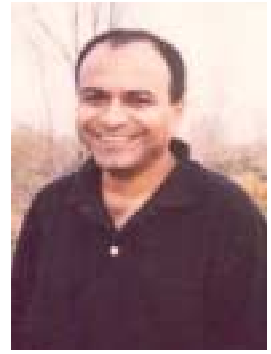
BIOGRAPHICAL SKETCH

VASANT G. HONAVAR



Current Position
Professor of Computer Science
Professor and Associate Chair, Bioinformatics and Computational Biology
Director, Artificial Intelligence Research Laboratory
Department of Computer Science
226 Atanasoff Hall, Iowa State University
Ames, Iowa 50011-1040.
honavar@cs.iastate.edu
Voice: 515 294 1098
Fax: 515 294 0258
http://www.cs.iastate.edu/~honavar/

EDUCATION:

- **Ph.D. in Computer Science and Cognitive Science (Artificial Intelligence)**, 1990; Thesis advisor: Professor Leonard Uhr; University of Wisconsin-Madison, Madison, Wisconsin, U.S.A.
- **Course on Neural Networks Foundations and Applications**, 1990; Boston University.
- **Master of Science in Computer Sciences**, 1989; University of Wisconsin-Madison, Madison, Wisconsin, U.S.A.
- **Summer Institute in Cognitive Neuroscience**, 1989; Dartmouth College.
- **Connectionist Models Summer School**, 1988; Carnegie Mellon University.
- **Summer Workshop in Parallel Computing**, 1988; Argonne National Laboratory.
- **Master of Science in Electrical and Computer Engineering**, Drexel University, Philadelphia, Pennsylvania, U.S.A.
- Bachelor of Engineering in Electronic Engineering, Bangalore University, Bangalore, India.

PROFESSIONAL EXPERIENCE

- **Professor** (2001-); Department of Computer Science, Iowa State University, Ames, Iowa. http://www.cs.iastate.edu/
- **Director** (1990 -) Artificial Intelligence Laboratory, Dept. of Computer Science, Iowa State University. http://www.cs.iastate.edu/~honavar/aigroup.html
- **Visiting Professor** (2002); Department of Biostatistics and Medical Informatics (http://www.biostat.wisc.edu/) and Computer Sciences Department (http://www.cs.wisc.edu/), University of Wisconsin, Madison, Wisconsin.
- **Co-Editor-in-Chief** (1999-); *Journal of Cognitive Systems Research* Published by Elsevier. **Member of Editorial Board**, *Machine Learning Journal* (2002-2005). **Co-Editor**, *Machine Learning Journal* Special Issue on *Automata induction, Grammar inference, and Language acquisition,* **Member of Editorial Board**, *Encyclopedia of Cognitive Science*.

- **Director of Research** (2002-), Department of Computer Science, Iowa State University.
- **Professor** (2001-); Graduate Program in Bioinformatics and Computational Biology, Iowa State University, Ames, Iowa. http://www.bcb.iastate.edu/
- **Associate Chair** (2001-); Graduate Program in Bioinformatics and Computational Biology, Iowa State University, Ames, Iowa. http://www.bcb.iastate.edu/
- **Project Director** (Bioinformatics) (1999-); Computational Molecular Biology (CMB) Training Program, Iowa State University, Ames, Iowa. http://www.bioinformatics.iastate.edu/IGERT/
- **Professor** (2001-); Information Assurance Graduate Program, Iowa State University, Ames, Iowa. http://www.issl.iastate.edu/infas.html
- **Professor** (2001-); Graduate Program in Neuroscience, Iowa State University, Ames, Iowa. http://www.grad-college.iastate.edu/neuroscience/neuromain.html
- **Program Chair** Conference on Computational Biology and Genome Informatics, Durham, North Carolina, USA. http://www.cs.iastate.edu/~cbgi/cbgi.html
- **Member** (1999-); Center for Bioinformatics and Biological Statistics, Iowa State University, Ames, Iowa. http://www.bioinformatics.iastate.edu/
- **Member** (1998-); Virtual Reality Applications Center, Iowa State University, Ames, Iowa. http://www.vrac.iastate.edu/
- **Member** (1999-); Information Security Laboratory, Iowa State University, Ames, Iowa. http://www.issl.org
- **Member** (1998 -); Information Institute, Information Directorate, Air Force Rome Labs, Rome, New York.
- **Visiting Professor** (September 1998 - December 1998); School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- **Associate Professor** (1996 - 2001); Department of Computer Science, Iowa State University, Ames, Iowa. http://www.cs.iastate.edu/
- **Associate Professor** (1999 - 2001); Graduate Program in Bioinformatics and Computational Biology, Iowa State University, Ames, Iowa. http://www.bcb.iastate.edu/
- **Associate Professor** (1996-2001); Graduate Program in Neuroscience, Iowa State University, Ames, Iowa. www.grad-college.iastate.edu/neuroscience/neuromain.html
- **Assistant Professor** (1990 - 1996); Department of Computer Science, Iowa State University, Ames, Iowa.
- **Assistant Professor** (1992- 1996); Graduate Program in Neuroscience, Iowa State University, Ames, Iowa. www.grad-college.iastate.edu/neuroscience/neuromain.html
- **Consultant** on Intelligent Systems technologies for several industrial and commercial organizations.

- **Referee** for journals (including Connection Science, Applied Intelligence, IEEE PAMI, TKDE, TEC, TNN; Neural Computation, Neural Networks, Information Sciences)

- **Program committee member** for conferences (including WCNN, GP, IJCNN, ICML, ICGI, AAAI).

- **Organizer and Co-Chair**, Workshop on Learning from Sequential and Temporal Data, International Conference on Machine Learning, Stanford, CA, 2000.

- **Conference Program Chair**, Fourth International Colloquium on Grammatical Inference, Ames, Iowa, 1998; Conference on Computational Biology and Genome Informatics (CBGI-02), 2002. Artificial Life, Agents, and Adaptive Behavior Track, Genetic and Evolutionary Computing Conference, 1999, 2002.

- **Proposal reviewer** for National Science Foundation, Department of Energy, National Scientific and Engineering Research Council (Canada), Irish National Science Foundation, Dutch Research Council, United States Civilian Research and Development Foundation.

- **Invited Participant**, Connectionist Models Summer School, Carnegie Mellon University, 1988; Summer Workshop on Parallel Computing, Argonne National Laboratory, 1988; Summer Institute on Cognitive Neuroscience, Dartmouth College, 1989; NSF Workshop on Decision-Based Design, Sacramento, 1997; Snowbird Workshop on Learning, 1999.

- **Working Groups, Panels, and Advisory Boards**: NSF CISE IIS Review Panel, 2002; NIH Special Study Section (Bioinformatics) Member, 2001, 2002; CISE Research Instrumentation Proposal Review Panel, 1994; NSF Integrative Graduate Education and Research Training (IGERT) Proposal Review Panel, 1999; Intelligent Knowledge-Based Systems Working Group, Information Institute, Information Directorate, Air Force Research Labs, 1998; Intelligent Agents Working Group, Information Institute, Information Directorate, Air Force Research Labs, 1999; Member of Advisory Board, Multi-University EPSRC sponsored multi-university research project on Emergent Computational Neural Network Architectures (Emernet), Universities of Edinburgh, York, and Sunderland, United Kingdom, 1999-.

- **Co-Chair**, NIPS Post-conference workshop on Symbolic and Subsymbolic Information Processing in Biological Neural Systems, 1992; ICML workshop on automata induction, grammar inference, and language acquisition, 1997; AAAI Workshop on Computation with Neural Systems, 1999.

- **Research Assistant.** Summer 1987-Spring 1990. Responsible for the design, implementation, and performance evaluation of parallel algorithms for computer vision and constructive neural network learning algorithms for pattern classification. Computer Sciences Department, University of Wisconsin-Madison.

- **Teaching Assistant.** Fall 1985 - Spring 1987. Responsible for courses in Computer Architecture, Artificial Intelligence, Computer Vision, and Computer Programming. Computer Sciences Department, University of Wisconsin-Madison.

- **Teaching Assistant.** Fall 1984 - Spring 1985. Department of Electrical and Computer Engineering, University of Wisconsin-Madison. Responsible for courses in Circuits and Systems, and Digital Logic Design.

- **Research Assistant.** Summer 1983 - Spring 1984. Department of Electrical and Computer Engineering, Drexel University. Responsible for design, implementation, and evaluation of image processing software for automated analysis of dark field autoradiographs and inspection of VLSI chips.

RESPONSIBILITIES

My primary responsibilities at Iowa State University include Teaching (50%), Research (50%) and Service.

TEACHING AND CURRICULUM DEVELOPMENT

I typically teach 3 courses a year (not counting seminar courses). I have developed and taught the following courses in Computer Science:

- Principles of Artificial Intelligence (Dual-listed as ComS 572 and ComS 472; Taught every Fall)
- Machine Learning (ComS 573X Taught every Spring)
- Computational Models of Learning (ComS 672 Taught in alternate Spring Semesters)
- Distributed Intelligent Multi-Agent Systems (ComS 673; Taught in Alternate Spring Semesters).
- Intelligent Multi-Agent Systems (ComS 574;To be taught in Alternate spring semesters)
- Neural Computation (ComS 474; Taught every Spring between 1994 and 1998)
- Artificial Intelligence Research Seminar (ComS 610; Offered every semester).

Teaching Statement

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.* My teaching and curriculum development activities in computer science complement, and are sustained by, my research in artificial intelligence, bioinformatics and computational biology, and related areas.

Over the past ten 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 models of learning, distributed intelligent information networks, intelligent systems in molecular biology, and complex adaptive systems.

I have played an active role in the establishment of an interdepartmental graduate program in Bioinformatics and Computational Biology, with funding from the National Science Foundation through an Integrative Graduate Education and Research Training (IGERT) award. I have helped establish an Interdepartmental Graduate Minor in Complex Adaptive Systems. I also serve on the faculty of the Interdepartmental Graduate Program in Neuroscience.

The undergraduate courses developed by me 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.

Upon successful completion of an undergraduate course in artificial intelligence, the students have a good understanding of some of the key problems in, and approaches to, the design and analysis of intelligent agents. They are also able to develop and evaluate intelligent agents and multi-agent systems (e.g., automated problem solvers, and learning agents) of moderate complexity. Similarly, upon successful completion of the undergraduate course on neural computation and machine learning, students have good grasp of key computational models of neurons and networks of neurons (or neural networks). They are familiar with several key neural network and other machine learning algorithms (e.g., decision tree learning, reinforcement learning, etc.) and are able to apply these tools for building intelligent agents for applications in science and engineering. 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 also make an attempt to present such knowledge in the broader context of computer science as an intellectual discipline.

The graduate courses are designed to introduce students to fundamental research problems and approaches in intelligent agents and multiagent systems, machine learning, data mining and knowledge discovery, computational models of learning, neural computation, intelligent systems in molecular biology, and related areas. The topics covered in these courses vary from year to year. The emphasis of such courses is on the development of the graduate students into creative thinkers and problem-solvers, be it in academic research or advanced technology development. In addition to the regular courses, current research topics are explored in depth in research seminars which I organize or coorganize (with other faculty in Computer Science and Bioinformatics and Computational Biology) with the help of my graduate students.

The nature of the material taught in my courses requires a delicate balance between theory and experimentation. Adequate laboratory facilities are essential to support experiments, exercises, and projects that enhance the students' understanding of key concepts covered in the lectures. For this reason, I have had to put in considerable effort (with the help of my graduate students) to develop adequate laboratory facilities to support instruction in artificial intelligence and related areas.

It is my belief that strong communication skills are essential for the success of students in both academia and industry. For this reason, 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 problemsolving. 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 produced results that were eventually published in refereed national and international conferences.

In a fast-paced field like computer science in general and artificial intelligence 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. My research in artificial intelligence and bioinformatics and computational biology help me shape the course materials towards this end.

The syllabi and other materials for the courses (including lecture notes) that I have developed and taught at Iowa State University during 1990-2000 are available at http://www.cs.iastate.edu/honavar/courses.html .

GRADUATE, POSTDOCTORAL AND UNDERGRADUATE RESEARCH SUPERVISION

Student Mentoring Statement

I am primarily interested in exceptional Ph.D. students with diverse backgrounds (ranging from very theoretical to very experimental) whose research interests match the research foci of my lab. Occasionally, I accept highly qualified M.S. students and undergraduates interested in research. Students in my group benefit from strong mentoring and close interaction on a daily basis within a collaborative research environment designed to prepare each student for a productive and rewarding research career. Research-based training in our graduate programs in general, and my research group in particular, emphasizes: identification of fundamental research problems; development of creative and innovative solutions; and dissemination of research results to the community through publications in top journals and conferences. In addition to providing strong technical skills in the relevant research area(s), my group also fosters the development of strong writing and presentation skills.

Graduate students who join my lab typically have a broad-based training in Computer Science or a closely related discipline. Some of them also have training in biological sciences. Many of them 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. Close interaction among students in the group through research seminars and collaborative research projects is encouraged.

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 subdiscipline (e.g., Machine Learning) when necessary to solve a research problem. 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.

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).

Major Professor, Ph.D.

Current Ph.D. Students

1. Carson Andorf, In progress. Areas: Bioinformatics and Computational Biology; Ontology Assisted Data Mining and Data Visualization for Characterization of Macromolecular Structure-Function Relationships.

2. Anna Atramentov, In progress. Areas: Artificial Intelligence, Computational Learning Theory, Machine Learning, Data Mining.

3. Facundo Bromberg, In progress. Areas: Artificial Intelligence, Complex Systems, Intelligent Agents and Multi-Agent Systems.

4. Doina Caragea, In progress. Areas: Machine Learning, Bioinformatics, Distributed Learning, Incremental Learning, Data Visualization.

5. Bao Jie, In progress. Areas: Machine Learning, Multi-Agent Systems.

6. Jaime Reinoso-Castillo, In progress. Areas: Multi-Agent Systems, Machine Learning, Data Integration, Ontologies.

7. Adrian Silvescu, In progress. Areas: Intelligent Agents, Multi-Agent Systems, Machine Learning, Complex Adaptive Systems, Bioinformatics.

8. Changhui Yan, In progress. Areas: Bioinformatics and Computational Biology; Data Mining and Data Visualization Algorithms for Characterization of Macromolecular Structure-Function Relationships.

9. Feihong Wu, In progress. Areas: Bioinformatics and Computational Biology; Data Mining and Data Visualization Algorithms for Characterization of Macromolecular Structure-Function Relationships.

10. Jun Zhang, In progress. Areas: Bioinformatics, Machine Learning, Ontologies, Data Integration.

11. Dae-Ki Kang, In progress. Areas: Distributed Information Infrastructure.

Ph.D. Graduates

1. Jihoon Yang, *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.

2. Karthik Balakrishnan, *Biologically Inspired Computational Structures and Processes for Autonomous Agents and Robots*, 1998. Initial Employment: Senior Research Scientist, Datamining Group, Allstate Research and Planning Center, Menlo Park, CA. USA. Current Employment: Data Mining Group, Fireman's Insurance, Menlo Park, CA.

3. Rajesh Parekh, *Machine Learning of Automata and Neural Network Pattern Classifiers*, 1998. Initial Employment: Senior Research Scientist, Datamining Group, Allstate Research

and Planning Center, Menlo Park, CA. USA. Current Employment: Senior Research Scientist, Data Mining, Blue Martini Software, Menlo Park, CA.

4. Chun-Hsien Chen, *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: Assistant Professor, Department of Information Management, Chang Gung University, Taiwan.

5. Armin Mikler (joint supervision with Johnny Wong), *Quo Vadis - A Framework for Intelligent Routing in Large Communication Networks*, 1995. Initial Employment: Research Associate, Scalable Computing Laboratory, DOE Ames Lab, Ames, Iowa. Current Employment: Assistant Professor of Computer Science, University of North Texas, Denton, TX, USA,

Major Professor, M.S.

Current M.S. Students

1. Zhong Gao, Bioinformatics and Computational Biology. *Computational Discovery of Sequence Correlates of Structural Features of Proteins with Applications to Function Classification and Structure Prediction.* In progress.
2. Sergio Ferrero, Computer Science. In progress.
3. Jyotishman Pathak, Computer Science. In progress.

M.S. Graduates

1. Jaime Reinoso-Castillo, *Ontology-Driven Query-Centric Information Integration from Heterogeneous, Distributed, Autonmous Data Sources for Computer Assisted Scientific Discovery.* 2002.

2. Hector Leiva, *Learning Classifiers from Relational Data*. 2002.

3. Ziaosi Zhang, *Identification of Functionally Related Genes from Gene Expression Data*. 2002.

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

5. 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.

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

7. Dake Wang, *Data-Driven Generation of Decision Trees for Motif-Based Assignment of Protein Sequences to Functional Families.*, 2001. Initial Employment: Lumicyte, Inc.

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

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

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

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

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

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

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

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

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

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

18. Richard Spartz, 1992. *Speeding Up Backpropagation Using Expected Source Values.* Initial Employment: IBM.

Host, Visiting Researchers

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

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

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

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

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

6. Vadim Kirillov (Fullbright Scholar from Ukraine, 1995), *Constraint-based Reasoning under uncertainty*.

Supervisor, Undergraduate Honors Project

1. Eric Barsness (1993), *An Object-Oriented Implementation of a Genetic Algorithms Testbed.*

2. Daniel Graves (1992), *Parallel Architectures for Artificial Intelligence.*

Supervisor, Undergraduate Research

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

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

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

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

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

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

7. Gabriel Ki (1996-), *Situated Robotics.*

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

9. Jouko Rytilahti (1994), *Explorations in Evolutionary Algorithms.*

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

Mentor, Freshman Honors Study

1. Jeffrey Schroeder (1997) *Topics in Artificial Intelligence.*

2. Matthew Potter (1997) *Topics in Artificial Intelligence.*

3. Brian George (1994), *Topics in Neural Computing.*

4. Adam Johnson (1994), *Topics in Neural Computing.*

5. Marcus Ryan (1993), *Topics in Artificial Intelligence.*

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.*

Research Collaborators

1. Venkataramana Ajjarapu, Electrical and Computer Engineering, Iowa State University.
2. Amy Andreotti, Biochemistry and Biophysics, Iowa State University.
3. Karthik Balakrishnan, Obongo, Inc. Menlo Park, CA.
4. Albert Baker, Computer Science, Iowa State University.
5. Thomas Barta, Industrial and Manufacturing Systems Engg., Iowa State University.
6. Philip Becraft, Zoology and Genetics, Iowa State University.
7. Olivier Bousquet, Ecole Normale Superieure des Telecommunications, France.
8. Volker Brendel, Zoology and Genetics, Iowa State University.
9. Chun-Hsien Chen, Industrial Technology Research Institute, Taiwan.
10. Mok-Dong Chung, Computer Science, Pokyang National University, Korea.
11. Di Cook, Statistics, Iowa State University.
12. Drena Dobbs, Zoology and Genetics, Iowa State University.
13. Karin Dorman, Statistics, Iowa State University.
14. David Fernandez-Baca, Computer Science, Iowa State University.
15. Shashi Gadia, Computer Science, Iowa State University.
16. Xun Gu, Agronomy, Iowa State University.
17. Philip Haydon, Zoology and Genetics, Iowa State University.
18. Guy Helmer, Palisade Systems, Inc., Ames, Iowa.
19. Colin de la Higuera, Computer Science, Univ. St. Etienne, France.
20. Sarah Nusser, Statistics, Iowa State University.
21. Kai-Ming Ho, Physics, Iowa State University.
22. Srdija Jeftinija, Veterinary Anatomy, Iowa State University
23. Robert Jernigan, Biochemistry and Biophysics, Iowa State University
24. Wolfgang Kleimann, Mathematics, Iowa State University.
25. Steven LaValle, Computer Science, University of Illinois at Urbana-Champaign.
26. Jan Leach, Plant Pathology, Kansas State University.
27. Gary Leavens, Computer Science, Iowa State University.

28. Hei Leung, Plant Pathology, Consultative Group on International Agricultural Research.
29. Jack Lutz, Computer Science, Iowa State University.
30. G. Manimaran, Electrical and Computer Engineering, Iowa State University.
31. John Mayfield, Zoology and Genetics, Iowa State University.
32. James McCalley, Electrical and Computer Engineering, Iowa State University.
33. Armin Mikler, Computer Science, University of North Texas.
34. Les Miller, Computer Science, Iowa State University.
35. Jamie Morris, Ames Laboratory, Iowa State Univerity.
36. Gavin Naylor, Computer Science, Iowa State University
37. T. Neubert, Structural Biology and Pharmacology, New York University.
38. Codrin Nichitiu, Ecole Normale Superieure de Lyon, France.
39. Gregg Oden, Psychology and Computer Science, University of Iowa.
40. Reid Palmer, USDA, Ames, Iowa.
41. Rajesh Parekh, Allstate Research and Planning Center, Menlo Park, CA.
42. Jo Anne Powell-Coffman, Zoology and Genetics, Iowa State University.
43. Steve Rodermel, Botany, Iowa State University.
44. Pamela Ronald, Plant Pathology, University of California at Davis.
45. Patrick Schnable, Agronomy, Iowa State University.
46. Jude Shavlik, Computer Science, University of Wisconsin-Madison.
47. Gerry Sheble, Electrical and Compuiter Engineering, Iowa State University.
48. Giora Slutzki, Computer Science, Iowa State University.
49. Arun Somani, Electrical and Computer Engineering, Iowa State University.
50. Ron Sun, Computer Science and Psychology, University of Alabama.
51. David Touretzky, Computer Science, Carnegie Mellon University.
52. L. Udpa, Electrical and Computer Engineering, Iowa State University.
53. S. Udpa, Electrical and Computer Engineering, Iowa State University.
54. Leonard Uhr, Computer Science, University of Wisconsin-Madison.
55. Vijay Vittal, Electrical and Computer Engineering, Iowa State University.
56. Cai-Zhong Wang, Ames Laboratory, Iowa State University.

57. Guo-Liang Wang, Plant Pathology, Ohio State University.

58. Heather West-Greenlee, Veterinery Medicine, Iowa State University.

59. Zhijun Wu, Mathematics, Iowa State University

60. Johnny Wong, Computer Science, Iowa State University.

61. Jihoon Yang, Hughes Research Laboratory, Malibu, CA.

Member of Program of Study Committee

Ph.D. Supervisory Committees:

1. Jaekyung Yang (Industrial and Manufacturing Systems Engg.)
2. Guy Helmer (Computer Science, 2001);
3. Curtis Clifton (Computer Science);
4. Olivier Fedrigo (Genetics);
5. Cizhiong Jiang (Genetics);
6. Steven Jenkins (Computer Science);
7. Marybeth Gurski (Computer Science);
8. Brooke Peterson (Molecular Biology and Genetics / Bioinformatics);
9. Vincent Van Acker (Electrical and Computer Engineering, 2000);
10. Chun-Fu Chen (Economics, 1999);
11. Victoria Bascunana (Chemical Engineering, 1999);
12. Guozhong Zhou (Electrical and Computer Engineering, 1998);
13. Cheng-Chi Tai (Electrical and Computer Engineering, 1998);
14. James Lathrop (Computer Science, 1997);
15. Krishna Dhara (Computer Science, 1997);
16. Babak Fourouraghi (Computer Science, 1995);
17. Timothy Wahls (Computer Science, 1995);
18. Chang-Chun Tsai (Industrial and Manufacturing Systems Engineering, 1995);
19. Sonmez Rifat (Civil and Constructional Engineering, 1995);
20. Richa Agrawala (Computer Science, 1994);
21. Bamshad Mobasher (Computer Science, 1994); j

22. Hun Kang, (Electrical and Computer Engineering, 1993).

M.S. Supervisory Committees

1. Patricia Lonosky (Genetics, 2002)
2. Jeremy Patterson (Computer Science, 2001);
3. Sa Lin (Computer Science, 2001);
4. Vijay Viswanathan (Electrical and Computer Engineering, 2001);
5. Fengmei Liu (Computer Science, 2001);
6. Xinhua Dong (Computer Science, 2001);
7. Mark Slagell (Computer Science, 2001);
8. Hao Dong (Computer Science, 2001);
9. Jun Li (Computer Science, 2001);
10. Thai-Tin Huang (Computer Science, 2000);
11. Ran Liu (Computer Science, 2000);
12. Xumei Lu (Computer Science, 2000);
13. Nanchang Yang (Computer Science, 2000);
14. Peng Han (Botany, 1999);
15. Jeffrey Yakey (Computer Science, 1999);
16. Sunitha Kothapalli (Electrical Engineering, 2000);
17. Raghunandan Havaldar (Computer Science, 1998);
18. Guy Helmer (Computer Science, 1998);
19. Ngee Jenn Lee (Mechanical Engineering, 1998);
20. Laura Nelson (Computer Science, 1998);
21. Jibin Xiang (Computer Science, 1998);
22. Venkataraman Naganathan (Computer Science, 1998);
23. Dean Stevens (Computer Science, 1998);
24. Prashant Pai (Computer Science, 1998);
25. Abhinav Rawat (Nuclear Engineering, 1998);
26. Rishi Nayar (Computer Science, 1997);

27. Marcie Goodman (Computer Science, 1997);
28. Jonathan Schultze-Hewett (Computer Science, 1997);
29. Chin Khor (Mechanical Engineering, 1997);
30. Qiang-lin Zhao (Computer Science, 1996);
31. Chi-Chuan Chen (Agricultural Engineering, 1996);
32. Mahesh Subramaniam (Computer Science, 1996);
33. Glen Holt (Computer Science, 1996);
34. Niranjan Vaidya (Computer Science, 1996);
35. Thirumalai Anandapillai (Industrial and Manufacturing Engineering, 1995);
36. Thomas DeWulf, (Electrical Engineering, 1994);
37. Raghav Trivedi (Computer Science, 1994);
38. Arun Barboza, (Computer Science, 1994);
39. Brian Schmidt, (Electrical and Computer Engineering, 1993);
40. Brian Peterson, (Computer Science, 1992);
41. Salim Chandani, (Industrial and Manufacturing Systems Engineering, 1992);
42. Prerana Vaidya, (Computer Science, 1991);
43. James Wittry, (Computer Science, 1991);
44. Srinivas Boddu, (Electrical Engineering, 1991);
45. Simanta Mitra, (Computer Science, 1991);

RESEARCH

Research Statement

I work on a number of basic and applied research problems in artificial intelligence Artificial Intelligence and Cognitive Science (Artificial Neural Networks; Autonomous Robots; Automata Induction; Computational Learning Theory; Computational Organizational Theory; Data Mining, Knowledge Discovery and Visualization; Decision Support Systems; Distributed Knowledge Networks; Intelligent Agents, Mobile Agents, and Multi-Agent Systems; Knowledge Representation and Inference; Machine Learning; Parallel and Distributed AI); Biological Computation (Computational Neuroscience; Evolutionary, Cellular, and Neural Computation); Bioinformatics and Computational Biology (Genetic Regulatory Networks; Protein Structure-Function Prediction; Computational Genomics; Metabolic Pathways; Distributed Knowledge Networks for Bioinformatics); Complex Adaptive Systems; Distributed Information Systems (Distributed Knowledge Networks; Distributed Databases, Mediators, and Data Warehouses; Intelligent Agents, Mobile Agents,and Multi-Agent Systems); Applied AI (Diagnosis;

Bioinformatics; Complex Systems Monitoring and Control; Intrusion Detection). Some of this research is being carried out in collaboration with faculty in Computer Science, Bioinformatics and Computational Biology, and Electrical and Computer Engineering.

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.

More detailed information about some of my current research projects can be found at my laboratory's web page at http://www.cs.iastate.edu/~honavar/aigroup.html

Current Research Grants

- ITR: Algorithms and Software for Knowledge Acquisition from Heterogeneous Distributed Data. National Science Foundation, Vasant Honavar, (2002-2005). $210,000.
- Training Grant, NIH-NSF BBSI Summer Institute in Bioinformatics and Computational Biology - Iowa State University. National Science Foundation, Vasant Honavar (with Volker Brendel, Robert Jernigan, Karin Dorman, and Xun Gu). (2002-2006). $645,000.
- Interactive Visual Overviews of Large, Multi-Dimensional Datasets, *National Science Foundation*, Co-Principal Investigator, 1999-2002, $370,000.
- Graduate Research Fellowships in Bioinformatics and Computational Biology, *Pioneer Hi-Bred, Inc.* 2002-2004. *Major professor,* Adrian Silvesu and Carson Andorf, $80,000.
- IBM Graduate Research Fellowship in Computer Science, IBM Inc., Major Professor, Doina Caragea, 2002-2003, $30,000 (approx.)
- Automated Integration of Condition Monitoring with an Optimized Maintenance Scheduler for Circuit Breakers and Power Transformers. *Power Systems Engineering Research Center (a National Science Foundation Industry-University Research Center)*, 2002-2005. Co-Principal Investigator (2002-2005). $255,000.
- IGERT: Computational Molecular Biology Training Program. *National Science Foundation*, 1999-2004. Co-Principal Investigator, $2,374,597 (plus $1,161,010 in matching funds).
- Center for Computational Inference, Learning, and Discovery (CCILD). Seed Grant. Iowa State University. 2002. $30,000.
- SGER: Distributed Knowledge Networks to Support Security-Economy Decisions in Stressed Electric Power Systems. *National Science Foundation*, 2000-2001, Co-Principal Investigator, $99,999.
- 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.

Past Research Grants

- Constructive Neural Network Learning Algorithms for Pattern Classification, *National Science Foundation*, 1994-1999, Principal Investigator, $111,537 (plus $10,000 in matching funds).
- 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.
- Intelligent Multi-Agent Systems for Intrusion Detection, *National Security Agency*, 1998-2000, Co-Principal Investigator, $199,769.
- Data Mining of Electric Power Usage Data to Develop Customer Profiles. Cooperative Research Proposal. Power Domain, Inc. (2001-2002). Vasant Honavar. $43,639.
- Artificial Intelligence Applications to Power System Management and Control, *Electric Power Research Institute,* Co-principal investigator. 1998-2000, $151,000
- Distributed Knowledge Networks, *John Deere Foundation*, 1999-2001, Principal Investigator, $30,000.
- Development of Algorithmic Approaches to Gene Expression Analysis from Microarray Data, *Carver Foundation*, 2000-2001, Principal Investigator, $25,000.
- Development of Protein Structure Prediction Algorithms. *Carver Foundation*, 1999-2000, Co-Principal Investigator, $25,000.
- Genetic algorithms for protein structure prediction. *Ames Laboratory*, Co-Principal Investigator, 1999-2000. $35,700.
- A Gene-Specific DNA Chip for Exploring Molecular Evolutionary Change, *Carver Foundation*, 1998-1999. Co-Principal Investigator, $17,120.
- Intelligent Diagnosis Systems, *John Deere Foundation*, 1995-1998, Principal Investigator, $30,000.
- Graduate Fellowship (Data Mining and Knowledge Discovery), *IBM Corporation*, 1997-1998, Principal Investigator, $20,800.

Pending Grant Proposals

- Discovering Protein Sequence-Structure-Function Relationships (with Robert Jernigan and Drena Dobbs), National Institutes of Health, Principal Investigator, $1,022,000.
- Gene Expression Analysis in the Developing Retina (With Heather West-Greenlee and Jan Buss), National Institutes of Health, Co-Principal Investigator, $438,000.
- Effects of Microgravity on Osteoblasts and Differentiation of Mesenchymal Stem Cells, (with Don Sakaguchi, Robert Doyle, Hajime Takano), Co-Principal Investigator, National Aeronautics and Space Administration, $902,373.
- Intelligent Multi-Agent System for Intrusion Detection and Countermeasures, Department of Defense, Co-Principal Investigator, $236,880.

Current Research Projects

Distributed Knowledge Networks

This research seeks to develop, implement, and evaluate algorithmic and systems solutions for *distributed knowledge networks* multi-agent organizations consisting of stationary as well as mobile intelligent software agents designed to support the utilization of heterogeneous, distributed data and knowledge sources for automated data-driven knowledge acquisition and decision support. This includes:

- Design, implementation, and experimental evaluation of intelligent software agent architectures and algorithms for customizable, adaptive, proactive, reactive, context-sensitive information retrieval, extraction, transformation, and fusion using heterogeneous, distributed data and knowledge sources.

- Design, implementation, and experimental evaluation of learning algorithms and data transformation tools for automated (or semi-automated) incremental data-driven knowledge discovery and theory refinement using heterogeneous, distributed (including mobile) data sources.

- Design, implementation, and experimental evaluation of multi-agent organizations and coordination structures for distributed knowledge networks.

Results of this research include new algorithms and software tools which will help translate the recent advances in automated data acquisition, digital storage, computers and communications into fundamental gains in scientific understanding (e.g., through data-driven, computer-supported discovery in biological sciences), and information technology for organizational decision support, complex distributed systems monitoring and control, and related applications.

Learning from Distributed, Dynamic Data Sources

Translating recent advances in high throughput data acquisition and storage technologies and networks into fundamental gains in understanding of respective domains (e.g., in biological sciences, organizational decision support) call for the development of powerful new tools for knowledge acquisition. For example, by examining data gathered by sensors located at different network hosts (e.g., system logs that contain records of various system calls) and known cases of coordinated attacks on the network, a knowledge acquisition agent can infer useful, a-priori unknown predictive relationships that can be subsequently employed for predicting, detecting, and counteracting intrusions. Similarly, bioinformatics knowledge discovery agents can learn regularities that characterize molecular structure-function relationships. The acquired knowledge, in addition to being of immediate value to the users, would also be used by software agents to hypothesize likely events based on information available and then seek out additional data to support or refute the hypothesis (e.g., in the context of data-driven scientific discovery).

Machine learning is currently perhaps the most practical approach to automated or semi-automated data-driven knowledge acquisition and theory refinement. However, most algorithms available today require that the entire dataset be available for processing at a single location before knowledge acquisition can begin.

However, many data sources of interest are large and dynamic. Thus it is desirable to use mobile software agents that transport themselves to the data repositories, or stationary software agents that reside at the repositories, to perform as much analysis as possible where the data are located, and return only the results of analysis in order to conserve network bandwidth.

Efficient and scalable approaches to data-driven knowledge acquisition from distributed, dynamic data sources call for algorithms that can modify knowledge structures (e.g., pattern classifiers) in an incremental fashion without having to revisit previously processed data (examples). We have recently developed several formulations of the problem of learning from distributed and dynamic data sources under different assumptions regarding the data sources and the properties of the learning algorithms.

This research builds on this work to investigate several approaches to incremental learning from dynamic, distributed data sources as well as cumulative, multi-task learning in open-ended environments. These include:

- Design of knowledge representations that lend themselves to incremental update of knowledge structures using only the new data and the design of efficient update algorithms; The algorithm that we have designed for incremental induction of support vector machines provides an example of this approach.
- Design of online learning algorithms based on Stochastic Approximations. Our algorithm for incremental acquisition of spatial maps using Kalman Filters provides an example of this approach.
- Design of serial or parallel aggregation schemes for effectively combining multiple hypotheses learned using small subsets of data using techniques based on various weighting and voting schemes akin to boosting and bagging methods. We have conducted some preliminary experiments along these lines with encouraging results.
- Design of hybrid algorithms which synergistically combine statistical summaries of data, identification and memorization of informative instances, and incremental knowledge structure update.

Results of this research include new algorithmic and systems solutions for large-scale automated knowledge acquisition from large, dynamic, distributed data sources. The resulting software tools are likely to find use in a variety of applications including intrusion detection in computer systems) and and data-driven knowledge discovery in bioinformatics.

An Agent-Based Environment for Integrating and Analyzing Genomic Databases

Recent advances in high throughput (and often automated or semi-automated) data acquisition technologies, digital storage technologies, computers and communications have made it possible to gather and store large amounts of data on biological organisms. In order to translate these advances in high-throughput data acquisition into fundamental gains in scientific understanding calls for development of software tools that allow integrated retrieval and analysis of the individual databases. The design and implementation of such tools has to address several challenges in computer and information sciences:

- In many instances, the data sources are physically distributed (e.g., at multiple laboratories and data repositories). This calls for the use of *information assistants* or *software agents* for intelligent, selective, and context-sensitive data gathering, information extraction, and data assimilation prior to large scale data analysis. Since scientific data sources are dynamic (i.e., they change rapidly as data items are added or modified), there is a need to monitor the data sources, propagate the changes, and trigger the necessary updates in the affected data and knowledge repositories. For example, as experimental data about mRNA expression and localization are gathered using microarray technology and *in situ* hybridization, it would be extremely beneficial to annotate the corresponding DNA sequences in an automated or semi-automated fashion. Since the information of interest is user, problem, and context-dependent, such tools have to be customizable.

- Given the large volumes of data involved, it is desirable to perform as much analysis as feasible at the sites where the data is located and transmit only the results of analysis rather than flooding the network with data. This calls for the use of *mobile* software agents that can transport themselves to appropriate sites, carry out the computation on site, and return with useful results.

- Since the data sources are often autonomously owned and operated, and reside on heterogeneous hardware and software platforms, their effective use requires a sufficient degree of interoperability among the different data sources (despite their heterogeneity). For example, specific bioinformatics applications might have to access and use data from multiple genome or protein data banks.

- The data sources contain multiple types of data (text, images, relational databases, sequence data, spectrograms, protein structures, etc.) This calls for sophisticated tools for extracting, transforming, and assimilating relevant information from heterogeneous data sources into a *data warehouse* where it can be further analyzed to faciliate knowledge discovery.

- The large volumes of data, the range of scientifically relevant but complex interrelationships that need to be discovered, and the diversity of data sources challenge state-of-the-art approaches to *data mining* and *knowledge discovery*. In particular, current statistical and artificial intelligence tools have to be extended and new efficient algorithms developed to handle data-driven knowledge acquisition and incremental theory refinement from multiple heterogeneous, structured as well as semi-structured scientific data and knowledge sources.

This research brings together scientists with complementary expertise in computer science, molecular biology, and computational biology to design, implement, and evaluate a modular, flexible, and extensible multi-agent system for selective information retrieval, information extraction, information assimilation, and data-driven scientific knowledge discovery using heterogeneous, distributed, data and knowledge sources. A number of carefully chosen bioinformatics problems in plant genomics are being collectively used to identify the functionality required of the multiagent tools as well as to evaluate and further develop the tools.

Multi-Agent Systems for Integrated Host and Network Based Intrusion Detection

Complex Distributed Systems (e.g., computer systems, communication networks, power systems) are equipped with sensors and measurement devices that gather and store, a variety of data that is useful in monitoring and controlling the operation of such systems. For instance, system logs gathered by multiple computers connected to a network contain information t hat is useful in detecting anomalies and intrusions. Analysis of such system log s over time can lead to discovery of useful knowledge to detect intrusions on the basis of observed activity. An example of an attack involving more than one subsystem would be a combined NFS and rlogin attack wherein an attacker would determine an NFS file handle for an .rhosts file or /etc/hosts.equiv file (assuming that the appropriate file systems are exported by the UNIX system), using the NFS handle rewrite the file to gain login privileges to the attacked host. To detect and respond to such multistage or concerted attacks, the intrusion detection system must have support for gathering and operating on data and knowledge sources from the entire observed system.

This research is aimed at developing, implementing, and evaluating multi-agent systems for integrated host and network based monitoring of large distributed computer and communication networks for intrusions. A system of stationary and mobile software agents will:

- monitor different processes, resources, users, events,
- extract relevent information from system logs,
- integrate information from disparate sources over multiple space and time scales,
- detect anomalous patterns of activity,
- selectively share information with other agents,
- adapt monitoring functions to observed patterns of activity,
- perform data mining to learn predictive rules for intrusion detection, and
- recommend or execute appropriate countermeasures

Results of this research include new algorithmic and systems solutions for monitoring of large distributed systems in general, and detection of coordinated or concerted attacks on distributed computing systems in particular. This research is closely integrated with education and training of graduate and undergraduate students in Computer Science at Iowa State University.

Agent-Based Approaches to Intelligent Traffic Management in Large Communication Networks

With the unprecedented growth in size and complexity of modern distributed systems such as communication networks, the development of intelligent and adaptive approaches to system management (including such functions as routing, congestion control, traffic/load management, etc.) have assumed considerable theoretical as well as practical significance. Knowledge representation and heuristic techniques of artificial intelligence, decision-theoretic methods, as well as techniques of adaptive control offer a broad range of powerful tools for the design of intelligent, adaptive, and autonomous communication networks.

Routing in a communication network refers to the task of propagating a message from its source towards its destination. For each message received, the routing algorithm at each node must select a neighboring node to which the message is to be sent. Such a routing algorithm may be required to meet a diverse set of often conflicting performance requirements (e.g., average message delay, network utilization, etc.), thus making it an instance of a multi-criterion optimization problem.

For a network node to be able to make an optimal routing decision, as dictated by the relevant performance criteria, it requires not only up-to-date and complete knowledge of the state of the entire network but also an accurate prediction of the network dynamics during propagation of the message through the network. This, however, is impossible unless the routing algorithm is capable of adapting to network state changes in almost real time. In practice, routing decisions in large communication networks are based on imprecise and uncertain knowledge of the current network state. This imprecision is a function of the network dynamics, the memory available for storage of network state information at each node, the frequency of, and propagation delay associated with, update of such state information.

Motivated by these considerations, we have developed a multi-agent system of agents each of which maintains and updates a small knowledge base of constant size (independent of the size of the network). This knowledge base summarizes the state of the network from the agent's point of view. It provides an accurate picture of the network in the immediate neighborhood of the agent and a spatio-temporally averaged summary of the network state in distant neighborhoods. This mechanism takes advantage of the fact that the number of available paths (and hence the flexibility of routing decisions) grows as a function of distance between the source and destination.

Experimental and theoretical analysis of this approach have demonstrated several desirable properties including minimization of delay and load balancing over the entire network without access to accurate global network state information.

A long-term objective of this research is the design of completely autonomous self-managing, intelligent, low-overhead, robust and adaptive traffic management mechanisms for very large high speed communication networks of the future. Towards this objective, mechanisms that dynamically adapt network management policies in response to changes in network dynamics are of interest. This, however, requires an understanding of the complex interactions that exist between different measures of network performance and resource requirements and the development of a coherent framework that facilitates a smooth tradeoff of some of the performance measures and resource requirements against others on demand.

Coordination and Control of Multi-Agent Systems

This research seeks to develop, implement, and evaluate algorithms for coordination and control of multi-agent organizations in the context of *distributed knowledge networks* consisting of stationary as well as mobile intelligent software agents designed to support the utilization of heterogeneous, distributed data and knowledge sources for automated data-driven knowledge acquisition and decision support.

Modular and open-ended design of distributed knowledge networks implies that the resulting system consists of multiple more or less autonomous intelligent agents with different capabilities. Each agent is responsible for a particular, usually fairly narrowly defined function. Effective use of such agents in distributed problem solving (e.g., in computer-aided scientific discovery in bioinformatics) intrusion detection in distributed computing systems), require mechanisms for control and coordination the behavior of individual agents in a way that leads to the desired global behaviors. Both natural systems (e.g., cells, brains, immune systems, evolution, groups, social organizations, economies, societies) and artificial systems (computers, multi-computers, computer networks, programs, factories) offer rich sources of examples of a wide variety of coordination and control mechanisms that can be beneficially incorporated into the design of complex information processing systems in general, and multi-agent systems in particular: coordination that emerges from interaction among large number of agents that exhibit relatively simple behaviors inspired by organizations such as the ant colonies; hierarchical control where the flow of control follows the structure of the hierarchy (e.g., in the military); coordination that emerges from interaction (including communication and negotiation) among self-interested agents as exemplified in the contract net protocol and related negotiation mechanisms; control that emerges from competition for resources under the influence of environmental rewards as exemplified by evolutionary processes modeled by genetic algorithms.

We have designed, implemented, and analyzed coordination strategies for multi-agent systems in the context of distributed routing by a collection of utility-driven agents in communication networks

Work in progress is aimed at the high-level specification, design, implemention, and evaluation of different multi-agent organizations and inter-agent coordination mechanisms for interaction among intelligent mobile agents in distributed knowledge networks. Our approach to specification of coordination structures includes specification of *roles* that different agents can play based on their capabilities in *interactions* with other agents. Some multi-agent organizations that will be examined include:

- horizontal modular organization;
- hierarchical organization wherein agents at higher levels coordinate and control agents at lower levels
- a layered organization based on behavior subsumption;
- collectives of independent, self-interested (utility-driven) agents;
- collectives of communicating self-interested agents;
- collectives of self-interested negotiating agents.

Anticipated this research include new software tools for coordination and control of multi-agent systems for a variety of applications including: monitoring and control of distributed power systems and computer systems, organizational decision support systems, and distributed knowledge networks for bioinformatics.

Interactive Visual Overviews of Large, Multi-Dimensional Datasets

Recent advances in high throughput data acquisition, digital storage, computer and communications technologies have made it possible to gather, store, and transmit large volumes of data. Translating the advances in data acquisition and storage technologies into fundamental gains in understanding of the respective domains, requires the development of sophisticated computational tools to assist in the knowledge discovery process. Given the large volumes of data, and the broad range of scientifically relevant and potentially complex interrelationships that might be of interest, machine learning or data mining algorithms offer one of the most practical and cost-effective approaches to data-driven knowledge discovery. However, fully automated knowledge discovery is beyond the current state of the art in artificial intelligence, and we still need the "little-understood ability of human beings to `see the big picture' and `know where to look' when presented with visual data". This research seeks to develop sophisticated dynamic graphics tools for interactive exploratory analysis of very large datasets. These tools, when used in conjunction with data mining algorithms, will enable the user to overview the data space as well as the complex relationships discovered by the data mining algorithms. This would significantly enhance the utility of machine learning algorithms for interactive data-driven knowledge discovery from large, high dimensional datasets. This research brings together a team of researchers with complementary research interests and expertise in statistics and visualization, artificial intelligence, machine learning, and bioinformatics, databases and information management to develop a modular and extensible software toolbox for user-driven, computer-assisted, interactive exploration of extremely large, high-dimensional datasets.

This research focuses on tools for visualizing large, multi-dimensional data. The objective is to scale up the proven visual methods to work with larger amounts of data. The researchers plan to integrate the dynamic graphics tools of the grand tour with machine learning algorithms to provide an interactive, user-centered, collection of tools for knowledge discovery from very large data sets. Specifically, machine learning algorithms including dimensionality reduction techniques (e.g., principal component analysis, independent component analysis, clustering), neural networks, statistical methods (discriminant analysis, density estimation), decision tree or rule induction algorithms, feature extraction, feature subset selection, and feature extraction techniques will be an integral part of the knowledge discovery toolbox, along with the dynamic graphics tools. To scale the tour algorithm up to any size data set two approaches will be explored. The first involves the use of texture maps/images to display projected data, the maps created by a separate process that generates grand tour projections of the data which we store in a stack, and the second involves selecting useful subsets of cases (for example, using support vector machines and related machine learning algorithms) and intelligent selection of projections. The visual interface for the tour will pull projections/images out of the stack as needed to provide a smooth movie of the data, and provide interaction tools such as like a video recorder, play, forward/backward, freeze frame, zoom and pan. In addition, ways to compensate for overplotting, and ways to interact with images to facilitate drill down will be explored. The integration of these tools with data mining algorithms will enable the user to:

- detect local departures from an overall trend,
- detect structure in sparse high-dimensional data,
- select data-dependent preprocessing steps (e.g., scaling, normalization, feature extraction) to enhance the effectiveness of automated numeric or symbolic "machine learning" algorithms,
- refine solutions obtained with machine learning algorithms,
- explore the differences in inductive and representational biases of alternative machine learning algorithms (e.g., decision trees versus neural networks), and
- overview the data space as well as the complex relationships discovered by the data.

Constructive Neural Network Learning Algorithms for Pattern Classification

Induction of pattern classifiers from data is an important area of research in machine learning which finds applications in diverse areas including automated diagnosis, bioinformatics, design of customizable information assistants, intrusion detection in computer systems, among others. Artificial neural networks, because of their potential for massive parallelism and fault and noise tolerance, offer an attractive approach to the design of trainable pattern classifiers.

Constructive learning algorithms, which avoid the guesswork involved in deciding a suitable network architectures for different pattern classification problems by growing a network by recruiting neurons as needed can be effectively trained to solve complex pattern classification problems. Furthermore, it is possible to generalize (and provide convergence guarantees for) a large family of such algorithms designed for 2-class binary pattern classification problems to handle classification problems involving real-valued patterns and an arbitrary number of classes.

Different constructive neural network algorithms as well as the algorithms used to train the neuron weights have different inductive and representational biases making their performance sensitive to specific characteristics of the datasets. Thus, it is useful to experiment with a toolkit of multiple algorithms for practical applications. It is also possible to design hybrid algorithms that exploit the synergy between different algorithms. Visualization of decision boundaries constructed by the different algorithms can often provide useful insights into their behavior. This argues for systems that combine machine learning algorithms with visualization techniques in high dimensions to facilitate interactive knowledge discovery.

A simple inter-pattern distance based, provably convergent, polynomial time constructive neural network algorithm compares very favorably with computationally far more expensive algorithms in terms of generalization accuracy. The generalization accuracy of this algorithm (as expected), is sensitive to the choice of attributes used to represent the patterns and can be improved by using feature subset selection.

Fairly simple algorithms for incorporating prior knowledge can be used to enhance the performance of constructive algorithms for designing pattern classifiers. This raises the possibility of using constructive algorithms for knowledge transfer across similar tasks to facilitate multitask learning in complex environments.

A constructive learning procedure, augmented with a Kalman filter mechanism for estimation, can be effectively used for place learning and localization in a-priori unknown environments in the presence of sensor and motion uncertainties. The resulting computational model successfully accounts for a large body of behavioral and neurobiological data from animal experiments and offers several testable predictions. Preliminary results indicate that constructive learning algorithms and related approaches to machine learning can be adapted for data-driven knowledge discovery from distributed, dynamic data sources in a number of domains including bioinformatics, monitoring and control of complex distributed systems (computer systems, communication networks, power systems) and decision support systems.

Automata Induction, Grammar Inference, and Language Acquisition

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 modelling, prediction, natural language acquisition, data mining and knowledge discovery.

Regular grammars are the simplest class of formal grammars in the Chomsky hierarchy. An understanding of the issues and problems encountered in learning regular languages (or equivalently, identification of the corresponding DFA) are therefore likely to provide insights into the problem of learning more general classes of languages. Consequently, regular grammar inference from examples has received a great deal of attention in theoretical computer science, machine learning, syntactic pattern recognition, computational learning theory, and grammar inference communities.

Exact learning of the target DFA from an arbitrary presentation of labeled examples is a hard problem. Gold showed that the problem of identifying the minimum state DFA consistent with a presentation S comprising of a finite non-empty set of positive examples S+ and possibly a finite non-empty set of negative examples S- is NP-hard. Under the standard complexity theoretic assumption P not equal to NP, Pitt and Warmuth showed that no polynomial time algorithm can be guaranteed to produce a DFA that has approximately the same number of states as the target DFA from a set of labeled examples corresponding to a DFA.

In the face of these negative results, efficient learning algorithms for identification of DFA assume that additional information is provided to the learner. Trakhtenbrot and Barzdin described a polynomial time algorithm for constructing the smallest DFA consistent with a *complete labeled sample* i.e., a sample that includes all strings up to a particular length and the corresponding label that states whether the string is accepted by the target DFA or not. The *regular positive and negative inference* (RPNI) algorithm identifies in time that is polynomial in the sum of lengths of the examples, a DFA that is consistent with the given sample S. Further, as shown by Oncina and Garcia, if S is a superset of a characteristic set for the target DFA (which roughly means that each state transition is sampled at least once in some sample) then the DFA output by the RPNI algorithm is guaranteed to be equivalent to the target.

When examples are drawn at random (as in the PAC setting), results proved by Kearns and Valiant suggest that an efficient algorithm for learning DFA would entail efficient algorithms for solving problems such as breaking the RSA cryptosystem, factoring Blum integers, and detecting quadratic residues, all of which are known to be hard under standard cryptographic assumptions. Using a variant of Trakhtenbrot and Barzdin's algorithm, Lang empirically demonstrated that random DFAs are approximately learnable from a sparse uniform sample. However, exact identification of the target DFA was not possible even in the average case with a randomly drawn training sample. Some researchers have investigated the learnability of concept classes under restricted classes of distributions. Li and Vitanyi proposed a model for PAC learning with *simple* examples called the *simple PAC* model wherein the class of distributions is restricted to *simple* distributions (wherein examples are drawn according to the Universal distribution where strings with low Kolmogorov complexity are sampled with a high probability).

Our work has shown (among other things) that the class of DFA that have low Kolmogorov complexity are efficiently learnable if samples are drawn according to the Universal distribution. Some directions that are currently being explored include: efficient algorithms for learning interesting subclasses of grammars; application of grammar inference to characterizing macromolecular sequence families (DNA, RNA, and protein sequences), and incremental grammar induction.

Gene Expression Analysis

The central dogma of modern biology states that the functional state of an organism is determined largely by the pattern of expression of genes. Thus, the function of a cell, how well a cell performs its function, and even the determination of a cell's type is controlled by the level at which genes are expressed in the cell. Different gene expression levels can, for example, differentiate a neuron from a muscle cell or a normal cell from a cancerous cell of the same type. However, many biological processes of interest (e.g., cellular differentiation during development, aging, disease) are controlled by complex interactions over time between hundreds of genes. Furthermore, each gene is involved in multiple functions. Given the fact that thousands of genes are involved in determining the functional state of an organism, the task of assigning functions to genes, identifying underlying genetic signaling pathways, genetic control and regulatory networks is a formidable task.

The advent of microarray technology provides biologists with the ability to measure the expression levels of thousands of genes in a single experiment. Microarrays come in several flavors. In a common type of microarray, typically called a DNA microarray, several thousand DNA samples are fixed to a glass slide, each at a known position in the array. Each probe (a DNA sample) corresponds to a single gene within the organism under investigation. Messenger RNA samples are then collected from a population of cells under a given set of experimental conditions. These samples are converted into cDNA via reverse transcription and labeled with a fluorescent dye. The cDNA sample is hybridized with the microarray. The level of expression of a particular gene is roughly proportional to the amount of cDNA that hybridizes with the DNA affixed to the slide. By repeating this process under different experimental conditions, using different fluorescent dyes to distinguish between any pair of experimental conditions, it is possible to measure the expression levels of different thousands of genes under the chosen conditions. Thus, for example, the relative levels of expression of a particular gene under any pair of experimental conditions can be obtained by measuring the ratio of the amount of the two dyes present at the position of each DNA sequence on the slide using laser scanning technology.

With the increasing use of DNA microarray and related technologies for gathering gene expression data from plants and animals, there is a growing need for sophisticated computational tools for extracting biologically significant information from gene expression data, assigning functions to genes, and identifying shared signalling pathways and control circuits (e.g., genetic signalling and genetic regulatory networks). Against this background, our research is aimed at precise characterization of the information requirements of these tasks and the design and implementation of suitable algorithms for gene expression analysis. Algorthms are being developed for:

- Clustering of gene expression patterns to identify genes that are coexpressed and possibly coregulated.
- Prediction of gene function based on expression patterns.
- Discovery of statistically significant associations between expression patterns of a chosen gene and a subset of other genes.
- Inference of unknown genetic signalling networks from gene expression time series data.

The resulting computational tools for gene expression analysis are being applied to a variety of problems in computational biology including:

- Understanding the relationship between gene expression in the nervous system (e.g., in different neuron and glial cells) under various developmental, pathological, and functional states (in collaboration with Phil Haydon's laboratory).
- Identification of co-regulated genes, assignment of genes to functional families based on expression patterns, and working out underlying genetic signalling pathways in plants (in collaboration with Steve Rodermel’s laboratory).

Macromolecular Structure-Function Prediction from Sequences

Characterizing the structure and function of biological macromolecules (e.g., proteins) is a central problem in modern biology. Proteins are abundant in all organisms and are indeed fundamental to life. The diversity of protein structure underlies the very large range of their function. Proteins are linear heteropolymers of fixed length. That is, a single type of protein always has the same number and composition of monomers, but different proteins have a range of monomer units, from a few tens to approximately a thousand. The monomers are amino acids, and there are 20 types, which themselves have a range of chemical properties. There is therefore a great diversity of possible protein sequences. The linear chains fold into specific three-dimensional conformations, which are determined by the sequence of amino acids; proteins are generally self-folding. The three-dimensional structures of proteins are therefore also extremely diverse, ranging from completely fibrous, to globular. Protein structures can be determined to an atomic level by X-ray diffraction and neutron-diffraction studies of crystallized proteins, and more recently, by nuclear magnetic resonance (NMR) spectroscopy of proteins in solution. Recent advances in genome sequencing projects have led to enormous increase in protein sequence data. However, the database for protein structures has lagged far behind the number of known sequences. Protein sequences are encoded in DNA, the holder of genetic information, which is itself a linear molecule composed of four types of *bases* (monomers which act as *letters* of the genetic alphabet. In principle, it should therefore be possible to translate a gene sequence into an amino acid sequence, and to predict the three-dimensional structure of the resulting chain from this amino acid sequence by performing a detailed simulation of molecular dynamics of protein folding based on the underlying physics and chemistry of interactions among the monomers. However, searching the entire space of possible conformations for a stable or low energy 3-dimensional structure is a computational impossibility. Consequently, successful approaches to protein structure/function prediction from sequence cannot rely on an explicit search of the conformation space. Instead, the problem, like its counterparts in the realm of complex systems, calls for careful discovery and exploitation of regularities at multiple spatial and temporal scales. Such regularities can be observed at the level of secondary structures (e.g., alpha helix, beta fold with an elementary unit e.g., a spiral, or a hairpin), the so-called foldons (which are believed to be independently folding units), highly conserved signature motifs in the sequence that are characteristic of protein families.

Against this background, our research is aimed at the design and implementation of algorithms for molecular structure/function prediction from sequence data. We have recently succeeded in using data mining approaches to design sequence classifiers that assign protein sequences to corresponding functional families. Our approach maps each protein sequence into an N-dimensional vector whose components encode the presence or absence (or other more sophisticated information measures related to the presence of absence) of sequence *motifs*. Proteins with known functions are used to generate a training set which is then used to construct a sequence classifier. New sequences are assigned to fuctional families based on complex inter-relationships between sequence motifs discovered by a machine learning algorithm. This work builds on recent work on a broad range of data mining and knowledge discovery algorithms (including those drawn from artificial intelligence, statistical pattern recognition, grammatical inference, neural networks, text classification, and evolutionary computing).

Evolutionary Synthesis of Complex Adaptive Systems

Evolutionary algorithms (genetic algorithms, evolution strategies, genetic programming, evolutionary programming) offer an attractive paradigm for automated synthesis of complex adaptive systems e.g., sensory, behavior, and control structures for robots and intelligent agents. Our current research in this area is focused on automated synthesis of neural networks, finite automata, and computational architectures for reactive and deliberative behavior in intelligent autonomous agents and robots under a variety of cost, performance, and environmental constraints. Related research explores evolution of communication, cooperation, and language in communities of intelligent agents.

Massively Parallel Architectures for Knowledge Representation and Inference

Artificial neural networks, because of their massive parallelism and potential for fault tolerance provide an attractive paradigm for the implementation of high performance computer systems for knowledge representation and inference for real-time applications. Our research on neural architectures for knowledge representation and reasoning focuses on fault-tolerant neural network architectures for associative memories, data storage and retrieval, syntax analysis, and deductive inference.

Computational Models of Learning, Memory, and Behavior

Our laboratory is involved in interdisciplinary research on learning and memory, with particular emphasis on design and theoretical and experimental analysis of and computational models of cognitive phenomena. This often leads to design and implementation of biologically inspired computational architectures for intelligence for applications in robotics and intelligent agents. Of particular interest are spatial learning and navigation in animals and computational investigation of the role of neuron-glia, glia-neuron, and glia-glia interactions in learning and memory.

Other Topics of Current Interest

Other topics of interest include: Computational Models of Discovery; Computational Models of Creativity; Computational Modeling and Simulation of Complex Systems (evolution, social systems, immune systems, communication networks, etc.); Hybrid Intelligent Systems; Philosophy of Mind; Intelligent Design and Manufacturing Systems; Applications of information theory and complexity theory (in particular, Kolmogorov complexity, minimum description length, and related topics) in computational learning theory and biology.

PROFESSIONAL ACTIVITIES

Editor

- Co-Editor-in-chief, *Journal of Cognitive Systems Research.* (Published by Elsevier) (1999-).
- Member of Editorial Board, *Machine Learning Journal*, (2002-2005)
- Guest Editor, *Machine Learning Journal*, Special Issue on Automata Induction, Grammar Inference, and Language Acquisition. (2001).
- Editor, *Encyclopedia of Cognitive Science.* (Published by Macmillan).

Conference Program Committee Member, Program Chair, etc.

- **Program Chair**, Conference on Computational Biology and Genome Informatics, Durham, North Carolina, USA. 2002.
- Program Committee Member, International Colloquium on Grammatical Inference (ICGI-2002), Amsterdam, Netherlands. 2002.

- Program Committee Member, Second International Conference on Hybrid Intelligent Systems (HIS 2002), Santiago, Chile. 2002.
- Program Committee Member, Second International Workshop on Intelligent Systems Design and Applications, Atlanta, Georgia, USA. 2002.
- Program Committee Member, Thirteenth Midwest Artificial Intelligence and Cognitive Science Conference (MAICS-2002), Chicago, Illinois, USA.
- Track Chair, Adaptive Behavior and Agents, Genetic and Evolutionary Computation Conference (GECCO-2002), New York, USA. 2002.
- Scientific Committee Member, Network Applications in Bioinformatics (NETTAB) Workshop on Agents in Bioinformatics, Bologna, Italy, 2002.
- Program Committee Member, Fourth International Workshop in Frontiers of Evolutionary Computation (FEA-2002), Durham, North Carolina, USA. 2002.
- Program Committee Member, International Conference on Machine Learning (ICML-2001), 2001.
- **Organizer and Chair**, IJCAI-2001 Workshop on Knowledge Discovery from Heterogeneous, Distributed, Dynamic, Autonomous Data and Knowledge Sources. International Joint Conference on Artificial Intelligence, 2001. Seattle, Washington, USA.
- **Organizer and Co-Chair**, Workshop on Cognitive Agents and Multi-Agent Interaction, International Conference on Cognitive Science, August 2001. Beijing, China.
- International Scientific Committee Member, International Symposium on Artificial Intelligence (ISAI 2001). Fort Panhala, Kolhapur, India.
- Program Committee Member, International Joint Conference on Neural Networks (IJCNN-2001), 2001.
- Program Committee Member, International Workshop on Hybrid Intelligent Systems (HIS-2001), Australian National Conference on Artificial Intelligence, 2001.
- Program Committee Member, Third Workshop on Mining Scientific Datasets. First Siam International Conference on Data Mining. Chicago, USA. 2001.
- Program Committee Member, Midwestern Artificial Intelligence and Cognitive Science Conference (MAICS 2001).
- Program Committee Member, Genetic and Evolutionary Computing Conference (GECCO 2001).
- Program Committee Member, GECCO-2001 Workshop on Gene Expression. Genetic and Evolutionary Computing Conference, 2001.
- Program Committee Member, Fifth International Colloquium on Grammatical Inference. Lisbon, Spain. 2000.

- **Organizer and Co-Chair**, Workshop on Learning from Sequential and Temporal Data, Seventeenth International Conference on Machine Learning, Stanford University, Palo Alto, CA. 2000.
- Program Committee Member, Genetic and Evolutionary Computing Conference (GECCO) 2000.
- Program Committee Member, Midwest Artificial Intelligence and Cognitive Science Conference (MAICS 2000).
- Program Committee Member, Sixteenth National Conference on Artificial Intelligence (AAAI 99). Orlando, Florida, July 1999.
- **Co-Chair**, AAAI Workshop on Computation with Neural Systems, Orlando, Florida. July 1999.
- **Track Chair**, Artificial Life, Agents, and Adaptive Behavior Track. Genetic and Evolutionary Computing Conference (GECCO 99), Orlando, Florida. July 1999.
- Member, Advisory Committee, International Conference on Evolutionary Computation, 1999.
- Program Committee Member, Midwest Artificial Intelligence and Cognitive Science Conference, 1999.
- **Chair**, Fourth International Colloquium on Grammatical Inference (ICGI 98), Ames, Iowa, July 1998.
- Program Committee Member, Genetic Programming (GP-98), Madison, Wisconsin, 1998.
- Program Committee Member, Midwest Artificial Intelligence and Cognitive Science Conference, 1998.
- Program Committee Member, International Conference on Machine Learning (ICML 97), Nashville, TN. 1997.
- Program Committee Member, Genetic Programming (GP-97), Stanford University, Stanford, CA. 1997.
- **Organizer and Co-chair**, Workshop on Automata Induction, Grammatical Inference, and Language Acquisition, International Conference on Machine Learning (ICML 97), Nashville, TN.
- Advisory Committee Member, International Conference on Artificial Intelligence Applications (ICAIA '97), Cairo, Egypt. July 1997.
- Program Committee Member, Midwest Artificial Intelligence and Cognitive Science Conference, Dayton, OH. 1997.
- Program Committee Member, *Genetic Programming (GP-96)*, Stanford University, Stanford, CA. July 1996.
- Program Commitee Member, and Session Co-Chair, World Congress on Neural Networks, San Diego, CA. 1996.

- Program Committee Member, Midwest Artificial Intelligence and Cognitive Science Society Conference, Bloomington, Indiana. April 1996.
- Program Committee Member, *Midwest Artificial Intelligence and Cognitive Science Society Conference*, Carbondale, Illinois. April 1995.
- Program Committee Member and Invited Session Co-Chair, World Congress on Neural Networks, Washington, D. C. July 1995.
- Program Committee Member, *International Simulation Technology Conference.* San Francisco, California. 1993.
- Program Committee Member, *Fifth UNB Artificial Intelligence Symposium*, Fredericton, Canada. 1993.
- Program Committee Member and Session Chair, *Learning and Optimization. SimTec/WNN*, Houston, Texas. November 1992.
- **Organizer and Chair**, *Workshop on Symbolic and Subsymbolic Information Processing in Biological Neural Circuits and Systems. Conference on Neural Information Processing Systems.* Vail, Colorado. December 1992.
- Program Committee Member and Session Chair, *Fourth UNB Artificial Intelligence Symposium*, Fredericton, Canada. September 1991.

Professional and Scientific Societies

- Member, Institution of Electrical and Electronic Engineers (IEEE)
- Member, Association for Computing Machinery (ACM)
- Member, American Association for Artificial Intelligence (AAAI)
- Elected Member, New York Academy of Sciences (NYAS)
- Member, American Association for the Advancement of Science (AAAS)
- Member, International Neural Networks Society (INNS)
- Associate, Behavioural and Brain Sciences (BBS)
- Elected Member, Sigma Xi (Scientific Research Honor Society)
- Member, Society for Neuroscience
- Member, Cognitive Science Society
- Member, Cognitive Science Society Committee on Cognitive Science Education
- Member, ACM Special Interest Group In Artificial Intelligence (SIGART)
- Member, Neural Networks Council (NNC)

- Organizer, Iowa State University (local) chapter of ACM Special Interest Group on Artificial Intelligence (SIGART).

Research Proposal Review

- Member of the Bioinformatics Special Study Section, *National Institutes of Health*, 2001-present
- Member of Review Panel, Information and Intelligent Systems Program, *National Science Foundation*, 2002
- Member of Review Panel, CISE SBIR Program, *National Science Foundation*, 2000.
- Member of Review Panel, *U.S. Civilian Research Development Foundation*, 2000.
- Member of Review Panel, Integrative Graduate Education and Research Training (IGERT) Program, *National Science Foundation*, 1999.
- Member of Review Panel, CISE Research Instrumentation Grants, *National Science Foundation*, 1994.
- Reviewer, *National Science Foundation*
- Reviewer, *Defense Advanced Research Projects Agency*
- Reviewer, *Department of Energy*
- Reviewer, *National Scientific and Engineering Research Council*, Canada.

Journal Referee

- Applied Intelligence
- Connection Science
- IEEE Expert
- IEEE Computer
- Information Sciences
- Information and Computation
- IEEE Transactions on Neural Networks
- IEEE Transactions on Knowledge and Data Engineering
- IEEE Transactions on Evolutionary Computation
- IEEE Transactions on Pattern Analysis and Machine Intelligence Integrated Computer-aided Engineering
- Journal of Machine Learning Research

- Neural Computation
- Neural Networks
- Machine Learning

Conference Referee

- National Conference on Artificial Intelligence (AAAI)
- Cognitive Science
- International Joint Conference on Neural Networks
- International Joint Conference on Artificial Intelligence
- International Conference on Machine Learning
- World Congress on Neural Networks
- Genetic Programming Conference
- IEEE Symposium on Frontiers of Massively Parallel Computation
- Midwest Artificial Intelligence and Cognitive Science Society Conference
- IEEE Conference on Computer Vision and Image Processing
- IEEE Conference on Communications
- UNB Artificial Intelligence Symposium
- International Conference on Intelligent Information Management Systems

Advisory Board Membership

- Member of Advisory Board, Multi-University Research Project on Emergent Computational Neural Network Architectures (Emernet) sponsored by the Engineering and Physical Sciences Research Council, Universities of Edinburgh, York, and Sunderland, United Kingdom. 1999-.
- Member of Search Committee, Senior Science and Technology Position in Defensive Information Warfare, Air Force Research Laboratory, Rome, New York. 1999.
- Reviewer, University of Georgia Board of Regents, Proposal for a Ph.D. Program in Artificial Intelligence.
- Member, Intelligent Knowledge-Based Systems Working Group, Information Institute, Information Directorate, Air Force Research Labs, 1998;
- Member, Intelligent Agents Working Group, Information Institute, Information Directorate, Air Force Research Labs, 1999.

- 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

Departmental Service

- Director of Research, Computer Science Department 2002-2003.
- Member, Computer Science Graduate Admissions Committee, 2002-2003.
- Member, Computer Science Graduate Committee, 2002-2003.
- Member, Computer Science Department Chair Search Committee, 2000-2002.
- Co-Chair, Computer Science Strategic Planning Committee (ad hoc), 2001.
- Member, Computer Science Faculty Search Committee, 2000-2001.
- Chair, Computer Science Graduate Admissions Committee, 2000-2002.
- Coordinator, Computer Science Research Methods and Writing Workshop for Graduate Students, 2000-2001.
- Member, Joint Computer Science and Computer Engineering Planning Committee for Computer and Information Sciences Research Institute at Iowa State University (ad hoc), 2000.
- Chair, Computer Science Faculty Search Committee, 1999-2000.
- Chair, Computer Science Graduate Admissions Committee, 1999-2000.
- Computer Science Department Representative, ISU Faculty Conference, Grinnell, 1998.
- Member, Computer Science Graduate Committee, Iowa State University, 1996-99.
- Member, Advisory Committee to the Chair on ISU Presidential Task Force on Computer and Information Science and Technology (ad hoc), Iowa State University, 1998.
- Member, Computer Science Research Infrastructure Proposal Committee (ad hoc), Iowa State University, 1996.
- Member, Partnership in Advanced Computing Infrastructure Committee (ad hoc), Iowa State University, 1996.
- Member, Computer Science Industry Day Organizational Committee (ad hoc), Iowa State University, 1996.
- Coordinator, Computer Science Departmental Seminar for Beginning Graduate Students (ComS 591), Iowa State University, 1996-97.

- Member, Computer Science Graduate Admissions Committee, Iowa State University, 1990-1996.
- Departmental Representative (ad hoc), Interdepartmental Cognitive Science Initiative, 1996-97.
- Faculty Secretary, Department of Computer Science, Iowa State University, 1995-1996.
- Member, Departmental Vision Committee (ad hoc), Department of Computer Science, Iowa State University, 1996.
- Member, Strategic Planning Committee (ad hoc), Department of Computer Science, Iowa State University, 1994-1995.
- Member, Center for Computing and Communications Research proposal committee (ad hoc), Iowa State University, 1991.

College and University Service

- Associate Chair, Bioinformatics and Computational Biology Graduate Program, 2001-2003
- Member, Provost's Committee on Advanced Computational Research and Infrastructure, 2002-2003
- Member, Supervisory Committee, Bioinformatics and Computational Biology Graduate Program, 1999-.
- Chair, Graduate Admissions Committee, Bioinformatics and Computational Biology Graduate Program, 1999-.
- Member, Supervisory Committee, Complex Adaptive Systems Graduate Minor, 1999-.
- Coordinator, ISU Complex Adaptive Systems Workshop, 1999.
- Chair, Complex Adaptive Systems Program Steering Committee, 1997-1999.
- Member, Computational Biology Steering Committee, 1997-
- Member, Carver and University Research Proposal Review Committee, 1996, 1997.
- Member, Supervisory Committee, Neuroscience Program, Iowa State University, 1995-2000.
- Advisor, Iowa State University Team, AAAI Robot Competition, 1996.
- Member, Liberal Arts and Sciences Honors Program Committee, Iowa State University, 1992-1994.

HONORS AND AWARDS

- National Merit Scholar, India (1975-82).
- National Science Talent Scholar, India (1977-82).
- Gold Medal for Academic Excellence, Bangalore University, India (1982).
- Connectionist Models Summer School Fellowship (1988).
- Argonne Parallel Computing Workshop Fellowship (1988).
- IJCAI Student Fellowship (1989).
- McDonnell-Pew Cognitive Neuroscience Workshop Fellowship (1989).
- Workshop on Human and Machine Cognition Fellowship (1990).
- NATO Advanced Workshop Travel Fellowship (1992).

PUBLICATIONS

Books

1. Honavar, V. & Slutzki, G. (Ed.) (1998). *Grammatical Inference* Vol. 1433. Lecture Notes in Computer Science. Berlin: Springer-Verlag.
2. Patel, M., Honavar, V. & Balakrishnan, K. (Ed.) (2001). *Advances in Evolutionary Synthesis of Intelligent Agents*. Cambridge, MA: MIT Press.
3. Honavar, V. & Parekh, R., & de la Higuera, C. (Ed.) (2003). Advances in Automata Induction, Grammar Inference, and Language Acquisition. Kluwer. To appear.
4. Honavar, V. Learning from Heterogeneous, Distributed, Autonomous Data Sources. (2004). (with contributions from D. Caragea, J. Reinoso-Castillo, A. Silvescu, H. Leiva, and J. Zhang) To appear.
5. Banzaf, W., Daida, J., Eiben, A. Garzon, M., Honavar, V., Jakiela, M., & Smith, R. (Ed.) (1999). Genetic and Evolutionary Computation. Proceedings of GECCO 99. San Mateo, CA: Morgan Kaufmann.
6. Honavar, V. & Uhr, L. (1994) (Ed). *Artificial Intelligence and Neural Networks: Steps Toward Principled Integration*. New York, NY: Academic Press.

Refereed Journal Papers

1. Andorf, C., Dobbs, D., and Honavar, V. Reduced Alphabet Representations of Amino Acid Sequences for Protein Function Classification. *Information Sciences.* In press.
2. Wang, X., Schroeder, D., Dobbs, D., and Honavar, V. (2002). Data-Driven Discovery of Rules for Protein Function Classification Based on Sequence Motifs. *Information Sciences*. In press.

3. 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*. In press.

4. Helmer, G., Wong, J., Honavar, V., and Miller, L. (2002). Lightweight Agents for Intrusion Detection. *Journal of Systems and Software*. In press.

5. 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.

6. 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.

7. 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.

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

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

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

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

18. 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.

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

20. Yang, J. & Honavar, V. (1998). Feature Subset Selection Using a Genetic Algorithm. **Invited** paper. *IEEE Intelligent Systems* vol. 13, no. 2., pp. 44-49.

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

22. 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.

23. 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.

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

25. 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.

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

27. 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.

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

Journal Submissions Under Review

1. Caragea, D., Silvescu, A., and Honavar, V. Learning Classification Trees from Distributed Data.

2. Cook, D., Caragea, D., and Honavar, V. Visualization for Classification Problems, With Examples Using Support Vector Machines.

3. Helmer, G., Wong, J., Slagell, M., Honavar, V., Miller, L., and Lutz, R. Software Fault Tree and Colored Petri Net Based Specification, Design and Implementation of Agent-Based Intrusion Detection Systems.

4. Yan, C., Dobbs, D., and Honavar, V. Prediction of Protein-Protein Interaction Sites Based on Sequence Characteristics.

5. Reinoso-Castillo, J., Silvescu, A., Gadia, G., and Honavar, V. A Federated Query-Centric Approach to Ontolgy-Driven Information Extraction and Integration from Autonomous, Heterogeneous, Distributed Data Sources.

6. Zhang, X., Lonosky, P., Dobbs, D., Rodermel, S., and Honavar, V. Clustering Algorithms for Gene Expression Analysis: A Comparative Study.

7. Lonosky, P., Zhang, X., Honavar, V., Dobbs, D., Fu, A., and Rodermel, S. A Proteomic Analysis of Chloroplast Biogenesis in Maize.

Invited or Refereed Book Chapters

1. Honavar, V., Andorf, C., Caragea, D., Dobbs, D., Reinoso-Castillo, J., Silvescu, A. Wang, X. (2002). **Invited** Chapter. Algorithmic and Systems Solutions for Computer Assisted Knowledge Acquisition in Bioinformatics and Computational Biology. In: *Computational Biology and Genome Informatics*. Wu, C., Wang,P., and Wang, J. (Ed.) World Scientific. In press.

2. 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.

3. 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. In press.

4. Honavar, V., Andorf, C., Caragea, D., Silvescu, A., and Sharma, T. (2001). **Invited** Chapter. Agent-Based Systems for Data-Driven Knowledge Discovery from Distributed Data Sources: From Specification to Implementation. In: *Intelligent Agent Software Engineering.* Plekhanova, V. and Wermter, S. (ed.). Idea Group Publisher. To appear.

5. 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.

6. 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.

7. 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.

8. 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.

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

Refereed Conference Papers

1. 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. Berlin: Springer-Verlag. In press.

2. Jhang, 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.

3. 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.

4. 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.

5. 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.*

6. 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.

7. 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*.

8. 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*.

9. Silvescu, A. and Honavar, V. (2001). Temporal Boolean Network Models of Genetic Networks and Their Inference from Gene Expression Time Series. In: Proceedings of the Atlantic Symposium on Computational Biology, Genome Information Systems & Technology.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

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

18. 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.

19. 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.

20. 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.

21. 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.

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

23. 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.

24. 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.

25. 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.

26. 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.

27. 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.

28. 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.

29. 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.

30. 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.

31. 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.

32. 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.

33. 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.

34. 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.

35. 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.

36. 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.

37. 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.

38. 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.

39. 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.

40. 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.

41. 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.

42. 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.

43. 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.

44. 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.

45. 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.

46. 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.

47. 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.

48. 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.

49. 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.

50. 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.

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

52. 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.

53. 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.

54. 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.

55. 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.

56. 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.

57. 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.

58. 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.

59. 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.

60. 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.

61. 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.

62. 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.

63. 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.

64. 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.

65. 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.

66. 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.

67. 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.

68. 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.

69. 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.

Refereed Extended Abstracts in Conferences

1. 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.

2. 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.

3. 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.

4. 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.

5. 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.

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

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.

Juried Conference Publications

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

2. 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.

3. 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.

4. 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).

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.

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.

Workshop Presentations

1. 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.

2. 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.

3. 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.

4. 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.

5. 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.

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

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

8. 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.

9. Honavar, V. (1990). Generative Learning Algorithms for Connectionist Networks. In: *NIPS-90 Post-Conference Workshop on Constructive and Destructive Learning Algorithms*. Keystone, CO.

INVITED LECTURES AND TUTORIALS

Invited Talks

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

2. **ACM Lecture,** Algorithmic and Systems Solutions for Computer Assisted Knowledge Discovery from Biological Data. Missouri, 2002 (Planned).

3. **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. (Planned)

4. **Invited Colloquium,** Learning Under Simple Distributions (or Learning from a Helpful Teacher). University of Wisconsin-Madison Computer Science Department. 2002. (Planned).

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

6. **Keynote Address**, Learning from Large, Distributed, Heterogeneous Data Sets. International Symposium on Artificial Intelligence (ISAI 2001), Kohlapur, India.

7. 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.

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

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

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

11. **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.

12. **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.

13. Distributed Knowledge Networks for Bioinformatics. *Iowa Computational Biology Laboratory*, Ames, Iowa. 1998.

14. Intelligent Agents *Department of Mechanical Engineering*, Iowa State University, Ames, Iowa, 1998.

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

16. Data Mining and Knowledge Discovery. *Irish Life*, Des Moines, Iowa. September 1996.

17. Computational Models of Learning. *Department of Electrical and Computer Engineering, Iowa State University*, 1996.

18. Experiments in Evolutionary Robotics. *Neurosciences Seminar, Iowa State University*, 1996.

19. Knowledge Acquisition Through Machine Learning. *Principal Mutual*, Des Moines, Iowa. January 1994.

20. **Invited** talk, Panel on Learning in Knowledge-Based Systems. *Second World Congress on Expert Systems.* Lisbon, Portugal (1994).

21. **Invited** talk, Panel on Hybrid Architectures for Intelligent Systems. *Second World Congress on Expert Systems.* Lisbon, Portugal (1994).

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

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

24. Generative Learning Structures and Processes for Generalized Connectionist Networks. *Cognitive and Learning Systems Laboratory, Siemens Research*, Princeton, NJ. (1990).

Invited Tutorials

1. *Computational Learning Theory*, Genetic Programming Conference, Stanford, 1997.

2. *Intelligent Agents*, Genetic Programming Conference, Madison, WI, 1998.

3. *Intelligent Agents and Multi-Agent Systems* IEEE Conference on Evolutionary Computation (CEC), Washington, DC. 1999.

REFERENCES AVAILABLE ON REQUEST