-cell epitopes, glycosylation and phosphorylation sites) 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.)

Some of the results to date include:

- Comprehensive Database of Protein-protein Interfaces (Jordan et al., 2012) and of Protein-RNA Interfaces (Lewis et al., 2010).

- Development of a state-of-the-art approach to predicting protein-RNA interface residues in protein sequences (Walia et al., 2012, Terribilini et al., 2006, 2007)

- Development of machine learning approaches and online servers for the prediction of protein-protein interface residues from amino acid sequence and when available, structural information (Yan et al., 2004; Towfic et al., 2011) including a state-of-the-art structure-based interface predictor (Jordan et al., 2012).

- Development of sequence homology based methods and online servers for protein interface prediction (Xue et al., 2011), including non partner-specific methods for predicting obligate interfaces and interfaces of disordered proteins and partner-specific methods for predicting transient interfaces.

- Development of sequence-based machine learning methods for predicting the approximate number of putative interaction partners of a protein (Andorf et al., 2013).

- Development of a novel approach and online server for scoring docked protein-protein complex conformations using predicted partner-specific protein-protein interfaces (Xue et al., 2011; 2012).

- 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) (Caragea et al., 2009).

- 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 (Andorf et al., 2007).

- Development of machine learning approaches and online servers for prediction of protein-DNA interface residues from amino acid sequence, and when available, structural information (Yan et al., 2006).
- Structural characterization of protein-protein and protein-RNA interfaces (Towfic et al., 2011).
- Development of machine learning methods and online servers for identification of posttranslational modification sites e.g., phosphorylation sites, glycosylation sites in amino-acid sequences (Caragea et al., 2007).
- Development of machine learning methods and online servers for predicting linear and B-cell epitopes from amino acid sequences (El-Manzalawy et al., 2008) including methods for predicting variable length and conformational B-cell epitopes.
- Demonstrations of the pitfalls of commonly used benchmark datasets for evaluating the performance of machine learning approaches to MHC-II binding site prediction (El-Manzalawy, 2008).
- Prediction of the designability of binary (H-P) protein sequences (Peto et al., 2008).
- 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 Lee et al., 2008).

The online web servers can be found at (http://www.cs.iastate.edu/~honavar/ailab/projects/proteins.html).

Biologically Inspired Algorithms for Knowledge Representation, Memory, Language Processing and Learning (Funded in part by a grant from the National Science Foundation).

Artificial neural networks, because of their potential for massive parallelism and fault and noise tolerance, offer an attractive approach to the design of associative memories, language processors, and trainable pattern classifiers. Constructive learning algorithms, which build arbitrarily complex decision boundaries needed for pattern classification (and in some ways, foreshadowed the recent development of support vector machines) were motivated by: the need to overcome the limitations of learning through parameter modification within an a priori fixed network topology; and to avoid the guesswork involved in deciding suitable network architectures for different pattern classification problems by dynamically growing the network to match the complexity of the underlying pattern classification task. Evolutionary algorithms offer a powerful means of exploring large search spaces for solutions that optimize multiple objectives e.g., feature subsets that maximize the predictive performance and minimize the complexity of the classifiers that use them. Against this background, we explored several closely related topics in biologically inspired (neural, evolutionary) algorithms and architectures for knowledge representation, language processing e.g., parsing, and learning. This work has led to:

- Generalization (with convergence guarantees) of a large family of constructive neural network learning algorithms designed for 2-class binary pattern classification problems to handle classification problems involving real-valued patterns and an arbitrary number of classes (Parekh et al., 2000).
- 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 (Yang et al., 1999).
- Development of algorithms for construction of robust, noise-tolerant neural memories for pattern storage and associative, content-based retrieval (Chen et al., 1995) and query processing (1996).
- Development of algorithms for construction of highly parallel neural architectures for syntax analysis (parsing of regular, context-free, and context-sensitive languages) (1999).
- 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 (Balakrishnan et al., 1998, 2000).
- Development of evolutionary algorithms for feature subset selection for classification problems (Yang et al., 1998) and sensors and controllers for adaptive robots (Balakrishnan and Honavar, 1996; 1998; 2001).

- Development of incremental neural network learning algorithms with applications in sensing and nondestructive evaluation (Polikar et al., 2001a, 2001b, 2004).

- Development of constructive neural network algorithms that take advantage of prior knowledge in the form of classification rules (Parekh et al., 1999).

- Hybrid neural-symbolic architectures for information processing (Honavar and Uhr, 1994; 1995).

Critical Infrastructure Monitoring and Protection (Funded in part by grants from the National Science Foundation and the US Department of Defense)

Formal methods offer a rigorous foundation for representing and reasoning about critical infrastructure for information and communication, transportation, and health. Data-driven approaches complement formal methods in domains where the available formal specifications are incomplete. Our work has used both approaches to solve problems that arise in monitoring and protecting critical infrastructure. This work has led to:

- Development of model checking approaches to containing infection propagation across networks (Santhanam et al., 2011).

- Development of tools for formal specification of intrusions software fault trees and of intrusion detection systems using colored Petri nets s for automated generation of multi-agent systems for coordinated intrusion detection in computer and communication networks (Helmer et al., 2002; 2006).

- 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 (Helmer et al., 2003Wang et al., 2006).

- Development and application of machine learning approaches for learning predictive rules for anomaly and misuse detection (Kang et al., 2005, 2006).

- Development of an electronic nose for detection and identification of odorants using machine learning (Polikar et al., 2001).

- Development and applications of machine learning methods for non-destructive inspection of nuclear power plant pipes using ultrasound (Polikar et al., 2006).

- Development of a service-oriented distributed software infrastructure for monitoring distributed power systems (Pathak et al., 2006; 2007; McCalley et al., 2007; Pham et al., 2009).

- Design, analysis, and evaluation 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) (Mikler et al., 1996; 1997; 1998; 2001).

PUBLICATIONS AND PRESENTATIONS

Plenary Lectures, Invited Keynotes or and Invited Talks at Conferences

1. Invited Talk, From Data Analytics to Discovery Informatics. NIH NIAID Symposium on Unlocking the Power of Big Data. National Institutes of Health, 2013.

2. Invited Keynote Talk, From Big Data Analytics to Discovery Informatics. Conference on Complex Adaptive Systems, Washington DC. November 2012.

3. Invited Talk, Computational Prediction of Protein Interfaces and Interactions. Conference on Modeling Protein Interactions, Lawrence, Kansas, November 2012.

4. Invited Keynote Talk, Learning Predictive Models from Distributed Data. Conference on Intelligent Data Understanding, Boulder, CO, October 2012.

5. Invited Keynote Talk, Humanities as Information Sciences. [Chicago Colloquium on Digital Humanities and Computer Science](), Chicago, November 2009.

6. Invited Keynote Talk, Aligning Macromolecular Networks. Sixth International Biotechnology and Bioinformatics Symposium (BIOT 2009), Lincoln, Nebraska, October 2009.

7. Invited Plenary Talk, Machine Learning in Bioinformatics, Annual Conference of the Italian Association for Artificial Intelligence (AI*IA 2008), Cagliari, Italy, September 2008.

8. Invited Keynote Talk, International Congress on Pervasive Computing and Management (ICPCM 2008), New Delhi, India, December 2008.

9. Invited Talk, Telluride Meeting on Characterizing the Landscape From Biomolecules to Cellular Networks, Telluride, Colorado, July 2008.

10. Invited Talk, Privacy-preserving Reasoning, Semantic Technology Conference, San Jose, CA, USA, May 2008

11. Invited Keynote Talk, Computational Structural Bioinformatics Workshop, IEEE Conference on Bioinformatics and Biomedicine, Silicon Valley, 2007.

12. Invited Talk, Making Biology and Medicine a Predictive Science. NSF Workshop on Biomedical Informatics. Oregon, 2007.

13. Invited Talk, Knowledge Acquisition from Semantically Disparate Distributed Data. NSF Workshop on Next Generation Data Mining and Cyber-Enabled Discovery, Baltimore, Maryland, 2007.

14. Invited Talk, On Selective Sharing and Reuse of Ontologies, Semantic Technology Conference, San Jose, CA, USA, May 2007

15. Invited Keynote Talk, Semantic Web for Collaborative e-Science, International Conference on Intelligent Sensing and Information Processing, Bangalore, India, 2006.

16. Invited Lecture: Querying Semantically Heterogeneous Data Sources from a User's Point of View, Semantic Technology Conference, San Jose, CA, USA, March 2006.

17. 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.

18. Plenary Talk, Data-Driven Discovery of Macromolecular Sequence-Structure-Function Relationships. International Conference on Intelligent System Design and Applications, 2003.

19. Invited Talk, Agent-Based Distributed Intelligent Information Networks for Computational Inference and Knowledge Discovery in Bioinformatics. In: Workshop on Agents in Bioinformatics, Italy, 2002.

20. Plenary Talk, Computational Discovery of Protein Sequence-Structure-Function Relationships, Diversity in Information Science and Technology, Nebraska EPSCOR Conference, 2002

21. Invited Keynote Talk, Learning from Large, Distributed, Heterogeneous Data Sets. International Symposium on Artificial Intelligence (ISAI 2001), Kolhapur, India.

22. Invited Talk, Distributed Intelligent Information Networks. Midwestern Conference on Artificial Intelligence and Cognitive Science, 2000.

23. Invited Talk, Cumulative Learning in Open Environments. International Workshop on Current Computational Architectures Integrating Neural Networks and Neuroscience. Durham Castle, United Kingdom. 2000.

24. 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.

Books Authored and Edited

1. Honavar, V., and Caragea, D., Koul, N., Silvescu, A., Zhang, J and Lin, H. (2012) 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 two papers that received "best paper" awards and several that are highly cited. According to Google scholar search (December 2012), the list includes: 1 article (Yang & Honavar, 1998) with over 844 citations which is among the 10 most cited publications in *IEEE Intelligent Systems,* 1 article with 318 citations, 3 articles with between 150 and 200 citations each, 3 articles with between 100 and 150 citations each, 23 articles with between 50 and 99 citations each, 5 articles with between 40 and 49 citations each, 27 articles with between 30 and 39 citations each, 33 articles with between 20 and 29 citations each, and 57 articles with between 10 and 19 citations each. The *h*-index[3] computed using Google Scholar citations data (as of December 2012) is 38, g-index is 69, and *i*10-index is 153.

1. Lee, S. and Honavar, V. (2013). m-Transportability: Transportability of a Causal Effect from Multiple Environments. In: Proceedings of the 27th Conference on Artificial Intelligence (AAAI 2013). In press.

2. Andorf, C., Honavar, V. and Sen, T. (2013). Predicting the Binding Patterns of Proteins: A Study Using Yeast Protein Interaction Networks. PLOS One 8(2): e56833, doi: 10.1371/journal.pone.0056833

3. Oster, Z., Santhanam, G., Basu, S. and Honavar, V. (2013). Model Checking of Qualitative Sensitivity Preferences to Minimize Credential Disclosure. International Symposium on Formal Aspects of Component Software. Revised Selected Papers. Berlin: Springer. Lecture Notes in Computer Science Vol. 7684, pp. 205-223.

4. Kumar, S; Nilsen, W.; Abernethy, A.; Atienza, A.; Patrick, K.; Pavel, M.; Riley, W. T; Shar, A.; Spring, B.; Spruijt-Metz, D.; Hedeker, D; Honavar, V.; Kravitz, R.L.; R. Lefebvre, C.; Mohr, D.C.; Murphy, S.A.; Quinn, C.; Shusterman, V.; Swendeman, D. (2013). Exploring Innovative Methods to Evaluate the Efficacy and Safety of Mobile Health. American Journal of Preventive Medicine, In Press.

5. Yasser El-Manzalawy, Drena Dobbs, Vasant Honavar: Predicting protective bacterial antigens using random forest classifiers. In Proceedings of the ACM Conference on Bioinformatics and Computational Biology (BCB 2012), pp. 426-433

6. Jordan, R., El-Manzalawy, Y., Dobbs, D., and Honavar, V. (2012). Predicting protein-protein interface residues using local surface structural similarity. BMC Bioinformatics 2012, 13:41 doi:10.1186/1471-2105-13-41. Highly Accessed.

---

[3] The *h*-index of an author is *n* if the author has *n* publications with at least *n* citations each. The *g*-index of an author is *m* if the author has *m* publications that taken together have at least m^2 citations.The i10-index of an author is *m* if the author has *m* publications with at least 10 citations each.

7.   Tao, J., Slutzki, G., and Honavar, V. (2012). PSpace Tableau Algorithms for Modalized ALC. Journal of Automated Reasoning. DOI: 10.1007/s10817-011-9232-3, 49(4): 551-582.

8.   Towfic, F., Gupta, S., Honavar, V., and Subramaniam, S. (2012). B-Cell Ligand Processing Pathways Detected by Large-Scale Gene Expression Analysis. Genomics, Proteomics, and Bioinformatics. 10(3): 142-152 (2012).

9.   Towfic, F., Kohutyuk, O., Greenlee, MHW., and Honavar, V. (2012). Bionetworkbench: Database and Software for Storage, Query, and Interactive Analysis of Gene and Protein Networks. Bioinformatics and Biology Insights 2012:6 235-246, doi: 10.4137/BBI.S9728

10.   Tu, K. and Honavar, V. (2012). Unambiguity Regularization for Unsupervised Learning of Probabilistic Grammars. In: Proceedings of EMNLP-CoNLL 2012 : Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. pp. 1324-1334

11.   Walia, R., Caragea, C., Lewis, B., Towfic, F., Terribilini, M., El-Manzalawy, Y., Dobbs, D., Honavar, V. (2012). Protein-RNA Interface Residue Prediction Using Machine Learning: An Assessment of the State of the Art. BMC Bioinformatics 13:89 doi:10.1186/1471-2105-13-89. **Highly Accessed.**

12.   El-Manzalawy, Y., Dobbs, D., and Honavar, V. (2011). Predicting MHC-II binding affinity using multiple instance regression. IEEE/ACM Transactions on Computational Biology and Bioinformatics. DOI: 10.1109/TCBB2010.94

13.   Lewis, B.A., Walia, R.R., Terribilini, M., Ferguson, J., Zheng, C., Honavar, V., and Dobbs, D. (2011). PRIDB: A Protein-RNA Interface Database. Nucleic Acids Research. D277-282. DOI: 10.1093/nar/gkq1108.

14.   Lin, H., Koul, N., and Honavar, V. (2011). Learning Relational Bayesian Classifiers from RDF data. In: Proceedings of the International Semantic Web Conference (ISWC 2011). Springer-Verlag Lecture Notes in Computer Science Vol. 7031 pp. 389-404.

15.   Muppirala, U.K., Honavar, V., Dobbs, D. (2011). Predicting RNA-Protein Interactions Using Only Sequence Information. BMC Bioinformatics 2011, 12:489, doi:10.1186/1471-2105-12-489. **Highly Accessed.**

16.   Santhanam, G., Basu, S., and Honavar, V. (2011). Representing and Reasoning with Qualitative Preferences for Compositional Systems. Journal of Artificial Intelligence Research Vol 42, pp. 211-274.

17.   Santhanam, G., Suvorov, Y., Basu, S., and Honavar, V. (2011). Verifying Intervention Policies for Infection Propagation over Networks: A Model Checking Approach. In: Proceedings of the Twenty-Fifth Conference on Artificial Intelligence (AAAI-2011). Pp. 1408-1414.

18.   Santanam, G., Basu, S., and Honavar, V. (2011). Identifying Sustainable Designs Using Preferences Over Sustainability Attributes. In: AAAI Spring Symposium on Artificial Intelligence. pp. 91-97.

19.   Silvescu, A. and Honavar, V. (2011). Abstraction Super-structuring Normal Forms: Towards a Theory of Structural Induction. In: The Proceedings of the Solomonoff 85th Memorial Conference. Springer-Verlag Lecture Notes in Artificial Intelligence. In press.

20.   Tu, K. and Honavar, V. (2011). On the Utility of Curricula in Unsupervised Learning of Grammars. In: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI 2011) pp. 1523-1528.

21.   Tu, K., Ouyang, X., Han, D., Yu, Y., and Honavar, V. (2011). Exemplar-based Robust Coherent Biclustering. In: Proceedings of the SIAM Conference on Data Mining (SDM 2011). pp. 884-895.

22.   Xue, L., Dobbs, D., and Honavar, V.. (2011). HomPPI: A Class of Sequence Homology Based Protein-Protein Interface Prediction Methods. BMC Bioinformatics 12:244 doi:10.1186/1471-2105-12-244

23.   Xue, L., Jordan, R., El-Manzalawy, Y., Dobbs, D., and Honavar, V. (2011). Sequence Based Partner-Specific Prediction of Protein- Protein Interfaces and its Application in Ranking Docked Models. In: ACM Conference on Bioinformatics and Computational Biology.

24.   Yakhnenko, O., and Honavar, V. (2011). Multi-Instance Multi-Label Learning for Image Classification with Large Vocabularies. In: Proceedings of the British Machine Vision Conference. http://dx.doi.org/10.5244/C.25.59.

25.   Barnhill, A.E., Hecker, L.A., Kohutyuk, O., Buss, J.E., Honavar, V. and Greenlee, H.W. (2010) Characterization of the Retinal Proteome During Rod Photoreceptor Genesis. BMC Research Notes 3:25.

26.   Caragea, C., Silvescu, A., Caragea, D. and Honavar, V. (2010). Abstraction-Augmented Markov Models. In: Proceedings of the IEEE Conference on Data Mining (ICDM 2010). IEEE Press. pp. 68-77.

27. Caragea, C. Silvescu, A., Caragea, D., and Honavar, V. (2010). Semi-supervised prediction of protein subcellular localization using abstraction augmented Markov models. BMC Bioinformatics. doi: 10.1186/1471-2105-11-S8-S6.

28. Caragea, C., Silvescu, A., Caragea, D., and Honavar, V. (2010). Semi-Supervised Sequence Classification Using Abstraction Augmented Markov Models. In: Proceedings of the ACM Conference on Bioinformatics and Computational Biology. pp. 257-264, doi: 10.1145/1854776.1854813. ACM Press.

29. El-Manzalawy, Y. and Honavar, V. (2010). Recent Advances in B-Cell Epitope Prediction Methods. Immunome Research Suppl. 2:S2. http://www.immunome-research.com/content/pdf/1745-7580-6-S2-S2.pdf.

30. Koul, N., Bui, N., and Honavar, V. (2010). Scalable, Updatable Predictive Models for Sequence Data. In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2010).

31. Koul, N. and Honavar, V. (2010). Learning in the Presence of Ontology Mapping Errors. In: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. pp. 291-296. ACM Press.

32. Pandit, S., and Honavar, V. (2010). Ontology-Guided Extraction of Complex Nested Relationships from Text. IEEE Conference on Tools With Artificial Intelligence (ICTAI 2010). pp. 173-178.

33. Sanghvi, B., Koul, N., and Honavar, V. (2010). Identifying and Eliminating Inconsistencies in Mappings across Hierarchical Ontologies. In: Springer-Verlag Lecture Notes in Computer Science Vol. 6427, pp. 999-1008. Berlin: Springer.

34. Santhanam, G., Basu, S., and Honavar, V. (2010). Dominance Testing Via Model Checking. In: Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI-10). AAAI Press. pp. 357-362.

35. Santhanam, G., Basu, S., and Honavar, V. (2010). Efficient Dominance Testing for Unconditional Preferences. In: Proceedings of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR 2010), pp 590-592.

36. Sun, H., Basu, S., Honavar, V., and Lutz, R. (2010). Automata-Based Verification of Security Requirements of Composite Web Services. In: Proceedings of the IEEE International Symposium on Software Reliability Engineering (ISSRE-2010). pp. 348-357, IEEE Press.

37. Tao, J., Slutzki, G., and Honavar, V. (2010). Secrecy-preserving Query Answering for EL. In: Proceedings of the 4th International Conference on Web Reasoning and Rule Systems (RR 2010). Berlin: Springer. In press.

38. Towfic, F., VanderPlas, S., Oliver, C.A., Couture, O., Tuggle, C.K., Greenlee, M.H.W., and Honavar, V. (2010). Detection of gene orthology from gene co-expression and protein interaction networks. BMC Bioinformatics, BMC Bioinformatics 2010, 11 (Suppl 3): S7 doi:10.1186/1471-2105-11-S3-S7

39. Towfic, F., Caragea, C., Dobbs, D., and Honavar, V. (2010). Struct-NB: Predicting protein-RNA binding sites using structural features. International Journal of Data Mining and Bioinformatics, Vol. 4, No.1 pp. 21 - 43

40. Tuggle, C.K., Bearson, S.M.D, Huang, T.H., Couture, O., Wang, Y., Kuhar, D., Lunney, J.K., Honavar, V. (2010). Methods for transcriptomic analyses of the porcine host immune response: Application to Salmonella infection using microarrays. Veterinary Immunology and Immunopathology. Vol. 138. pp. 282-291.

41. Bao, J., Voutsadakis G., Slutzki, G. Honavar, V. (2009). Package-Based Description Logics. In: Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization. Lecture Notes in Computer Science Vol. 5445, pp. 349-371.

42. Bromberg, F., Margaritis, D., and Honavar, V. (2009). Efficient Markov Network Structure Discovery from Independence Tests. Journal of Artificial Intelligence Research. Vol. 35. pp. 449-485.

43. Caragea, C., Sinapov, J., Dobbs, D., and Honavar, V. (2009). Mixture of experts models to exploit global sequence similarity on biomolecular sequence labeling. BMC Bioinformatics. doi:10.1186/1471-2105-10-S4-S4.

44. Caragea, C., Caragea, D., and Honavar, V. (2009). Learning Link-Based Classifiers from Ontology-Extended Distributed Data. In: Proceedings of the 8th International Conference on Ontologies, Databases, and Applications of Semantics (ODBASE). Springer-Verlag Lecture Notes in Computer Science vol. 5871, pp. 1139-1146.

45. Caragea, C., Caragea, D., and Honavar, V. (2009). Learning Link-Based Classifiers from Ontology-Extended Textual Data. In: Proceedings of the IEEE Conference on Tools with Artificial Intelligence. IEEE Press. pp. 354-361

46. Couture, O., Callenberg, K., Koul, N., Pandit, S., Younes, J., Hu, Z-L., Dekkers, J., Reecy, J., Honavar, V., and Tuggle, C. (2009). ANEXdb: An Integrated Animal Annotation and Microarray Expression Database. Mammalian Genome. Vol. 20. pp. 768-777.

47. El-Manzalawy, Y. and Honavar, V. (2009). MICCLLR: Multiple-Instance Learning using Class Conditional Log Likelihood Ratio. In: Proceedings of the 12th International Conference on Discovery Science (DS 2009). Springer-Verlag Lecture Notes in Computer Science vol. 5808. pp. 80-91.

48. Koul, N., and Honavar, V. (2009). Design and Implementation of a Query Planner for Data Integration. In: Proceedings of the IEEE Conference on Tools with Artificial Intelligence. IEEE Press. pp. 214-218.

49. Pham, H., Santhanam, G., McCalley, J., and Honavar, V. (2009). BenSOA: a Flexible Service-Oriented Architecture for Power System Asset Management. In Proceedings of the North American Power Symposium (NAPS).

50. Santhanam, G.R., Basu, S., and Honavar, V. (2009). Web Service Substitution Based on Preferences Over Non-functional Attributes. In: Proceedings of the IEEE International Conference on Services Computing (SCC 2009). IEEE Press. pp. 210-217.

51. Silvescu, A., Caragea, C., and Honavar, V. (2009). Combining Abstraction and Superstructuring for Sequence Classification. In: Proceedings of the IEEE Conference on Data Mining (ICDM 2009). IEEE press. pp. 986-991

52. Towfic, F., Greenlee, H., and Honavar, V. (2009). Aligning Biomolecular Networks Using Modular Graph Kernels. In: Proceedings of the 9th Workshop on Algorithms in Bioinformatics (WABI 2009). Berlin: Springer-Verlag: LNBI Vol. 5724, pp. 345-361.

53. Towfic, F., Greenlee, H., and Honavar, V. (2009). Detecting Orthologous Genes Based on Protein-Protein Interaction Networks. In: Proceedings of the IEEE Conference on Bioinformatics and Biomedicine (BIBM 2009). IEEE Press. pp. 48-53.

54. Yakhnenko, O., and Honavar, V. (2009). Multi-Modal Hierarchical Dirichlet Process Model for Predicting Image Annotation and Image-Object Label Correspondence. In: Proceedings of the SIAM Conference on Data Mining, SIAM. pp. 281-294.

55. Bao, J., Voutsadakis, G., Slutzki, G., and Honavar, V. (2008). On the Decidability of Role Mappings between Modular Ontologies. Proceedings of the 23nd Conference on Artificial Intelligence (AAAI-2008), AAAI, 404-409.

56. Caragea D., Cook, D., Wickham H. and Honavar, V. (2008). Visual Methods for Examining SVM Classifiers. Simeon J. Simoff, Michael H. Böhlen, Arturas Mazeika (Eds.): Visual Data Mining - Theory, Techniques and Tools for Visual Analytics. Springer-Verlag Lecture Notes in Computer Science Vol. 4404 pp.136-153

57. Caragea, C., Sinapov, J., Dobbs, D., and Honavar, V. (2008). Using global sequence similarity to enhance biological sequence labeling. Proceedings of the IEEE Conference on Bioinformatics and Biomedicine (BIBM-2008). IEEE Press. pp. 104-111.

58. Dunn-Thomas, T., Dobbs, D.L., Sakaguchi, D. Young, M.J. Honavar, V. Greenlee, H. M. W. (2008). Proteomic Differentiation Between Murine Retinal and Brain Derived Progenitor Cells. Stem Cells and Development. 17: 191-131.

59. El-Manzalawy, Y., Dobbs, D., and Honavar, V. (2008). Predicting linear B-cell epitopes using string kernels. Journal of Molecular Recognition, DOI:10.1002/jmr.893. vol. 21. pp. 243-255.

60. El-Manzalawy, Y., Dobbs, D., and Honavar, V. (2008). On Evaluating MHC-II Binding Peptide Prediction Methods. PLoS One, 3(9):e3268

61. El-Manzalawy, Y., Dobbs, D., and Honavar, V. (2008). Predicting Flexible Length Linear B-cell Epitopes, 7th International Conference on Computational Systems Bioinformatics, Stanford, CA. Singapore: World Scientific. pp. 121–132.

62. El-Manzalawy, Y., Dobbs, D., and Honavar, V. (2008). Predicting Protective Linear B-cell Epitopes using Evolutionary Information. Proceedings of the IEEE Conference on Bioinformatics and Biomedicine (BIBM-2008). IEEE Press. pp. 289-292

63. Hecker, L., Alcon, T., Honavar, V., and Greenlee, H. (2008). Analysis and Interpretation of Large-Scale Gene Expression Data Sets Using a Seed Network. Journal of Bioinformatics and Biology Insights. Vol. 2. pp. 91-102.

64. Hughes, L., Bao. J., Honavar, V., and Reecy, J. (2008). Animal Trait Ontology (ATO): the importance and usefulness of a unified trait vocabulary for animal species. Journal of Animal Science 86: 1485-1491

65. Jo, H., Na, Y-C., Oh, B., Yang, J., and Honavar, V. (2008). Attribute Value Taxonomy Generation Using Matrix Based Genetic Algorithm. In: Proceedings of the IEEE Conference on Tools with Artificial Intelligence (ICTAI 2008). IEEE Press. pp. 393-400.

66. Koul, N., Caragea, C., Bahirwani, V., Caragea, D., and Honavar, V. (2008). Using Sufficient Statistics to Learn Predictive Models from Massive Data Sets. Proceedings of the ACM/IEEE/WIC Conference on Web Intelligence (WI-2008). pp. 923-926.

67. Koul, N., Lathrop, J., Lutz, J., and Honavar, V. (2008). Complexes of Online Self-Assembly. IEEE International Conference on Electro-Information Technology. IEEE Press. pp. 448-452.

68. 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. (2008). Striking Similarities in Diverse Telomerase Proteins Revealed by Combining Structure Prediction and Machine Learning Approaches. Pacific Symposium on Bioinformatics, Vol. 13 pp. 501-512

69. Pathak, J., Basu, S., Lutz, R., and Honavar, V. (2008). MoSCoE: An Approach for Composing Web Services through Iterative Reformulation of Functional Specifications. International Journal on Artificial Intelligence Tools. Vol. 17. No. 1. pp. 109-138.

70. Pathak, J., Basu, S., and Honavar, V. (2008). Assembling Composite Web Services from Autonomous Components. In: Emerging Artificial Intelligence Applications in Computer Engineering. Soldatos, J. et al. (ed). IOS Press. In press.

71. Pathak, J., Basu, S., and Honavar, V. (2008). Composing Web Services through Automatic Reformulation of Service Specifications. IEEE International Conference on Services Computing IEEE Press. pp. 361-369.

72. Peto M., Kloczkowski A., Honavar V., Jernigan R.L. (2008). Use of machine learning algorithms to classify binary protein sequences as highly-designable or poorly-designable. BMC Bioinformatics, 9:487- **Highly Accessed.**

73. Santhanam, G., Basu, S., and Honavar, V. (2008). TCP-Compose* - A TCP-net based Algorithm for Efficient Composition of Web Services Based on Qualitative Preferences. Proceedings of the International Conference on Service-Oriented Computing. IEEE Press. pp 453-467.

74. Tu, K. and Honavar, V. (2008). Unsupervised Learning of Probabilistic Context-Free Grammar using Iterative Biclustering. Proceedings of the International Colloquium on Grammatical Inference (ICGI-2008). Springer-Verlag Lecture Notes in Computer Science vol. 5278 pp. 224-237.

75. Voutsadakis, G., Bao, J., Slutzki, G., and Honavar, V. (2008). F-ALCI: A Fully Contextualized, Federated Logic for the Semantic Web. Proceedings of the ACM/IEEE/WIC Conference on Web Intelligence (WI-2008). pp. 575-578.

76. Yan, C., Wu, F., Jernigan, R.L., Dobbs, D., Honavar, V. (2008) Analysis of Protein-Protein Interfaces. The Protein Journal 27 :59-70.

77. 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/

78. 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

79. 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, 791-797.

80. Caragea, C., Sinapov, J., Silvescu, A., Dobbs, D. And Honavar, V. (2007). Glycosylation Site Prediction Using Ensembles of Support Vector Machine Classifiers. BMC Bioinformatics. doi:10.1186/1471-2105-8-438 **Highly Accessed.**

81. 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.

82. 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.

83. 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.

84. Pathak, J., Basu, S., and Honavar, V. (2007). On Context-Specific Substitutability of Web Services. Proceedings of the IEEE International Conference on Web Services, pp. 192-199. IEEE press.

85. Pathak, J., Li, Y., Honavar, V., McCalley, J. (2007). A Service-Oriented Architecture for Electric Power Transmission System Asset Management. Second International Workshop on Engineering Service-Oriented Applications: Design and Composition, Lecture Notes in Computer Science, Berlin: Springer-Verlag, 2007.

86. 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

87. 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, CA.

88. 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.

89. 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.

90. 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.

91. 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.

92. 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.

93. 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.

94. 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**.

95. 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.

96. 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.

97. 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.

98. 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

99. 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.

100. 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

101. 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.**

102. 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.

103. 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.

104. 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.

105. 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.

106. 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.

107. 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.

108. 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.

109. 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.

110. 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.**

111. 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.

112. 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.

113. 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.

114. 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.

115. 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

116. 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

117.	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

118.	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.

119.	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.

120.	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.

121.	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.

122.	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.

123.	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.

124.	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).

125.	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), Springer-Verlag Lecture Notes in Computer Science vol. 3291. pp. 963-980.

126.	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

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

128.	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. IEEE Press. pp. 130-137.

129.	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.

130.	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). Springer-Verlag Lecture Notes in Computer Science Vol. 3372, pp. 41–56.

131.	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.

132.	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**.

133.	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.

134.   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.)

135.   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).

136.   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. IEEE Press. pp. 289-296

137.   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.

138.   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.

139.   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.

140.   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.

141.   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.

142.   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.

143.   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

144.   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.

145.   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.

146.   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.

147.   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

148.   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.

149.   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.

150.   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.

151.   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. In: Proceedings of the Atlantic Symposium on Computational Biology, Genome Information Systems & Technology.

152. 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.

153. 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.

154. 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.

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

156. 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.

157. 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.

158. 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: Proceedings of the Atlantic Symposium on Computational Biology, Genome Information Systems & Technology).

159. 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.

160. 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.

161. 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.

162. 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.

163. 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.

164. 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.

165. 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.

166. 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.

167. 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.

168. 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.

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

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

171. 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.

172. 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.

173. 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.

174. 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.

175. 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.

176. 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 Modeling. pp. 112-119. Nottingham University Press.

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

178. 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.

179. 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.

180. 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.

181. 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.

182. 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.

183. 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.

184. 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.

185. 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.

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

187. 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.

188. 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.

189. 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.

190. Yang, J. & Honavar, V. (1998). Feature Subset Selection Using a Genetic Algorithm. **Invited** paper. IEEE Intelligent Systems vol. 13, no. 2., pp. 44-49. **Among the 10 most cited articles in IEEE Intelligent Systems.**

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

192. 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.

193. 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.

194. 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.

195. 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.

196. 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.

197. 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.

198. 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.

199. 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.

200. 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.

201. 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.

202. 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.

203. 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.

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

205. 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.

206. 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.

207. 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.

208. 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.

209. 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.

210. 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.

211. 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.

212. 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.

213. 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.

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

215. 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.

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

217. 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.

218. 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.

219. 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.

220. 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.

221. 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.

222. 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.

223. 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.

224. 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.

225. 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.

226. 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.

227. 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.

228. 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.

229. 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.

230. 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.

231. 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.

232. 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.

233. Honavar, V. & Uhr, L. (1989). Brain-Structured Connectionist Networks that Perceive and Learn. Connection Science 1 139-159.

234. 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.

235. 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. Tuggle, C., Towfic, F., and **Honavar, V**. (2011). An Introduction to Computational Systems Biology for Animal Scientists. In: Systems Biology For Animal Scientists. In: te Pas, M.F.W., Woelders, H., and Bannick, A. (Ed). Wiley. In press.

2. Caragea, C. and **Honavar, V**. (2009). Machine Learning in Computational Biology. In: *Encyclopedia of Database Systems.* Raschid, L. (ed). Springer. Pp. 1663-1667

3. Caragea, D. and **Honavar, V**. (2009). 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.

4. **Honavar, V.** and Caragea, D. (2008). Towards a Semantics-Enabled Infrastructure for Knowledge Acquisition from Distributed Data. In: Next Generation Data Mining. Kargupta, H. et. al., CRC Press. pp. 317-337.

5. Caragea, D. and **Honavar, V**. (2008). *Learning Classifiers from Semantically Heterogeneous Data*. In: Encyclopedia of Data Warehousing and Mining. Wang, J. (Ed).

6. Pathak, J., Basu, S., and **Honavar, V.** (2008). Assembling Composite Web Services from Autonomous Components. In: Emerging Artificial Intelligence Applications in Computer Engineering. Soldatos, J. et al. (ed). IOS Press. pp. 394-405.

7. **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.

8. 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.

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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.

19. 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.

20. 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.

21. 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.

22. 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, Lightly Refereed Papers, Extended Abstracts, and Posters in Conferences and Workshops

1. Santhanam, G., Basu, S., and Honavar, V. (2011). Identifying Sustainable Designs Using Preferences Over Sustainability Attributes. In: AAAI Spring Symposium on Artificial Intelligence in Sustainable Design. Stanford, CA.

2. Caragea, C., Silvescu, A., Caragea, D. and Honavar, V. (2009). Abstraction Augmented Markov Models. In: NIPS Workshop on Machine Learning in Computational Biology (MLCB).

3. Tu, K., and Honavar, V. (2009). An Empirical Study of Hierarchical Dirichlet Process Priors for Grammar Induction. In: NIPS Workshop on Grammar Induction, Representation of Language, and Language Learning.

4. Yakhnenko, O. and Honavar, V. (2009). Multiple label prediction for image annotation with multiple kernel correlation models. In: CVPR Workshop on Visual Context Learning.

5. Santhanam, G., Basu, S., and Honavar, V. (2008). On Utilizing Qualitative Preferences in Web Service Composition: A CP-net based approach 2nd International Workshop on Web Service Composition and Adaptation (WSCA) 2008

6. 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*, 2007.

7. 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.

8. 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.

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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).

19. 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.

20. 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).

21. 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).

22. 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).

23. 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).

24. 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.

25. 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).

26. 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.

27. 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.

28. 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.

29. 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).

30. 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.

31. 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.

32. 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.

33. Silvescu A., and Honavar V. (2003) Ontology Elicitation: Structural Abstraction = Structuring + Abstraction + Multiple Ontologies. Poster presentation. Learning Workshop, Snowbird, Utah, 2003.

34. 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/

35. 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.

36. 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.

37. 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.

38. 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.

39. 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.

40. 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.

41. 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.

42. 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.

43. 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.

44. Honavar, V. (1997). The Design Process: A Computational Perspective. In: Proceedings of the NSF Workshop on Decision-based Design, Sacramento, CA.

45. 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.

46. 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.

47. Parekh, R., Yang, J. & Honavar, V. (1996). In: Proceedings of the Thirteenth National Conference on Artificial Intelligence. AAAI Press. vol. 2. pp. 1398.

48. 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.

49. 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.

50. 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.

51. Honavar, V. (1991). Generative Learning in Generalized Connectionist Networks. In: Constructive Induction Session - Eighth International Workshop on Machine Learning. Evanston, IL.

52. Honavar, V. (1991). Language and Knowledge: Communication, Acquisition, and Evolution. Invited presentation in: Second International Workshop on Human and Machine Cognition. Perdido Key, Florida.

53. 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.

54. 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).

55. 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.

56. 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.

Selected Invited Colloquia

1. **Distinguished Lecture,** Computational Prediction of Protein Interfaces and Interactions, Georgia State University. January 2013.

2. **Distinguished Lecture,** From Big Data Analytics to Discovery Informatics. Pennsylvania State University, December 2012.

3. **Invited Colloquium,** Computational Prediction of Protein Interfaces and Interactions, University of California Irvine, September 2012.

4. **Invited Colloquium,** Computational Prediction of Protein Interfaces and Interactions. University of North Texas, June 2012.

5. **Invited Talk,** Towards Infrastructure for Collaborative Discovery. ICiS Workshop on Integrating, Representing, and Reasoning over Human Knowledge, August 2010.

6. **Invited Talk**, Knowledge Acquisition from Semantically Disparate, Distributed Data. CISE (IIS), National Science Foundation, May 2010.

7. **Invited Lecture**, From Annotating Sequences to Aligning Networks. Computation in Biology and Medicine Annual Retreat, University of Wisconsin-Madison, October 2009.

8. **Invited Colloquium,** Transforming Biology From a Descriptive Science into a Predictive Science, Indian Institute of Information Technology, Bangalore, India, January 2009.

9. **Invited Colloquium,** Transforming Biology From a Descriptive Science into a Predictive Science: Predictive Models of Macromolecular Function and Interaction. Bioinformatics Center, University of Pune, India, December 2008.

10. **Invited Colloquium,** Semantics-Enabled Infrastructure for Collaborative, Integrative e-Science. School of Information Technology, Jawaharlal Nehru University, New Delhi, India, December 2008.

11. **Invited Talk,** Computational Sciences. High Performance Computing Center, Jawaharlal Nehru University, New Delhi, India, December 2008.

12. **Invited Colloquium,** Semantics-enabled infrastructure for collaborative, integrative e-science. Yahoo!, Bangalore, India, January 2008.

13. **Invited Colloquium,** Algorithms and Software for Knowledge Acquisition from Semantically Heterogeneous, Distributed Data Sources. Dept. of Electrical and Computer Engineering. University of Iowa. 2006.

14. **Invited Colloquium,** Algorithms and Software for Collaborative Discovery in Systems Biology. Dept. Biostatistics, Bioinformatics & Epidemiology. Medical University of South Carolina, 2006.

15. **Invited Talk,** Algorithms and Software for Knowledge Acquisition from Semantically Heterogeneous, Distributed, Autonomous Information Sources. Google Research, 2005.

16. **Invited Talk,** All Science is Computer Science. Iowa Undergraduate Consortium. Drake University, 2004.

17. **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.

18. **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.

19. **Invited Talk,** Neuromimetic Adaptive Autonomous Intelligent Systems. Institute for Computer Applications in Science and Engineering. NASA-Langley Research Center. Hampton, VA. September 28, 1999.

20. **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.

21. **Invited Colloquium,** Experiments in Evolutionary Robotics. Department of Mathematics and Computer Science, Grinnell College, Iowa. October 1996.

22. **Invited Lecture,** Data Mining and Knowledge Discovery. Irish Life, Des Moines, Iowa. September 1996.

23. **Invited Lecture,** Knowledge Acquisition through Machine Learning. Principal Mutual, Des Moines, Iowa. January 1994.

24. **Invited Colloquium,** Generalized Connectionist Networks and Processes for Intelligent Systems. International Computer Science Institute, Berkeley, *CA.* (1990).

25. **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 Big Data and Discovery Informatics, AAAI Fall Symposium on Discovery Informatics, Washington DC, November 2012.

2. Panel on Big Data Research Opportunities and Challenges in Biological Sciences, ACM Conference on Bioinformatics and Computational Biology, Orlando, Florida, 2012.

3. Panel on Digital Humanities and Computer Science, Chicago Colloquium on Digital Humanities and Computer Science, Chicago, November 2009.

4. Panel on Data Mining Careers, SIAM Conference on Data Mining (SDM 2009), April 2009.

5. Panel on Interdisciplinary Research and Training in Bioinformatics and Computational Biology, Annual Conference of the Italian Association for Artificial Intelligence (AI*IA 2008), Cagliari, Italy, September 2008.

6. Panel on Semantic Data Integration. NSF Workshop on Biomedical Informatics, 2007.

7. Panel on Learning in Knowledge-Based Systems. Second World Congress on Expert Systems. Lisbon, Portugal (1994).

8. Panel on Hybrid Architectures for Intelligent Systems. Second World Congress on Expert Systems. Lisbon, Portugal (1994).

9. Panel on Hybrid Intelligent Systems (SIGHI meeting) *World Congress on Neural Networks*. San Diego, U.S.A. (1994).

SOFTWARE, DATABASES, SERVERS

- INDUS: Algorithms and Software for Learning Predictive Models from Distributed Data:
    - Indus Learning Framework http://code.google.com/p/induslearningframework/
    - Indus Integration Framework http://code.google.com/p/indusintegrationframework/
- iPref-R: A Qualitative Preference Reasoner: http://fmg.cs.iastate.edu/project-pages/preference-reasoner/
- Biomolecular Network Alignment Toolkit: http://www.cs.iastate.edu/~ftowfic/twiki/bin/view/Projects/BinaToolkit
- Bionetworkbench: http://bionetworkbench.sourceforge.net/
- Database of Protein-protein Interfaces: http://protindb.cs.iastate.edu/
- Database of Protein-RNA Interfaces: http://pridb.gdcb.edu
- NPS-HomPPI: Non Partner-Specific Sequence Homology-Based Protein-Protein Interface Prediction Server: http://einstein.cs.iastate.edu/NPSHOMPPI/
- PS-HomPPI: Partner-Specific Sequence Homology Based Protein-Protein Interface Prediction Server: http://einstein.cs.iastate.edu/PSHOMPPI/
- PrISE: A Local Structural Similarity Based Protein-Protein Interface Prediction Server: http://prise.cs.iastate.edu/
- DockRank: Ranking Docked Models Based on Predicted Interfaces: http://einstein.cs.iastate.edu/DockRank/
- ANExDB: Animal Gene Expression Database: http://www.anexdb.org/
- Protein-RNA Interface Prediction Server: http://einstein.cs.iastate.edu/RNABindR/
- Protein-DNA Interface Prediction Server: http://turing.cs.iastate.edu:8080/PredDNA/
- Glycosylation Site Prediction Server: http://turing.cs.iastate.edu/EnsembleGly/
- MHC-II Binding Affinity Prediction Server: http://ailab.cs.iastate.edu/mhcmir/
- B-cell epitope Prediction Server: http://ailab.cs.iastate.edu/bcpreds/
- Epitope Prediction Toolkit: http://ailab.cs.iastate.edu/epit/

STUDENT MENTORING

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. My students benefit from strong mentoring and close interactions within a very collaborative research group. 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 (Past)

1. **Cornelia Caragea** (2010) Research Associate, Center for Computational Intelligence, Learning, and Discovery (CCILD), Iowa State University. Topic: Machine Learning in Computational Biology. Cornelia was supported in part on a grant from the NSF and in part by CCILD. Current position: Research Associate, Pennsylvania State University.

2. **George Voutsadakis,** Research Associate, Center for Computational Intelligence, Learning, and Discovery (CCILD), Iowa State University. Topics: Federated Knowledge Bases, Semantic Web, Secrecy-Preserving Reasoning. George was supported in part by the ISU NSF Industry-University Cooperative Research Center in Computer and Network Security and in part by CCILD. Current position: Associate Professor of Mathematics and Computer Science, Lake Superior State University.

3. **Jae-Hyung Lee** (2008-2009). Research Associate, Center for Computational Intelligence, Learning, and Discovery (CCILD), Iowa State University. Topic: Construction and Analysis of Macromolecular Interaction Networks. Jae-Hyung was supported in part by CCILD. Current Position: Research Associate, UCLA.

4. **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 was supported in part by a grant from the National Science Foundation and in part by CCILD. Current Position: Research Associate (with Jim Hendler), RPI.

5. **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.

6. **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.

Post-Doctoral Fellows (Current)

1. **Yasser El-Manzalawy** (Computer Science). Research Associate, Center for Computational Intelligence, Learning, and Discovery (CCILD), Iowa State University. Topic: Machine learning approaches in Immune Informatics. Yasser is supported in part by CCILD.

2. **Ganesh Ram Santhanam** (Computer Science), Research Associate, Center for Computational Intelligence, Learning, and Discovery, Iowa State University. Topics: Representing and reasoning about preferences. Model checking approaches to finding and reasoning about intervention policies in networks.

3. **Li Xue** (Bioinformatics and Computational Biology). Topic: Computational Prediction of Protein Interfaces and Interactions.

4. **Jia Tao** (Computer Science) Research Associate, Center for Computational Intelligence, Learning, and Discovery. Iowa State University. Topics: Epistemic Description Logics, Secrecy-preserving Query Answering.

Ph.D. Graduates

1. **Li Xue** (Bioinformatics and Computational Biology; with Drena Dobbs). Sequence-Homology Based Methods for Protein Interface Prediction and their Application to Ranking Docked Conformations. Graduate Research Excellence Award, Ph.D., 2012. Li was supported in part by research assistantships funded by a grant from the National Institutes of Health and in part by the Iowa State University Center for Computational Intelligence, Learning, and Discovery. Initial Employment: Postdoctoral Research Associated, Iowa State University.

2. **Hsien-chao Chou** (Bioinformatics & Computational Biology; with Volker Brendel), Local assembly and pre-mRNA splicing analysis by high-throughput sequencing data. Ph.D., 2012.

3. **Raphael Jordan** (Computer Science). Structure-Based Prediction of Protein-Protein Interaction Sites. Graduate Research Excellence Award, Ph.D., 2012. Raphael was supported in part by a research assistantship funded by a grant from the National Science Foundation and in part by a teaching assistantship in Computer Science. Current Employment: Data Mining Research Scientist, Quantcast.

4. **Kewei Tu** (Computer Science), Ph.D., 2012. Unsupervised Learning of Probabilistic Grammars. Kewei was supported by in part by a teaching assistantship in Computer Science and in part by a research assistantship funded by a grant from the National Science Foundation. Initial Employment: Postdoctoral Research Associate, Department of Statistics and Department of Computer Sceince, University of California Los Angeles.

5. **Jia Tao** (Computer Science, with Giora Slutzki). Topics in Knowledge Bases: Epistemic Description Logics and Secrecy-preserving Reasoning. Ph.D., 2012. Initial Employment, Postdoctoral Research Associate, Iowa State University.

6. **Fadi Towfic** (Bioinformatics and Computational Biology, with **M. Heather West Greenlee**). Ph.D., 2011. Modular Algorithms for Biomolecular Network Alignment. Fadi was supported in part by a fellowship funded by an Integrative Graduate Education and Research Training (IGERT) grant from the National Science Foundation. Initial employment: Postdoctoral Research Associate, The Broad Institute of MIT and Harvard.

7. **Neeraj Koul** (Computer Science). Ph.D., 2011. Learning Predictive Models from Massive, Semantically Disparate Data. Neeraj was supported in part by a research assistantship funded by a grant from the National Science Foundation. Initial Employment: Machine Learning Research Scientist, GILT Group.

8. **Ganesh Ram Santhanam** (Computer Science), Ph.D., 2010. Representing and Reasoning with Qualitative Preferences for Compositional Systems. Ganesh was supported in part by research assistantships funded by a grant from the National Science Foundation. Initial Employment: Postdoctoral Research Associate, Center for Computational Intelligence, Learning, and Discovery, Iowa State University.

9. **George Voutsadakis** (Computer Science, with Giora Slutzki), Ph.D., 2010. Federated Description Logics for the Semantic Web. George was supported in part by research assistantships funded by a grant from the National Science Foundation. Current Position: Associate Professor of Mathematics and Computer Science, Lake Superior State University.

10. **Cornelia Caragea** (Computer Science), Ph.D., 2009. Abstraction-Based Probabilistic Models for Sequence Classification. Cornelia was supported in part by research assistantships funded by grants from the National Science Foundation and the National Institutes of Health. Current Position: Postdoctoral Research Associate, Center for Computational Intelligence, Learning, and Discovery, Iowa State University.

11. **Oksana Yakhnenko** (Computer Science), Ph.D., 2009. Learning from Text and Images: Probabilistic Generative and Discriminative Models for Partially Labeled Data. Oksana was supported in part by a graduate teaching assistantship in Computer Science and in part by a research assistantship funded by a grant from the National Science Foundation. Current Position: Postdoctoral Research Associate, INRIA, Rhone-Alpes, France.

12. **Kent Vander Velden** (Bioinformatics & Computational Biology; with Peter Reilly), Ph.D., 2009. Modeling, Simulation, Synthesis, and Optimization of Biochemical Networks. Kent was supported in part by a graduate fellowship funded by an NSF IGERT award and by Pioneer Hi-Bred. Current Position: Senior Research Scientist, Pioneer Hi-Bred.

13. **Yasser El-Manzalawy** (Computer Science), Ph.D., 2008. Thesis: Machine Learning Approaches to Epitope Prediction. Recipient of an Iowa State University Graduate Research Excellence Award. Yasser was supported by a fellowship from the Egyptian Government. Current Position: Assistant Professor, Al-Azhar University, Egypt.

14. **Michael Terribilini**, (Bioinformatics & Computational Biology; with Drena Dobbs), Ph.D., 2008. Thesis: Deciphering the Protein-RNA Recognition Code. Michael was supported by a Multidisciplinary Graduate Education and Training (MGET) fellowship funded by the USDA. Current Position: Assistant Professor of Biology, Elon College, North Carolina.

15. **Feihong Wu** (Bioinformatics & Computational Biology; with Robert Jernigan), Ph.D., 2009. Thesis: Protein-Protein Interfaces: Databases, Analysis, and Prediction. Feihong was supported by research assistantships funded by the Iowa State University Graduate College and a grant from the National Institutes of Health. Current Position: Research Scientist, Yahoo! Inc.

16. **Adrian Silvescu** (Computer Science), Ph.D., 2008; Thesis: Inductive Learning Via Abstraction and Superstructuring. Adrian was supported through a teaching assistantship from the Department of Computer Science and research assistantships funded by Pioneer Hi-Bred and the National Science Foundation. Current Position: Research Scientist, Yahoo! Inc.

17. **LaRon Hughes** (Bioinformatics and Computational Biology, with James Reecy). Ph.D., 2008. Thesis: Design, Implementation, and Some Applications of an Animal Trait Ontology. LaRon was supported in part by an NSF Integrative Graduate Education and Research Training (IGERT) Fellowship in Bioinformatics and Computational Biology. Initial Employment: Research Scientist, Genome Quest.

18. **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, Professor Jim Hendler's group, Department of Computer Science, Rensselaer Polytechnic Institute.

19. **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.

20. **Jyotishman Pathak** (Computer Science). Ph.D., 2007. Thesis: Interactive and Verifiable Web Service Composition, Reformulation, and Adaptation. Recipient of an Iowa State University Graduate Research Excellence Award. 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. Current Employment: Assistant Professor of Biomedical Informatics, Mayo College of Medicine, Rochester, Minnesota.

21. **Dae-Ki Kang** (Computer Science). Ph.D., 2007. 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.

22. **Jun Zhang** (Computer Science). Ph.D., 2005. Ontology Aware Learning Algorithms. Recipient of an Iowa State University Graduate Research Excellence Award. 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.

23. **Changhui Yan** (Bioinformatics and Computational Biology, with Drena Dobbs). Ph.D., 2004. Computational Approaches to Prediction of Protein-Protein, Protein-DNA and Protein-RNA Interfaces. 2005. Recipient of an Iowa State University Graduate Research Excellence Award. Changhui was supported by a Plant Sciences Fellowship and 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.

24. **Doina Caragea** (Computer Science). Ph.D., 2004. Learning Classifiers from Semantically Heterogeneous, Distributed, Autonomous Data Sources. 2004. Recipient of an Iowa State University Graduate Research Excellence Award. Initial

Employment: Research Associate, Center for Computational Intelligence, Learning, and Discovery, Iowa State University. Current position: Assistant Professor of Computer Science, Kansas State University.

25. **Jihoon Yang** (Computer Science). Ph.D., 1999. 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.

26. **Karthik Balakrishnan** (Computer Science). Ph.D., 1998. Biologically Inspired Computational Structures and Processes for Autonomous Agents and Robots, Recipient of an Iowa State University Graduate Research Excellence Award. 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.

27. **Rajesh Parekh** (Computer Science). Ph.D., 1998. Recipient of an Iowa State University Graduate Research Excellence Award. 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!

28. **Chun-Hsien Chen** (Computer Science). Ph.D., 1997. 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.

29. **Armin Mikler** (Computer Science, with Johnny Wong), Ph.D., 1995. Quo Vadis - A Framework for Intelligent Routing in Large Communication Networks. 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. **Carson Andorf** (Bioinformatics & Computational Biology; with Drena Dobbs), In progress. Interests: Genome Annotation, 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: Spring 2013. Current employment: Research Scientist, USDA, Ames, Iowa.

2. **Ngot Bui** (Computer Science), In progress. Interests: Semantic technologies, Knowledge Representation, Machine Learning, Learning from Structured Data. Expected graduation: Spring 2014.

3. **Sangchack Lee** (Computer Science). In progress. Interests: Machine Learning. Expected graduation: Spring 2014.

4. **Harris Lin** (Computer Science). In progress. Interests: Relational Learning, Learning from Structured Data, Knowledge Representation, Semantic Technologies. Harris is supported in part by a research assistantship funded in part by a grant from the National Science Foundation. Expected graduation: Fall 2012.

5. **Rasna Walia** (Bioinformatics & Computational Biology; with Drena Dobbs), In progress. Interests: Macromolecular Interactions, Computational Systems Biology, Machine Learning. Expected Graduation: Spring 2014.

Major Professor, M.S.
Current M.S. Students

1. **Munish Gopal** (Computer Science). Interests: Machine Learning. Expected graduation: Spring 2013.

2. **Yiming Zhang** (Bioinformatics & Computational Biology; with Volker Brendel), In progress. Interests: Genomics. Expected graduation: Spring 2013.

M.S. Graduates

1. **Sateesh Kodavalli** (Computer Science). M.S., 2010. Extensible Problem Specific Tutor (xPST): Easy Authoring of Intelligent Tutoring Systems.

2. **Bhavesh Sanghvi** (Computer Science). M.S., 2010. Identifying and eliminating inconsistencies in mappings across hierarchical ontologies. Current Employment: Microsoft.

3. **Sushain Pandit** (Computer Science), M.S., 2010. Ontology-guided extraction of structured information from unstructured text: Identifying and capturing complex relationships. Current Employment: IBM.

4. **Tim Alcon** (Bioinformatics & Computational Biology; with Heather Greenlee)**.** Using a Seed-Network to Query Multiple Large-Scale Gene Expression Datasets from the Developing Retina in Order to Identify and Prioritize Experimental Targets. Tim was supported by a Multidisciplinary Graduate Education and Training (MGET) fellowship funded by the USDA.

5. **Lucas Bonansea** (Human-Computer Interaction; with Stephen Gilbert), 2009. 3-d hand gesture recognition.

6. **Flavian Vasile** (Computer Science), 2008. Uncovering the structure of hypergraphs through tensor decomposition: An application to folksonomy analysis. Flavian was supported in part by a teaching assistantship in Computer Science. Current Employment: Software Engineer, Yahoo! Inc.

7. **Oksana Kohutyuk** (Computer Science), 2007. Thesis: Retina Workbench: Database and software tools for comparative analysis and querying of gene and protein networks. Oksana was supported by a research assistantship funded by the National Institutes of Health and a teaching assistantship in Computer Science. Current position: Software Engineer, Cisco.

8. **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.

9. **Anna Atramentov**, A Multi-Relational Decision Tree Learning Algorithm – Implementation and Experiments. 2003. Current position: Ph.D. Student, University of Illinois at Urbana-Champaign

10. **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.

11. **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.

12. **Hector Leiva**, Learning Classifiers from Relational Data. 2002. Initial Employment: Research Scientist, Research Scientist, Universidad Nacional de San Luis. Argentina.

13. **Xiaosi Zhang**, Identification of Functionally Related Genes from Gene Expression Data. 2002.Initial Employment: Papajohn Center for Entreprenuership, Ames, Iowa.

14. **Xiangyun Wang**, Data Mining Approach to Discovery of Protein Sequence-Structure-Function Relationships. 2002. Initial Employment: Astra-Zeneca Inc.

15. **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.

16. **Neeraj Koul**, Clustering With Semi-Metrics, 2001. Initial Employment: Motorola.

17. **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.

18. **Rushi Bhatt**, Spatial Learning and Localization: A Computational Model and Behavioral Simulations, 2001. Ph.D. Program, Boston University.

19. **Fajun Chen**, Learning Information Extraction Patterns from Text, 2000. Initial Employment: Ericsson.

20. **Tarkeshwari Sharma** Agent Toolkit for Distributed Knowledge Networks, 2000. Initial Employment: Motorola, Inc.

21. **Asok Tiyyagura**, Alternative Criteria for Association Rule Mining, 2000. Cisco Systems, Inc.

22. **Di Wang**, 1997. Mobile Agents for Information Retrieval.

23. **Shane Konsella**, 1996. Trie Compaction Using Genetic Algorithms. Initial Employment: Hewlett-Packard.

24. **Karthik Balakrishnan**, 1993. Faster Learning Approximations of Backpropagation by Handling Flat-Spots. Continued as a Ph.D. student.

25. **Jayathi Janakiraman**, 1993. Adaptive Learning Rate for Increasing Learning Speed in Backpropagation Networks. Initial Empoyment: Motorola.

26. **Priyamvadha Thambu**, 1993. Automated Knowledge-Base Consistency Maintenance in an Evolving Intelligent Advisory System. Initial Employment: Inference Corporation.

27. **Rajesh Parekh**, 1993. Efficient Learning of Regular Languages Using Teacher-Supplied Positive Examples and Learner-Generated Queries. Continued as a Ph.D. student.

28. **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.	Scott Boyken	Bioinformatics and Computational Biology	In progress
3.	Haitao Cheng	Bioinformatics and Computational Biology	In progress
4.	Ataur Katebi	Bioinformatics and Computational Biology	In progress
5.	Alexei Kroujiline	Economics	In progress
6.	Jivko Sinapov	Computer Science	In progress
7.	Ru He	Computer Science	In progress
8.	Chris Strasburg	Computer Science	In progress
9.	Zack Oster	Computer Science	In progress
10.	Xia Zhang	Bioinformatics and Computational Biology	In progress
11.	Yao Fu	Bioinformatics and Computational Biology	In progress
12.	Erin Philips	Human-Computer Interaction	In progress
13.	Oliver Couture	Genetics	2011
14.	Tu-Liang Lin	Computer Science	2011
15.	Brian Patterson	Computer Science	2011
16.	Aimin Yan	Bioinformatics and Computational Biology	2011
17.	Deepak Reyon	Bioinformatics and Computational Biology	2011
18.	Scott Broderick	Material Science and Engineering	2009
19.	Changsung Kang	Computer Science	2008
20.	Jeff Sander	Bioinformatics and Computational Biology	2008
21.	Lei Yang	Bioinformatics and Computational Biology	2008
22.	Facundo Bromberg	Computer Science	2007
23.	Jae-Hyung Lee	Bioinformatics and Computational Biology	2007
24.	Yu Cao	Computer Science	2007
25.	Laura Hecker	Neuroscience	2007
26.	Oleksiy Atramentov	Physics	2006
27.	Xiaonan Li	Industrial and Manufacturing Systems Engineering	2006
28.	Di Wu	Bioinformatics and Computational Biology	2006
29.	Yungok Ihm	Bioinformatics and Computational Biology	2004
30.	Cizhiong Zhang,	Bioinformatics and Computational Biology	2004
31.	Zhong Zhang,	Electrical and Computer Engineering	2004
32.	Haibo Cao	Physics	2003
33.	Brooke Peterson	Genetics	2003
34.	Marybeth Gurski	Computer Science	2001

35.	Guy Helmer	Computer Science	2001
36.	Robi Polikar	Electrical and Computer Engineering	2001
37.	Vincent Van Acker	Electrical and Computer Engineering	2000
38.	Chun-Fu Chen	Economics	1999
39.	Victoria Bascunana	Chemical Engineering	1999
40.	Guozhong Zhou	Electrical and Computer Engineering	1998
41.	Cheng-Chi Tai	Electrical and Computer Engineering	1998
42.	James Lathrop	Computer Science	1997
43.	Krishna Dhara	Computer Science	1997
44.	Babak Fourouraghi	Computer Science	1995
45.	Timothy Wahls	Computer Science	1995
46.	Chang-Chun Tsai	Industrial and Manufacturing Engineering	1995
47.	Sonmez Rifat	Civil and Constructional Engineering	1995
48.	Richa Agrawala	Computer Science	1994
49.	Bamshad Mobasher	Computer Science	1994
50.	Hun Kang	Electrical and Computer Engg.	1993

M.S. Committees

1.	Sandeep Krishnan	Computer Science	2009
2.	Zack Oster	Computer Science	2009
3.	Georgi Batinov	Economics	2007
4.	Inya Nlenanya	Agricultural Engineering	2005
5.	Jing Xu	Psychology	2005
6.	Suxing Cheng	Computer Science	2005
7.	Kyongryun Lee	Computer Science	2005
8.	Haitao Cheng	Computer Science	2004
9.	Patricia Lonosky	Genetics	2002
10.	Mallika Bachan	Statistics	2002
11.	Melinda Vander Velden	Electrical and Computer Engineering	2002
12.	Jeremy Patterson	Computer Science	2001
13.	Sa Lin	Computer Science	2001
14.	Vijay Viswanathan	Electrical and Computer Engineering	2001
15.	Fengmei Liu	Computer Science	2001
16.	Xinhua Dong	Computer Science	2001
17.	Mark Slagell	Computer Science	2001
18.	Hao Dong	Computer Science	2001
19.	Jun Li	Computer Science	2001
20.	Thai-Tin Huang	Computer Science	2000
21.	Ran Liu	Computer Science	2000
22.	Xumei Lu	Computer Science	2000
23.	Nanchang Yang	Computer Science	2000
24.	Peng Han	Botany	1999
25.	Jeffrey Yakey	Computer Science	1999
26.	Sunitha Kothapalli	Electrical Engineering	2000

27.	Raghunandan Havaldar	Computer Science	1998
28.	Guy Helmer	Computer Science	1998
29.	Ngee Jenn Lee	Mechanical Engineering	1998
30.	Laura Nelson	Computer Science	1998
31.	Jibin Xiang	Computer Science	1998
32.	Venkat Naganathan	Computer Science	1998
33.	Dean Stevens	Computer Science	1998
34.	Prashant Pai	Computer Science	1998
35.	Abhinav Rawat	Nuclear Engineering	1998
36.	Rishi Nayar	Computer Science	1997
37.	Marcie Goodman	Computer Science	1997
38.	Jon Schultze-Hewett	Computer Science	1997
39.	Chin Khor	Mechanical Engineering	1997
40.	Qiang-lin Zhao	Computer Science	1996
41.	Chi-Chuan Chen	Agricultural Engineering	1996
42.	Mahesh Subramaniam	Computer Science	1996
43.	Glen Holt	Computer Science	1996
44.	Niranjan Vaidya	Computer Science	1996
45.	Thirumalai Anandapillai	Industrial and Manufacturing Engineering	1995
46.	Thomas DeWulf	Electrical Engineering	1994
47.	Raghav Trivedi	Computer Science	1994
48.	Arun Barboza	Computer Science	1994
49.	Brian Schmidt	Electrical and Computer Engineering	1993
50.	Brian Peterson	Computer Science	1992
51.	Salim Chandani	Industrial and Manufacturing Engineering	1992
52.	Prerana Vaidya	Computer Science	1991
53.	James Wittry	Computer Science	1991
54.	Srinivas Boddu	Electrical Engineering	1991
55.	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. Letao Qi, Iowa State University, 2012-2013.
2. Lionel Barrow, Bard College, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2010
3. Kim Eue, Iowa State University, Summer Research Student, 2010
4. Kiran Gustavson, Iowa State University, Summer Research Student, 2010
5. Ryan Pendergast, College of the Holy Cross, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2010
6. Marjie Volk, Iowa State University, Undergraduate Research Assistant, 2010-2011

7. Katie Wilkins, Case Western Reserve University, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2009

8. Casey Oliver, Penn State University, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2009

9. Susan Koons, Texas A&M University, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2009.

10. Kevin Yang, Cornell University, Summer Research Student, 2008.

11. Tyson Williams, Undergraduate Research Assistant, 2007-2008

12. David Gemperline, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2007.

13. Keith Callenberg, Participant, Summer Institute in Bioinformatics and Computational Systems Biology, 2007.

14. Remy Younes, Undergraduate Research Assistant. Topics in Data Integration. 2007.

15. Matt Miller, Topics in Learning Classifiers from Distributed Data, 2006-2007.

16. John Leacox, Topics in Data Integration, 2006.

17. Peter Wong, Topics in Collaborative Ontology Development, 2006.

18. Ryan Bruce (2004), Topics in Bioinformatics

19. Cody Pfau (2003). Topics in Data Mining.

20. Amy Nienaber (2003). Computational Discovery of Protein-Protein Interactions

21. Matthew Beard (2003). Computational Discovery of Protein-Protein Interactions

22. Diane Schroeder (2001-2002) Data Mining Approaches to Discovery of Protein Sequence Function *Relationships* (Graduate School: Stanford University)

23. Kent Vander Velden (1998-1999), Protein Structure Prediction. (Graduate School: Bioinformatics, Iowa State University)

24. Jeremy Ludwig (1996-1997), Topics in Neural Computing. (Graduate School: Intelligent Systems, University of Pittsburgh).

25. David DeYoe (1996-1997), Topics in Cognitive Modelling.

26. Carl Pecinovsky (1996-1997), Constructive Neural Network Learning Algorithms.

27. Brian Walenz, Topics in Genetic Algorithms (Graduate School: Computer Science, University of New Mexico)

28. Gabriel Ki (1996-), Situated Robotics.

29. Todd Lindsey (1995-96), Constructive Neural Network Learning Algorithms.

30. Jouko Rytilahti (1994), Explorations in Evolutionary Algorithms.

31. Leigh McMullen (1993), Adaptive Game-Playing Programs.

Mentor, Freshman Honors Study

1. Marjie Volk (2010) Topics in Machine Learning.

2. Jeffrey Schroeder (1997) Topics in Artificial Intelligence.

3. Matthew Potter (1997) Topics in Artificial Intelligence.

4. Brian George (1994), Topics in Neural Computing.

5. Adam Johnson (1994), Topics in Neural Computing.

6. 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. Professor Jihoon Yang, Sogang University, Korea, 2008.

2. Oswaldo Velesz-Langs, Madrid Polytechnic University, Madrid, Spain. 2003.

3. Professor Yigon Kim (on Sabbatical from Yosu National University, Yosu, Korea 2000-2001) Data Mining and Knowledge Discovery.

4. Professor Mok Dong Chung (on Sabbatical from Pukyong National University, Korea, 1999-2000) Agent-based systems and knowledge-based systems.

5. Olivier Bousquet (from Ecole Polytechnique, France, 1997) Topics in Cognitive Modeling and Robotics.

6. Codrin Nichitiu (from ENS Lyon, France, Summer 1996), Topics in Machine Learning.

7. Dimitri Kotchetkov (Visitor from Ukraine, Summer 1996), Topics in Robotics.

8. 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 20 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 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. 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 developed and 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 have developed and co-taught 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, among the first of its kind, was offered for the first time at ISU in the spring of 2008.

In addition to the regular courses, current research topics are explored in depth in research seminars, which I have organized or co-organized with the help and active participation of 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, which has been 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, data mining, big data analytics, 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, data analytics, 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.

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.

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) drawn from across all areas of science, engineering, and the humanities.

Brief Description of Recent Courses Developed and Taught

- **Principles of Artificial Intelligence**. Com S 572, Graduate. Specification, design, implementation, and selected applications of intelligent software agents and multi-agent systems. Computational models of intelligent behavior, including problem solving, knowledge representation, reasoning, planning, decision making, learning, perception, action, communication and interaction. Reactive, deliverative, rational, adaptive, learning and communicative agents. Artificial intelligence programming. Graduate credit requires a research project and a written report. Oral and written reports. Additional information can be found at http://www.cs.iastate.edu/~cs572/

- **Machine Learning**. Com S 573, Graduate. Algorithmic models of learning. Design, analysis, implementation and applications of learning algorithms. Learning of concepts, classification rules, functions, relations, grammars, probability distributions, value functions, models, skills, behaviors and programs. Agents that learn from observation, examples, instruction, induction, deduction, reinforcement and interaction. Computational learning theory. Data mining and knowledge discovery using artificial neural networks, support vector machines, decision trees, Bayesian networks, association rules, dimensionality reduction, feature selection and visualization. Learning from heterogeneous, distributed, dynamic data and knowledge sources. Learning in multi-agent systems. Selected applications in automated knowledge acquisition, pattern recognition, program synthesis, bioinformatics and Internet-based information systems. Oral and written reports. Additional information can be found at http://www.cs.iastate.edu/~cs573x

- **Computational Functional Genomics and Systems Biology.** BCB 570, Graduate. Algorithmic and statistical approaches in computational functional genomics and systems biology. Analysis of high throughput gene expression, proteomics, and other datasets obtained using system-wide measurements. Topological analysis, module discovery, and comparative analysis of gene and protein networks. Modeling, analysis, simulation and inference of transcriptional regulatory modules and networks, protein-protein interaction networks, metabolic networks, cells and systems: Dynamic systems, Boolean, and probabilistic models. Ontology-driven, network based, and probabilistic approaches to information integration. Additional details can be found at: http://groups.google.com/group/bcb570_2010?hl=en

- **Intelligent Multiagent Systems.** Com S 574, Graduate. Specification, design, implementation, and applications of multi-agent systems. Intelligent agent architectures; infrastructures, languages and tools for design and implementation of distributed multi-agent systems; Multi-agent organizations, communication, interaction, cooperation, team formation, negotiation, competition, and learning. Selected topics in decision theory, game theory, contract theory, bargaining theory, auction theory, and organizational theory. Selected topics in knowledge representation and ontologies. Agent-based systems and the Semantic Web. Applications in distributed intelligent information networks for information retrieval, information integration, inference, and discovery from heterogeneous, autonomous, distributed, dynamic information sources.

- **Advanced Topics in Computational Models of Learning.** Com S 672, Graduate. Selected topics in Computational Learning Theory (PAC learning, Sample complexity, VC Dimension, Occam Learning, Boosting, active learning, Kolomogorov Complexity, Learning under helpful distributions, Mistake Bound Analysis). Selected topics in Bayesian and Information Theoretic Models (ML, MAP, MDL, MML). Advanced statistical methods for machine learning. Selected topics in reinforcement learning.

- **Advanced Topics in Computational Intelligence.** Com S 673, Graduate. Advanced applications of artificial intelligence in bioinformatics, distributed intelligent information networks and the Semantic Web. Selected topics in distributed learning, incremental learning, multi-task learning, multi-strategy learning; Graphical models, multi-relational learning, and causal inference; statistical natural language processing; modeling the internet and the web; automated scientific discovery; neural and cognitive modeling.

- **Artificial Intelligence Graduate Research Seminar** Com S 610. Topics vary. Student-led discussion of research articles of current interest. Recent topics have included advanced topics in machine learning and knowledge

representation. Offered every semester.

- **Principles of Artificial Intelligence**. Com S 472, Undergraduate. Specification, design, implementation, and selected applications of intelligent software agents and multi-agent systems. Computational models of intelligent behavior, including problem solving, knowledge representation, reasoning, planning, decision making, learning, perception, action, communication and interaction. Reactive, deliverative, rational, adaptive, learning and communicative agents. Artificial intelligence programming. Graduate credit requires a research project and a written report. Oral and written reports.
- **Elements of Neural Computation.** Com S 474, Undergraduate. Introduction to theory and applications of neural computation and computational neuroscience. Computational models of neurons and networks of neurons. Neural architectures for associative memory, knowledge representation, inference, pattern classification, function approximation, stochastic search, decision making, and behavior. Neural architectures and algorithms for learning including perceptions, support vector machines, kernel methods, bayesian learning, instance based learning, reinforcement learning, unsupervised learning, and related techniques. Applications in Artificial Intelligence and cognitive and neural modeling. Hands-on experience is emphasized through the use of simulation tools and laboratory projects. Oral and written reports. Additional details can be found at: http://www.cs.iastate.edu/~cs474/

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

2011-2012	Board of Directors, ACM Special Interest Group on Bioinformatics
2012-	Interagency Working Group on Multi-Scale Modeling
2010	ICiS Workshop on Integrating, Representing, and Reasoning with Human Knowledge, Snowbird, Utah
2010	AFOSFR Workshop on Dynamic Data-Driven Application Systems, Arlington, VA
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

2012	Workshop Program Chair	ACM Conference on Bioinformatics and Computational Biology
2011	Co-Chair	ACM Immunoinformatics and Computational Immunology Workshop
2011	Workshop Program Chair	ACM Conference on Bioinformatics and Computational Biology
2011	Co-Chair	ACM Immunoinformatics and Computational Immunology Workshop

2010	Co-Chair	ACM Immunoinformatics and Computational Immunology Workshop
2010	Area Chair	IEEE Conference on Tools with Artificial Intelligence (ICTAI 2010)
2009	Sponsorship co-chair	SIAM Conference on Data Mining (SDM 2009)
2009	Area Chair	IEEE Conference on Tools with Artificial Intelligence (ICTAI 2009)
2009	Organizer and Chair	Doctoral Student Forum, SIAM Conference on Data Mining
2008	Vice Program Chair	International Semantic Web Conference (ISWC 2008)
2008	Track Chair	SIAM Conference on Data Mining (SDM 2008)
2007	Organizer and Co-Chair	IJCAI Workshop on Semantic Web for Collaborative Knowledge Acquisition, IJCAI-2007 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

2008	Advisory Board Member	Workshop on Web Service Composition and Adaptation, WSCA 2008
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

2013	Conference on Data Integration in Life Sciences (DILS 2013)
2013	IEEE Big Data Congress (BigData 2013)
2013	IEEE Conference on Bioinformatics and Biomedicine (IEEE BIBM 2013)
2013	IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)
2013	14th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL 2013)
2013	International Symp. on Network Enabled Health Informatics, Biomedicine and Bioinformatics (HI-BI-BI 2013)
2013	Fourth Immunoinformatics and Computational Immunology Workshop (ICIW 2013)
2013	International Conference on Healthcare Informatics (ICHI 2013)
2013	International Conference on Bioinformatics Models, Methods and Algorithms (Bioinformatics 2013)
2013	International Conference on Neural Computation Theory and Applications (NCTA 2013)
2012	International Semantic Web Conference (ISWC 2012)
2012	AAAI Conference on Artificial Intelligence (AAAI 2012)
2012	Conference on Data Integration in Life Sciences (DILS 2012)
2012	IEEE Conference on Tools with Artificial Intelligence (ICTAI 2012)
2012	International Conference on Bioinformatics Models, Methods, and Algorithms (Bioinformatics 2012)
2012	International Conference on Health Informatics (IHI 2012)
2012	IEEE Conference on Biomedical Computing 2012
2012	SIAM Data Mining Conference (SDM 2012)
2011	SIAM Data Mining Conference (SDM 2011)
2011	ACM Conference on Bioinformatics and Computational Biology (ACM-BCB 2011)
2011	Intelligent Systems in Computational Biology (ISMB / ECCB 2011)
2011	Pacific Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2011)
2011	IEEE Conference on Tools with Artificial Intelligence (ICTAI 2011)
2011	ACM/WIC/IEEE Conference on Intelligent Agent Technology (IAT 2011)
2011	ACM International Health Informatics Symposium (IHI 2011)
2010	National Conference on Artificial Intelligence (AAAI 2010)
2010	International Conference on Machine Learning (ICML 2010)
2010	IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2010)
2010	ACM International Health Informatics Symposium (IHI 2010)
2010	ACM/WIC/IEEE Conference on Intelligent Agent Technology (IAT 2010)
2010	ACM Conference on Bioinformatics and Computational Biology (ACM-BCB 2010)
2010	International Symposium on Bioinformatics Research and Applications (ISBRA 2010)
2010	Workshop on Modular Ontologies (WOMO 2010)
2010	SIAM Data Mining Conference (SDM 2010)
2010	Pacific Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2010)
2010	International Conference on Advances in Semantic Processing (SEMAPRO 2010)
2009	Intelligent Systems in Molecular Biology (ISMB 2009)
2009	ACM/WIC/IEEE Conference on Intelligent Agent Technology (IAT 2009)
2009	IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2009)
2009	International Conference on Intelligent Data Engineering and Automated Learning (IDEAL 2009)
2009	International Symposium on Bioinformatics Research and Applications (ISBRA 2009)
2009	International Conference on Advances in Semantic Processing (SEMAPRO 2009)
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

2005-2012	Director, Center for Computational Intelligence, Learning and Discovery		ISU
2010-2011	Member, Supervisory Committee, Bioinformatics and Computational Biology Graduate Program		ISU
2009-2010	Director of Research	Computer Science	ISU
2009-2010	Director, Center for Computational Intelligence, Learning and Discovery		ISU
2009-2010	Member, Research Computing Council		ISU
2009-2010	Chair, Promotion & Tenure Steering Committee	Computer Science	ISU
2009-2010	Member, Supervisory Committee, Bioinformatics and Computational Biology Undergraduate Program		ISU
2009-2010	Member, Supervisory Committee, Bioinformatics and Computational Biology Graduate Program		ISU
2009-2010	Director of Research	Computer Science	ISU
2008-2009	Member, Research Computing Council		ISU
2008-2009	Member, e-Science Steering Committee		ISU
2008-2009	Member, Promotion & Tenure Steering Committee	Computer Science	ISU
2008-2009	Member, Supervisory Committee, Bioinformatics and Computational Biology Undergraduate Program		ISU
2008-2009	Member, Supervisory Committee, Bioinformatics and Computational Biology Graduate Program		ISU
2007-2008	Member, Supervisory Committee, Bioinformatics and Computational Biology Undergraduate Program		ISU
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, Professional and Scientific Staff Hiring Committee (ad hoc)	Computer Science	ISU
2007-2008	Director of Research	Computer Science	ISU
2006-2007	Member, Graduate Committee	Computer Science	ISU
2006-2007	Chair, Research Infrastructure Committee	Computer Science	ISU
2006-2007	Director of Research	Computer Science	ISU
2006-2007	Director, Center for Computational Intelligence, Learning and Discovery	ISU	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	Director of Research	Computer Science	ISU
2005-2006	Director, Center for Computational Intelligence, Learning and Discovery	ISU	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