AN OVERVIEW OF RESEARCH ACTIVITIES IN THE MACHINE LEARNING AND INFERENCE LABORATORY: 1996-1997

Edited by

R. S. Michalski
Q. Zhang

Reports of the Machine Learning and Inference Laboratory, MLI 97-11, George Mason University, Fairfax, VA, 1997.

An Overview of Research Activities in the Machine Learning and Inference Laboratory: 1996-1997

(based on the MLI Laboratory Web pages)

P97-14
MLI 97-11

Ryszard S. Michalski and Qi Zhang (Eds.)

May 30, 1997

Reports

Machine Learning and Inference Laboratory

An Overview of Research Activities in the Machine Learning and Inference Laboratory: 1996-1997

(based on the MLI Laboratory Web pages)

P97-14
MLI 97-11

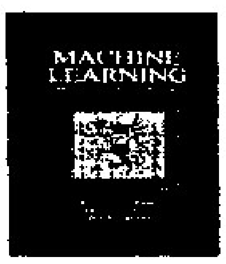
School of Information Technology and Engineering

George Mason University

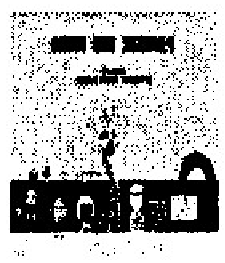
Machine Learning and Inference Laboratory

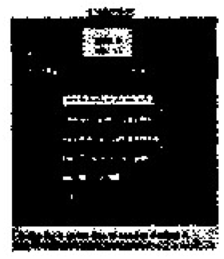
Mission
Research
People
Software
Publications
Colloquia
Events
Guest Book

Contributions



Papers
and
Reports





This site has been visited 6396 times. Last updated: 05/07/1997

URL: http://www.mli.gmu.edu

Contents

Mission

Machine Learning and Inference (MLI) Laboratory conducts fundamental and experimental research on the development of intelligent systems capable of advanced forms of learning and inference, and applies them to real-world problems. Major research areas include the development of theories and models of learning and inference, task-adaptive intelligent agents, software systems with learning capabilities, knowledge acquisition and discovery systems, machine vision through learning, and, generally, integrating learning with perception. The developed systems are experimentally applied in cooperation with industry to a wide spectrum of practical problems. Application areas of special interest include engineering design, information systems, communication networks, intelligent net surfing, geographic information systems, world economy, education, software engineering and computer vision. The Laboratory supports education, scholarship and research in these areas. It has a highly international team of researchers and the state-of-the-art computer facilities.

Previous Next

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Major Research Projects

- Theories of Learning, and Inference and Discovery
- Learning Systems
- Data Mining and Knowledge Discovery in Databases: INLEN
- Machine Vision through Learning
- Education

Theories of Learning, and Inference and Discovery

- Inferential Theory of Learning (Michalski, Wnek, Sklar, Alkharouf, Bloedorn, Kaufman, Utz)
- Multistrategy Task-Adaptive Learning: MTL (Michalski, Wnek, Kaufman, Utz, Vafaie, Zhang)
- Knowledge Representation Using Dynamically Interlaced Hierarchies (Michalski, Alkharouf, Utz)
- Cognitive Models of Plausible Reasoning (Michalski and Sklar)
- Learning Goals in Multistrategy Learning (Michalski and Utz)
- Inferential Theory of Design (Arciszewski, Michalski, Wnek)

Learning Systems

- AQ18-MOR and Natural Induction (Michalski, Zhang)
- Data-driven Constructive Induction: AQ17-DCI (Michalski, Bloedorn)
- Hypothesis-driven Constructive Induction: AQ17-HCI (Michalski, Wnek)
- Multistrategy Constructive Induction: AQ17-MCI (Michalski, Bloedorn, Wnek)

- Constructive Induction in Engineering Design (Arciszewski, Michalski, Wnek, Bloedorn)
- Constructive Induction Approach to Growing Neural Networks (Sazonow, Wnek)

Data Mining and Knowledge Discovery in Databases: INLEN

- Knowledge Discovery in Databases:INLEN (Michalski, Kaufman, Mitchell, Bloedorn, Kerschberg, Wnek, Imam, Ribeiro, Wozniak)
- Learning Problem-Optimized Decision Trees from Decision Rules (Michalski, Imam)
- Expert Systems with Learning Capabilities (Michalski, Kaufman, Imam, Ribeiro)

Machine Vision through Learning

- Multi-Level Image Sampling and Interpretation: MIST Methodology (Michalski, Duric, Zhang, Maloof)
- Machine Vision and Learning (Michalski, Duric, Maloof, Zhang, Wnek, Bloedorn) (with Computer Vision Laboratory of the University of Maryland at College Park, Rosenfeld, Aloimonos, Davis)
- Multistrategy Learning Vision Tasks by Integrating Symbolic and Neural Net Learning for Vision Tasks (Michalski, Zhang)
- Learning to Recognize Shapes (Michalski, Duric, Maloof)
- Dynamic Recognition (Michalski, Bloedorn)

Education

- Integrated Learning Systems for Education and Research: Emerald (Michalski, Kaufman, Lee, Wnek, Bloedorn, De Jong, Schultz)
- Diagrammatic Visualization of Learning Processes (DIAV) (Michalski, Wnek)

Previous Next

Inferential Theory of Learning

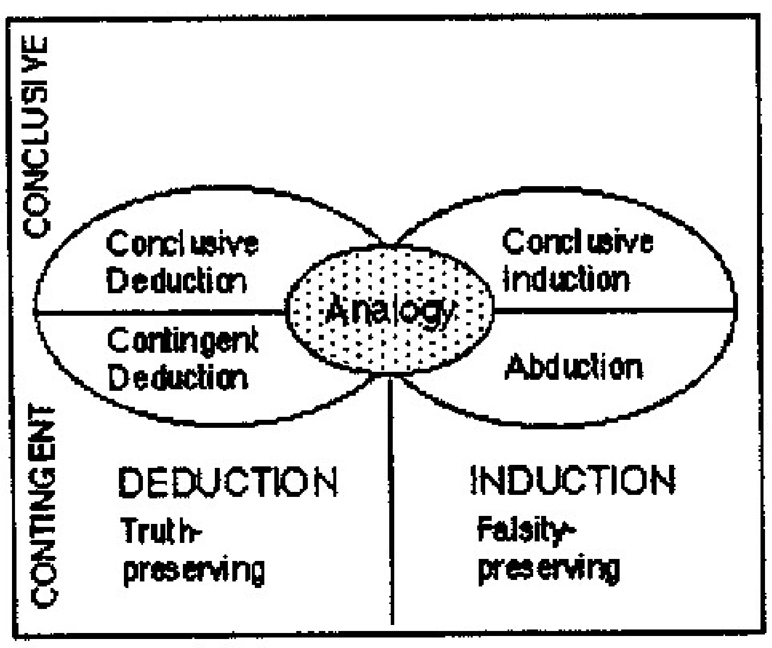
(Michalski, Wnek, Alkharouf, Bloedorn, Kaufman, Sklar, Utz)

This project aims at the development of the Inferential Theory of Learning (ITL) that views learning as a goal-oriented process of improving the learner's knowledge by exploring the learner's experience. The theory aims at understanding the competence aspects of learning processes, in contrast to the Computational Learning Theory that concerns their computational complexity. ITL addresses such questions as what types of inference and knowledge transformations underlie learning processes and strategies; what types of knowledge the learner is able to learn from a given input and from a given prior knowledge; what logical relationships exist among the learned knowledge, possible inputs and prior knowledge, etc.

The theory analyzes learning processes in terms of high level inference patterns called knowledge transmutations. Among basic transmutations are generalization, abstraction, similization, generation, insertion and replication. The central aspect of any transmutation is the type of underlying inference. If results of inference are found useful, then they are memorized. Thus, we have an "equation":

Learning = Inferencing + Memorizing

Since learning processes may involve any possible type of inference, the ITL postulates that a complete learning theory has to encompass a theory of inference. To this end, we have attempted to identify and classify all major types of inference.



A Classification of Major Types of Inference

The figure above illustrates the proposed classification. The first criterion divides inferences into deductive and inductive. To explain them in a general way, consider the fundamental equation for inference: P È BK |= C, where P stands for premise, BK for reasoner's background knowledge, |= for entailment, and C for consequent. Deductive inference is deriving C, given P and BK, and is truth-preserving. Inductive inference is hypothesizingP, given C and BK, and is falsity-preserving.

The second classification divides inferences into conclusive (strong) and contingent (weak). Conclusive inferences involve domain-independent inference rules, while contingent inferences involve domaindependent rules. Contingent deduction produces likely consequences of given causes, and contingent induction produces likely causes of given consequences. Analogy can be characterized as induction and deduction combined, and therefore occupies the central area in the diagram. Using this approach, we have clarified several basic knowledge transmutations, such as inductive and deductive generalization, inductive and deductive specialization, and abstraction and concretion. Generalization and specialization transmutations change the reference set of a description, and abstraction and concretion change its level-of-detail.

Selected References

Michalski, R.S., "Inferential Theory of Learning: Developing Foundations for Multistrategy Learning," in *Machine Learning: A Multistrategy Approach, Volume IV*, Morgan Kaufmann Publishers, 1994.

Michalski, R.S., "Inferential Theory of Learning as a Conceptual Basis for Multistrategy Learning," *Machine Learning*, Special Issue on Multistrategy Learning, Vol. 11, pp. 111-151, 1993.

Michalski, R.S., LEARNING = INFERENCING + MEMORIZING: Basic Concepts of Inferential Theory of Learning and Their Use for Classifying Learning Processes, in *Cognitive Models of Learning*, Chipman, S. and Meyrowitz, A. (Eds.), 1992.

For more references, see Publication section.

Multistrategy Task-Adaptive Learning:MTL

(Michalski, Wnek, Kaufman, Utz, Vafaie, J. Zhang)

This project is concerned with developing a novel methodology for multistrategy learning, based on the Inferential Theory of Learning. The proposed methodology, called multistrategy task-adaptive learning (MTL) integrates a range of learning strategies, in particular, two basic and mutually complementary learning paradigms: empirical learning and analytical learning (see Figure beside). Empirical learning assumes that the learner does not have much prior knowledge relevant to the task of learning, while analytic learning assumes that the learner has sufficient knowledge to solve the problem in principle, but that knowledge is not directly applicable or efficient. Empirical learning is based primarily on inductive inference from facts, while analytical learning is based primarily on deductive inference from prior knowledge.

Other major learning strategies that are integrated in MTL include constructive induction, analogical learning, and abstraction. Constructive induction employs background knowledge to generate problem-relevant descriptive concepts, and through them derives the most plausible inductive hypotheses. Analogical learning transfers knowledge from one problem domain to another through an analysis of similarities between concepts or problem solving methods. Abstraction transfers a description from a high-detail level to a low-detail and more goal-oriented level.

MTL postulates that the learning strategy, or a combination thereof, should be based on the analysis of the learning task at hand. A learning task is defined by the input, learner's prior knowledge and the learner's goal(s). The learning goal(s) are viewed as a central factor in controlling a learning process. This research provides foundations for building advanced learning systems, and applying them to such tasks as knowledge acquisition, planning, problem solving, intelligent robots and knowledge extraction from databases.

Selected References

Michalski, R.S., "Toward a Unified Theory of Learning: Multistrategy Task-adaptive Learning," in *Readings in Knowledge Acquisition and Learning: Automating the Construction and Improvement of Expert Systems*, B.G. Buchanan and D.C. Wilkins, Morgan Kaufmann, San Mateo, 1993.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Knowledge Representation Using Dynamically Interlaced Hierarchies

(Michalski, Alkharouf, Utz)

This project concerns a development of a new type of knowledge representation that facilitates all kinds of inferences and is thus particularly relevant to the development of multistrategy task-adaptive learning. Dynamic Interlaced Hierarchies (DIH) is based on psychological research into human semantic memory structure and utilizes hierarchies as its basic organizational principle. By storing new knowledge as links between hierarchically organized concepts, a conceptual framework is constructed that can represent very diverse and complex forms of knowledge as well as various knowledge transformations.

DIH uses type and part hierarchies of concepts as background knowledge, or knowledge considered to be relatively stable and unchanging. Statements or facts are stored as links between concepts and are considered dynamic knowledge, as these links are constantly being created and modified, strengthened or weakened. These links have numeric factors (or 'merit parameters') attached that affect the strength of the relationship between the various concepts. Rules and dependencies are bi-directional, each with a separate forward and backward 'strength'.

Inference patterns such as generalization/specialization, abstraction/ concretion, and similarity are easily visualized in DIH. Also these inferences are facilitated, since the procedure consists of manipulating links between hierarchies. Creating new links between concepts represents learning. In this way learning builds upon the background knowledge of the hierarchies and the dynamic knowledge already in place.

Selected References

Alkharouf, N.W. and Michalski, R.S., "Multistrategy Task-Adaptive Learning Using Dynamic Interlaced Hierarchies: A Methodology and Initial Implementation of INTERLACE," *Proceedings of the Third International Workshop on Multistrategy Learning (MSL-96)*, Harpers Ferry, WV, May 23-25, 1996, pp. 117-124.

Hieb, M.R. and Michalski, R.S., "Multitype Inference in Multistrategy Task-adaptive Learning: Dynamic Interlaced Hierarchies," *Informatica: An International Journal of Computing and Informatics*, Vol. 17, No. 4, pp. 399-412, December 1993.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Cognitive Models of Plausible Reasoning

(Michalski, Sklar)

The ability to reason plausibly, that is to derive useful conclusions from imperfect premises, is one of the most remarkable properties of the human mind, and a key to understanding intelligent behavior. In plausible reasoning, the premises may be incomplete, uncertain, imprecise or only partially relevant to the task. Yet, people are able to make useful conclusions from premises. The initial core theory of human plausible reasoning was developed by Collins and Michalski (1990--see see MLI publications). The goals of this research are to develop a computational theory and models of plausible reasoning, to validate the theory by experiments involving the models and human subjects, and to apply it to developing a new approach to knowledge representation, filling gaps in databases, and dynamic recognition.

Selected References

Boehm-Davis, D., Dontas, K. and Michalski, R.S., "A Validation and Exploration of Structural Aspects of the Collins-Michalski Theory of Plausible Reasoning," *Reports of the Machine Learning and Inference Laboratory*, MLI 90-5, School of Information Technology and Engineering, George Mason University, January 1990.

Collins, A. and Michalski, R.S., "The Logic of Plausible Reasoning: A Core Theory," *Cognitive Science*, Vol. 13, pp. 1-49, 1989.

Michalski, R.S., Dontas, K. and Boehm-Davis, D., "Plausible Reasoning: An Outline of Theory and Experiments," *Proceedings of the Fourth International Symposium on Methodologies for Intelligent Systems*, pp. 17-19, Charlotte, NC, October 1989.

Dontas, K., "APPLAUSE: An Implementation of the Collins-Michalski Theory of Plausible Reasoning," M.S. Thesis, Computer Science Department, University of Tennessee, Knoxville, TN, August 1988.

For more references, see Publication section.

Learning Goals in Multistrategy Learning

(Michalski, Utz)

Learning arises from an intelligent individual's inability to reason and comprehend with its current knowledge. From prior research, every learning task requires background knowledge, sufficient inputs and a learning goal to achieve success. In multistrategy learning, though, the pupil faces additional complexity: here, the pupil must derive several learning goals for sequential and possibly parallel application in the learning process. The pupil must use these goals then to choose relevant inputs and necessary strategies (among several) in a timely way in order to acquire the "right" target knowledge.

The aims of this research project are twofold: (1) to experiment with a formalism to specify goals for multistrategy learning and (2) to construct a processing mechanism to generate and apply learning goals appropriately. This project is based on the MTL methodology supporting the Inferential Theory of Learning. Such a formalism must be accurate and complete to enable the processing mechanism to create explicit learning goals to understand the context, direct the procedure, and evaluate the results of multistrategy learning tasks. The formalism must be domain-independent as well as task-adaptive.

Proper specification is essential for success. Learning goals must be specified to enable an examination of any newly acquired or more efficient knowledge. When the examination indicates that the results are implausible or incompatible with the target knowledge, the process must be capable of trying again. It should reselect inputs or learning strategies on the advice of the original learning goals or, where necessary, regenerate alternative learning goals to redirect the learning task.

Selected References

Michalski, R.S. and Ram A., "Learning as Goal-Driven Inference," in *Goal-Driven Learning*, A. Ram & D. B. Leake (Eds.), MIT Press/Bradford Books, Cambridge, MA, 1995.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Data-driven Constructive Induction: AQ17-DCI

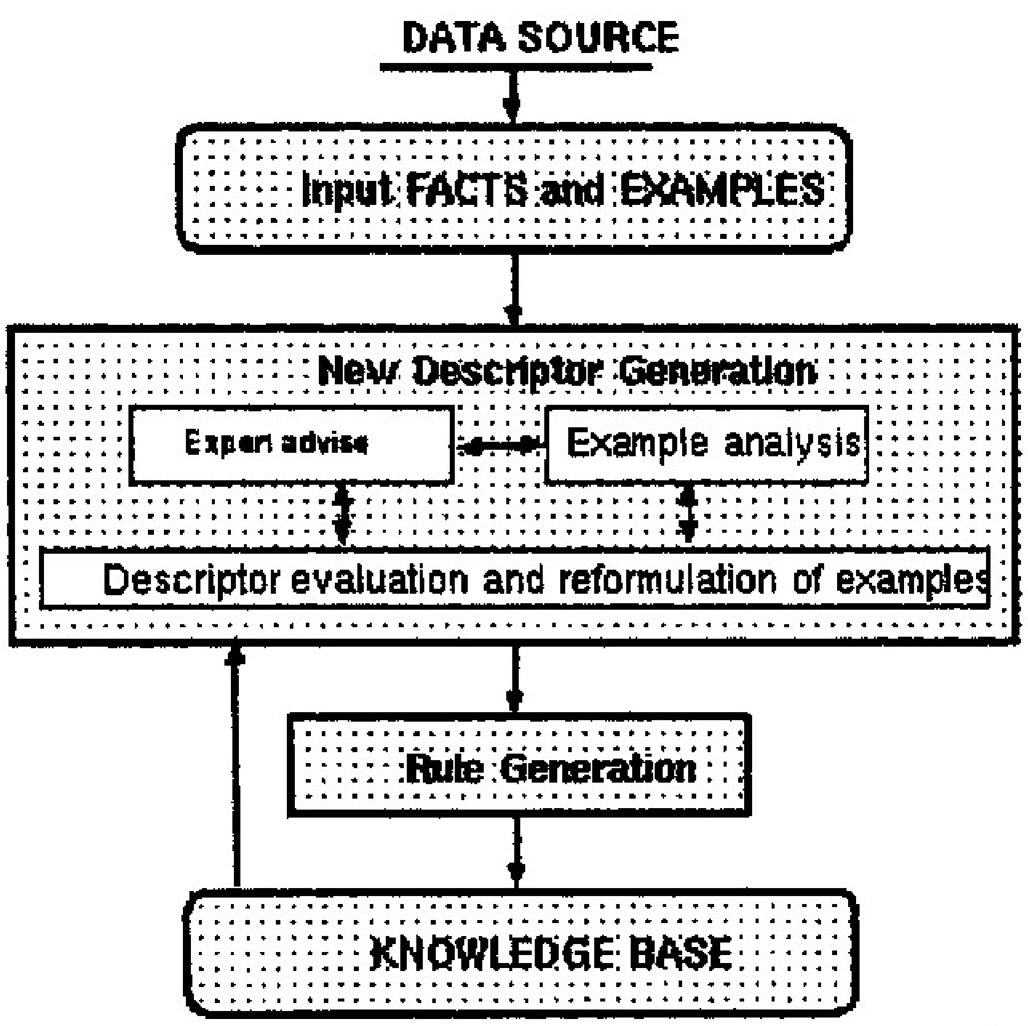
(Michalski, Bloedorn)

Most machine learning programs view the problem of learning an inductive hypothesis as a search for the "best" hypothesis in the given representation space. This works well if the problem is already well designed by some domain or machine learning expert using attributes which are relevant and simply related to the target concept.

However, finding a representation that is well suited to the problem is not a trival task. A number of different aspects have to be determined:

1) What attributes are relevant to the given task?
2) What values should those attributes take?
3) Are the concepts boundaries easily describable in the language and bias of the given learner?

In this project we have developed a method for automatically answering these questions. In our data-driven constructive induction approach we base decisions about the changes to make to the representation space on information and heuristics derived from the data. The data-driven approach can perform both expansions of the representation space through attribute construction, and reductions of the representation space through attribute removal, and abstraction.



A functional diagram of the DCI method. Changes to the Representation Space are based on Data and expert advise in the form of constraints provided by the user.

Data-driven constructive induction has been successfully applied to a number of different problems. These include artificial domains such as those in the 1st International Machine Learning Competition (Monk's Problems) to real-world domains involving predicting the voting pattern of members of the

House of Representatives to predicting the size of national Gross National Product (GNP) of countries around the world.

Selected References

Bloedorn, E. and Michalski, R.S., "The AQ17-DCI System for Data-Driven Constructive Induction and Its Application to the Analysis of World Economics," *Proceedings of the Ninth International Symposium on Methodologies for Intelligent Systems (ISMIS-96)*, Zakopane, Poland, June 10-13, 1996.

Bloedorn, E., and Michalski, R.S.,"Data-Driven Constructive Induction in AQ17-PRE: A Method and Experiments", *Proceedings of the IEEE International Conference on Tools for AI*, San Jose, CA, Nov. 1991. p. 30-27. Postscript, Compressed Poscript.

Bloedorn, E. and Michalski, R.S., "Constructive Induction from Data in AQ17-DCI: Further Experiments", *Reports of the Machine Learning and Inference Laboratory*, MLI91-12, George Mason University, Fairfax, VA, 1991. Postscript, Compressed Poscript.

For more references, see Publication section.

Hypothesis-driven Constructive Induction: AQ17-HCI

(Michalski, Wnek)

Traditional concept learning methods express the learned hypothesis using descriptors that are present in describing the training examples. In other words, they learn in the same representation space in which training examples are presented. For many practical problems this is a serious limitation, because concepts to be learned require descriptors that go beyond those originally provided.

To attack such problems, a constructive induction approach splits the learning process into two intertwined searches: one-for the most appropriate representation space for the given learning problem, and second -for the best inductive hypothesis in the newly created space.

A hypothesis-driven constructive induction method changes the concept representation spaces in the process of the concept learning. The changes involve expansion and contraction of the representation space, and are based on the analysis of consecutively created inductive hypotheses.

Selected References

Arciszewski, T., Michalski, R.S., Wnek, J., "Constructive Induction: the Key to Design Creativity," *Proceedings of the Third International Round-Table Conference on Computational Models of Creative Design*, Heron Island, Queensland, Australia, December 3-7, 1995.

Szczepanik, W., Arciszewski, T. and Wnek, J., "Empirical Performance Comparison of Two Symbolic Learning Systems Based On Selective And Constructive Induction," *Proceedings of the IJCAI- 95 Workshop on Machine Learning in Engineering*, Montreal, Canada, August, 1995.

Michalski, R.S. and Wnek, J., "Learning Hybrid Descriptions," *Proceedings of the 4th International Symposium on Intelligent Information Systems*, Augustow, Poland, June 5-9, 1995.

Wnek, J. and Michalski, R.S., "Conceptual Transition from Logic to Arithmetic," *Reports of Machine Learning and Inference Laboratory*, MLI 94-7, Center for MLI, George Mason University, Fairfax, VA, December 1994.

Wnek, J. and Michalski, R.S., "Discovering Representation Space Transformations for Learning Concept Descriptions Combining DNF and M-of-N Rules," *Working Notes of the ML-COLT'94 Workshop on Constructive Induction and Change of Representation*, New Brunswick, NJ, July 1994.

Bloedorn, E., Michalski, R.S., and Wnek, J., "Matching Methods with Problems: A Comparative Analysis of Constructive Induction Approaches",Reports of the Machine Learning and Inference Laboratory, MLI94-12, George Mason University, Fairfax, VA, 1994.

Arciszewski, E. Bloedorn, R.S. Michalski, M. Mustafa, J. Wnek, "Machine Learning in Conceptual Design: A Case Study on the Automated Acquisistion of Design Rules for Wind Bracings in Tall

Buildings Using Constructive Inductive Learning", ASCE Journal of Computing in CE, Vol. 8, No. 3. July 1994. George Mason University, Fairfax, VA, 1994.

Wnek, J. and Michalski, R.S., "Hypothesis-driven Constructive Induction in AQ17-HCI: A Method and Experiments," *Machine Learning*, Vol. 14, No. 2, pp. 139-168, 1994.

Wnek, J., *Hypothesis-driven Constructive Induction*, Ph.D. dissertation, School of Information Technology and Engineering, Reports of Machine Learning and Inference Laboratory, MLI 93-2, Center for Artificial Intelligence, George Mason University, (also published by University Microfilms Int., Ann Arbor, MI), March 1993.

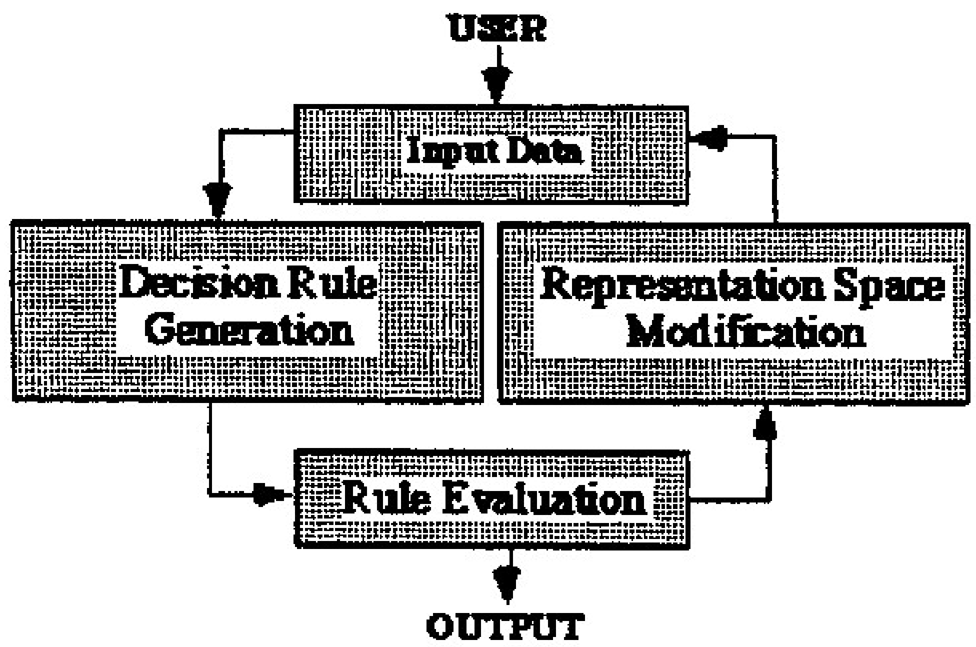
For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Multistrategy Constructive Induction: AQ17-MCI

(Michalski, Bloedorn, Wnek)

Conventional concept learning techniques generate hypotheses in the same representation space in which original training examples are presented. In many learning problems, however, the original representation space is inadequate for formulating the correct hypothesis. This inadequacy can be evidenced by a high degree of irregularity in the distribution of instances of the same class in the original representation space.



A functional diagram of the AQ17-MCI program.

We have been developing a methodology and a system, AQ17-MCI, for interpreting a range of strategies for an automated improvement of the knowledge representation spaces.

The system includes three basic mechanisms: (1) for accepting expert advice about the rules and procedures for generating new attributes;

(2) for analyzing learning examples and generating new attributes as logical or mathematical functions of the original attributes (implemented in AQ17-DCI version, which stands for data-driven constructive induction)

(3) for detecting "strong patterns" in the rules generated in one iteration of the rule generation module, and then using these patterns for proposing candidate attributes for a new iteration (implemented in AQ17-HCI version, which stands for hypothesis-driven constructive induction).

The attributes generated by these mechanisms are evaluated for their relevance to the problem at hand. If they pass the relevance test, they are used to reformulate original learning examples, and the rule generation module (based on the AQ algorithm) generates new rules. The quality of the rules is determined, and those that pass the quality criterion are stored in the knowledge base.

AQ17-MCI significantly extends current machine learning capabilities, as it is capable for "multi-mechanism" improvement of the original description space. It is a powerful program that represents a new generation of symbolic learning systems, and thus has a potential for important new applications. (See its performance on MONKS' problems, described in the project "A Comparative Study of Learning Methods.")

Selected Reference

Bloedorn, E.E., "Multistrategy Constructive Induction," *Ph.D. Dissertation,* School of Information Technology and Engineering, *Reports of Machine Learning and Inference Laboratory,* MLI 96-7, George Mason University, Fairfax, VA, 1996.

Bloedorn, E., Michalski, R.S., and Wnek, J., "Multistrategy Constructive Induction: AQ17-MCI", Proceedings of the 2nd International Workshop on Multistrategy Learning, May 26-29, 1993.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Constructive Induction in Engineering Design

(*Arciszewski, Michalski, Wnek, Bloedorn*)

The ultimate objective of this project is to develop a class of constructive induction methods for the applications to engineering design and a practical methodology for their use. A feasibility study has been completed and its results presented in the research report (Arciszewski et al 1992) published at the Center for Artificial Intelligence at George Mason University and in the ASCE Journal of Computing in Civil Engineering (Arciszewski et al, P94-16). The study was conducted in the area of conceptual design of wind bracings in steel skeleton structures of tall buildings.

Design rules were learned from a collection of 336 examples of minimum weight (optimal) designs of wind bracings. Constructive induction was used to produce design rules which explain how design requirements can be optimally (in terms of minimum steel weight) satisfied through the proper selection of individual components of a wind bracing structural system. All examples were prepared under identical design assumptions for a three-bay skeleton of a tall building in cooperation with practicing structural designers. Actual minimum-weight designs were produced using SODA, a computer system for optimization, analysis, and design of steel structures. The design rules obtained were divided into four classes corresponding to the value of the decision attribute: recommendation, standard, avoidance and infeasibility rules.

Two types of constructive induction have been used in the study: data-driven and hypothesis-driven constructive induction. The performance of both learning systems was formally measured by two empirical error rates: 1. the overall empirical error rate, 2. the omission error rate in accordance to the method of evaluation of performance of learning systems developed at the Laboratory and published in Arciszewski et al (1994). These error rates were calculated for the entire collection of examples using the leave-one-out resampling method. The error rates for constructive induction were compared with rates for the "traditional" induction, based on the use of the AQ15 algorithm. The individual error rates are shown in the table below. There is a significant improvement in performance (more than 50%) between the system based on the "traditional" induction and systems based on constructive induction. The difference in performance between two constructive induction-based systems is insignificant (less than 5%), but this may change as the research progresses.

Selected References

Chen, Q. and Arciszewski, T., "Machine Learning of Bridge Design Rules: A Case Study," *Proceedings of the 2nd ASCE Congress on Computing in Civil Engineering*, Atlanta GA, June, 1995.

Arciszewski, T., Bloedorn, E., Michalski, R.S., Mustafa, M. and Wnek, J., "Machine Learning of Design Rules: Methodology and Case Study," *ASCE Journal of Computing in Civil Engineering*, Vol. 8, No. 3, pp. 286-308, July 1994.

Arciszewski, T., "Machine Learning in Engineering Design," *Proceedings of the Conference on Intelligent Information Systems*, Institute of Computer Science, Polish Academy of Sciences, Wigry, Poland, 1994.

Wnek, J., Michalski, R.S. and Arciszewski, T., "An Application of Constructive Induction to Engineering Design," *Proceedings of the IJCAI-93 Workshop on AI in Design*, Chambery France, August 1993.

Arciszewski, T., Ziarko, W. and Khan, T.L., "Learning Conceptual Design Rules: A Rough Sets Approach," *Proceedings of the International Workshop on Rough Sets*, Banff, Alberta, Canada, 1993.

Arciszewski, T., Bloedorn, E., Michalski, R.S., Mustafa, M. and Wnek, J., "Constructive Induction in Structural Design," *Reports of the Machine Learning and Inference Laboratory*, School of Information Technology and Engineering, George Mason University, MLI 92-07, December 1992.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Constructive Induction Approach to Growing Neural Networks

(Sazonow, Wnek)

With most symbolic machine learning methods, if the given knowledge representation space is inadequate then the learning process will fail. This is also true with neural networks learning based methods. To overcome this limitation, a method for automatically "growing" neural network is being developed.

The BP-HCI method is a hypothesis-driven constructive induction for neural networks trained by the backpropagation algorithm. The method determines topology of a neural network and the initial connection weights based on patterns in the behavior of the neural network. The behavior of the neural network is captured by concepts called ACCORD and ANXIETY of a neural network.

The method was successfully applied to ten problems including such problems as learning t"exclusive-or" function, MONK2, parity-6BIT and inverse parity-6BIT.

Selected References

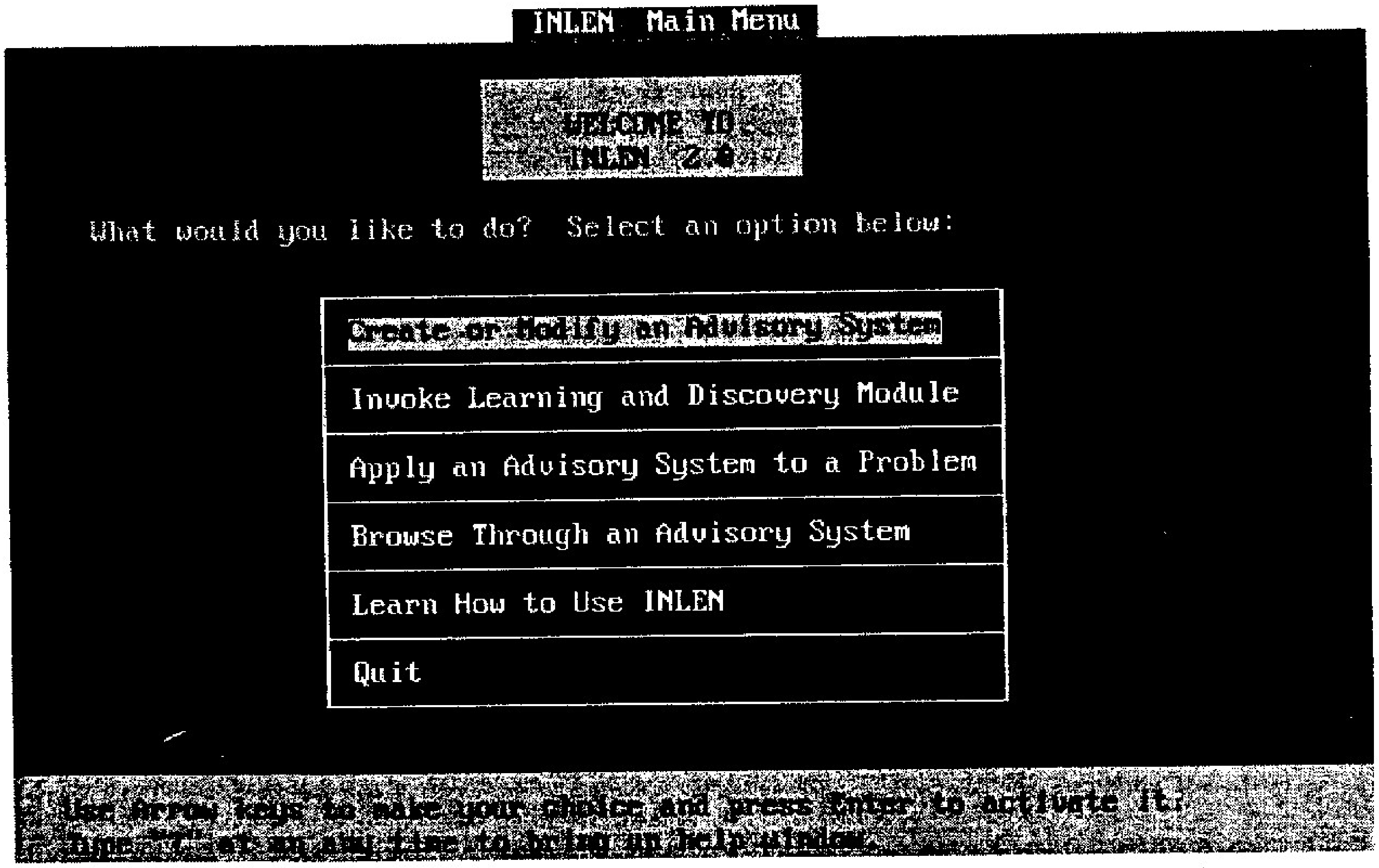
Sazonov, V.N. and Wnek, J., "Hypothesis-driven Constructive Induction Approach to Expanding Neural Networks," *Working Notes of the ML-COLT'94 Workshop on Constructive Induction and Change of Representation*, New Brunswick, NJ, July 1994.

For more references, see Publication section.

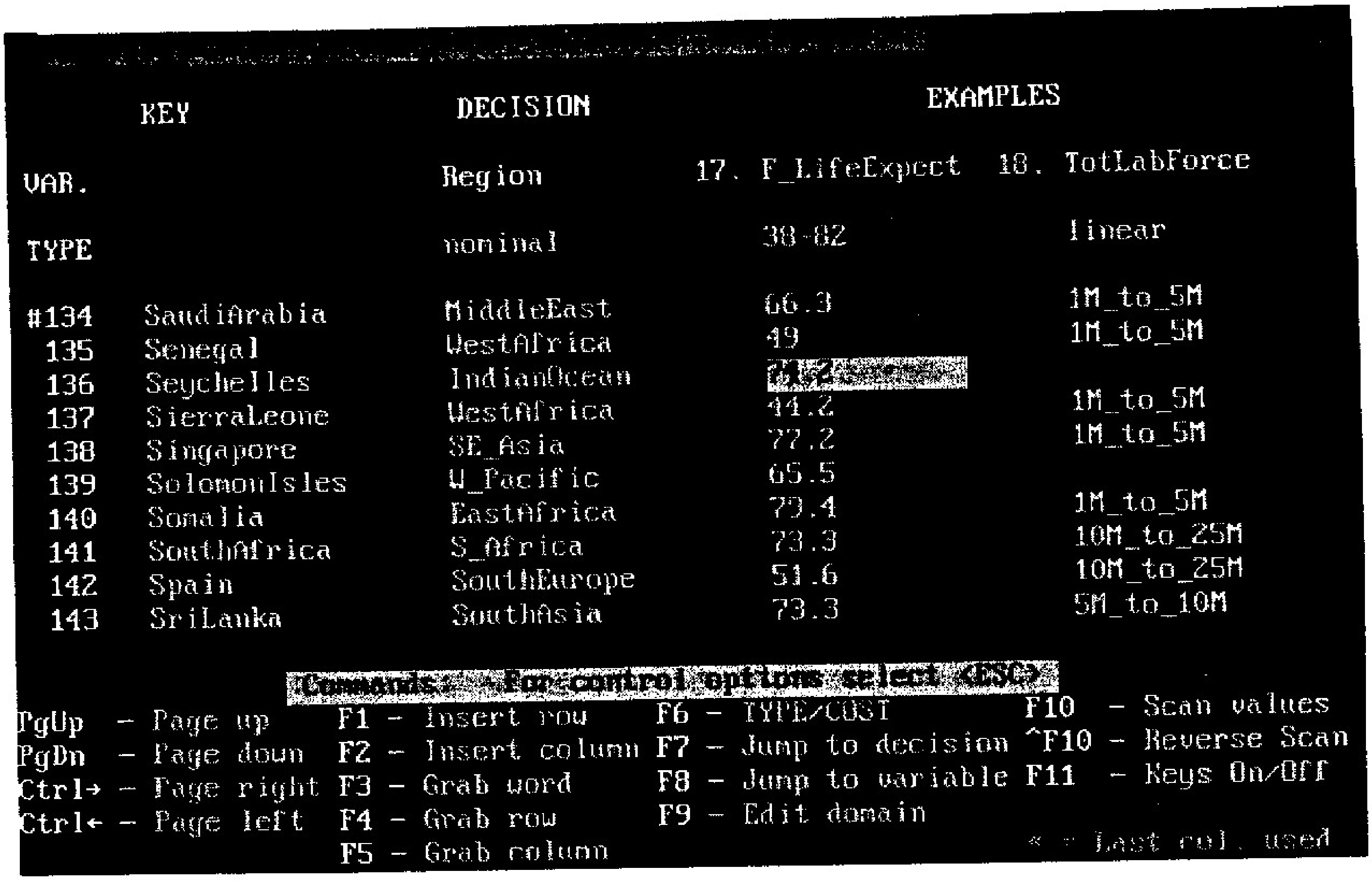
[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Knowledge Discovery in Databases: INLEN

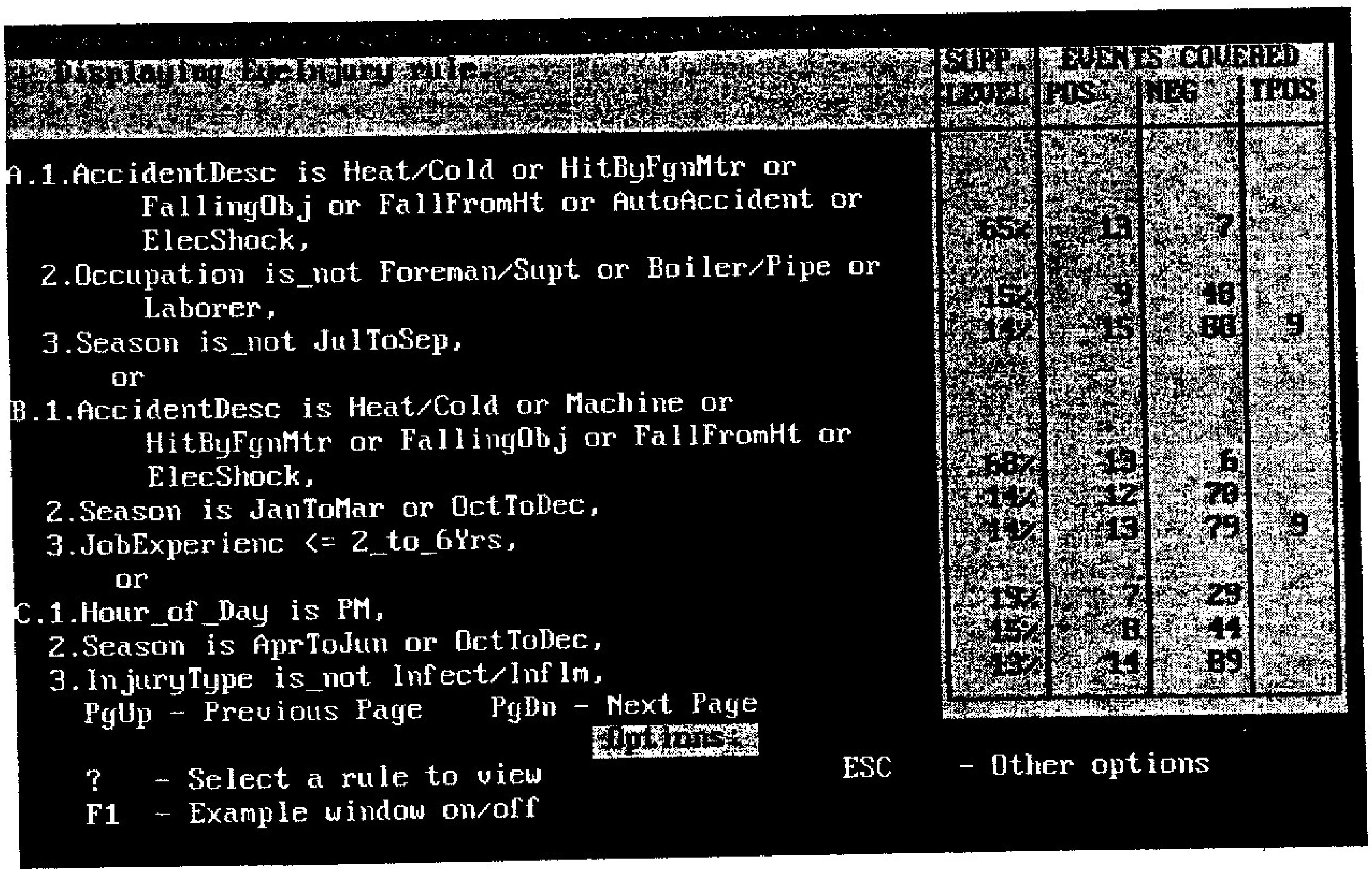
(Michalski, Kaufman, Mitchell, Kerschberg, Bloedorn, Wnek, Imam, Ribeiro)



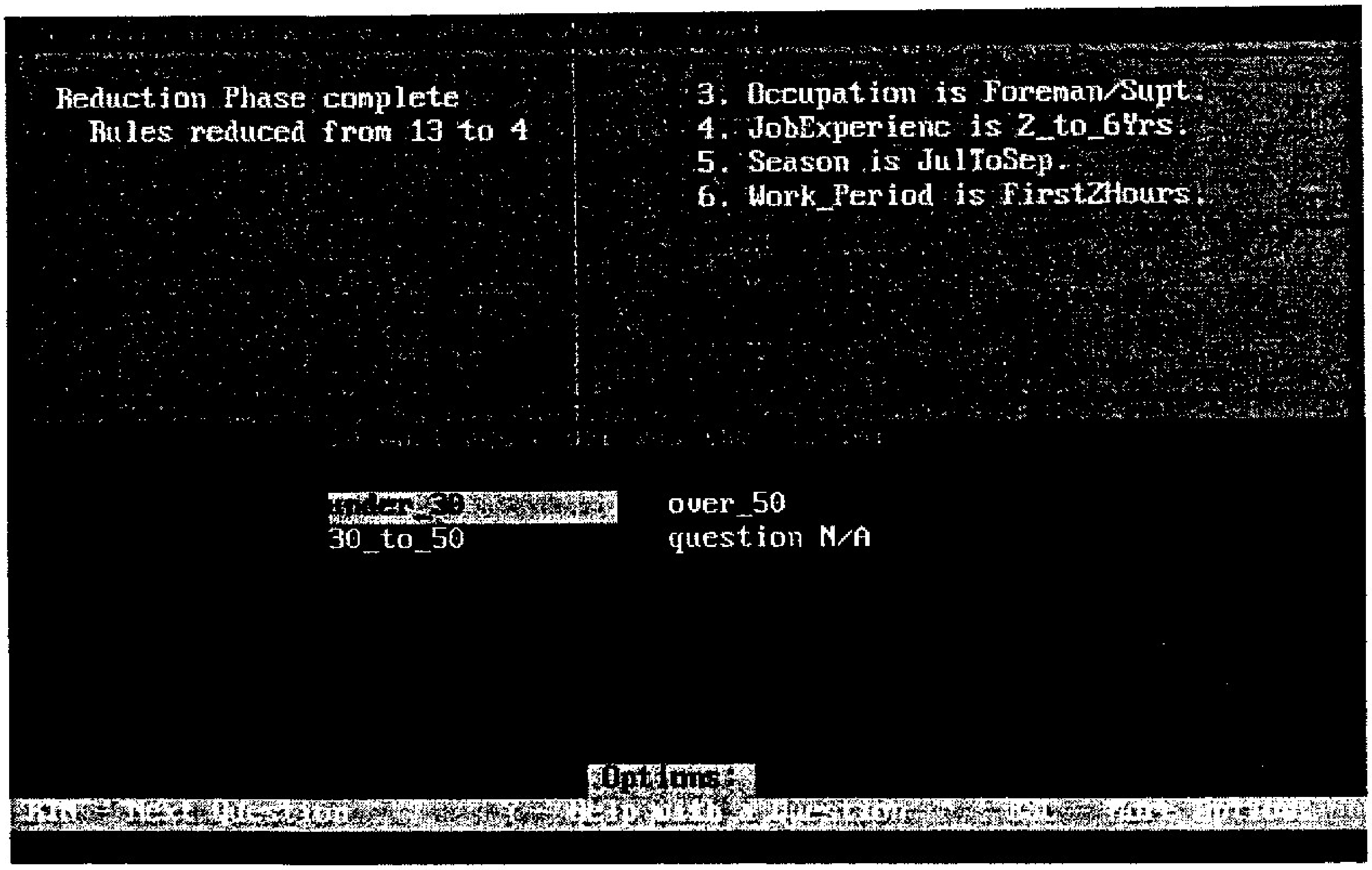
This project is concerned with the development of a large-scale multi-type reasoning system, called INLEN, for extracting knowledge from databases. The system assists a user in discovering general patterns or trends, meaningful relationships, conceptual or numerical regularities or anomalies in large databases. The volume of information in a database is often too vast for a data analyst to be able to detect such patterns or regularities. INLEN integrates symbolic learning and statistical techniques with database and knowledge base technologies. It provides a user with "knowledge generation operators" (KGOs) for discovering rules characterizing sets of data, generating meaningful conceptual classifications, detecting similarities and formulating explanations for the rules, generating rules and equations characterizing data, selecting and/or generating new relevant variables or representative examples, and testing the discovered rules on new data.



A screen in which the user may examine or modify the data set to be learned from. In this dataset each example describes a separate country. The Key field provides the country name (which is not learned from), and the values for the other attributes are presented in spreadsheet form



A screen displaying rules learned from an example set. In this example the rules are displayed which describe the conditions under which eye injuries have occurred on a construction site. The numbers of examples supporting each part of the rules are measured in the columns on the right.



An example screen from the INLEN Advisory Module, in which a user may be assisted in making a decision or can speculate on unknown data. The current question for the user is displayed in the middle of the screen, and INLEN's current best hypotheses are displayed in the top right.

Selected References

Kaufman, K.A. and Michalski, R.S., "KGL: A Language for Learning," *Reports of the Machine Learning and Inference Laboratory,* MLI 97-3, George Mason University, Fairfax, VA, 1997.

Kaufman, K. and Michalski, R.S., "A Method for Reasoning with Structured and Continuous Attributes in the INLEN-2 Knowledge Discovery System," *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*, Portland, OR, August, 1996, pp. 232-237.

Kaufman, K. and Michalski, R.S., "A Multistrategy Conceptual Analysis of Economic Data," Ein-Dor, P. (ed.), *Artificial Intelligence in Economics and Management: An Edithed Proceedings on the Fourth International Workshop*, Boston, Kluwer Academic Publishers, 1996, pp. 193-203.

Kaufman, K., "Addressing Knowledge Discovery Problems in a Multistrategy Framework," *Proceedings of the Third International Workshop on Multistrategy Learning (MSL-96)*, Harpers Ferry, WV, May 23-25, 1996, pp. 305-312.

Ribeiro, J., Kaufman, K. and Kerschberg, L., "Knowledge Discovery from Multiple Databases," *Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95)*, Montreal, Canada, August, 1995, pp. 240-245.

Michalski, R.S., Kerschberg, L., Kaufman, K.A. and Ribeiro, J.S., "Mining For Knowledge in Databases: The INLEN Architecture, Initial Implementation and First Results," *Intelligent Information Systems: Integrating Artificial Intelligence and Database Technologies*, Vol. 1, No. 1, pp. 85-113, August 1992.

Kaufman, K., Michalski, R.S. and Kerschberg, L., "Knowledge Extraction from Databases: Design Principles of the INLEN System," *Proceedings of the Sixth International Symposium on Methodologies for Intelligent Systems, ISMIS'91*, October 16-19, 1991.

Kaufman, K.A., Michalski, R.S. and Kerschberg, L., "Mining for Knowledge in Databases: Goals and General Description of the INLEN System," *Knowledge Discovery in Databases*, G. Piatetski-Shapiro and W.J. Frawley (Eds), AAAI Press/The MIT Press, Menlo Park, CA 1991.

Kaufman, K.A., Michalski, R.S. and Kerschberg, L., "Mining for Knowledge in Databases: Goals and General Description of the INLEN System," *Proceedings of IJCAI-89 Workshop on Knowledge Discovery in Databases*, Detroit, MI, August 1989.

Kaufman, K., Michalski, R.S., Zytkow, J. and Kerschberg, L., "The INLEN System for Extracting Knowledge from Databases: Goals and General Description," *Reports of the Machine Learning and Inference Laboratory*, MLI 89-6, School of Information Technology and Engineering, George Mason University, Fairfax, VA, 1989.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Learning Problem-Oriented Decision Structures from Decision Rules

(Michalski, Imam)

This project is concerned with learning problem-optimized decision trees from rules. A standard approach to determining decision trees is to learn them from examples. A disadvantage of this approach is that once a decision tree is learned, it is difficult to modify it to suit different decision making situations. Such problems arise, for example, when an attribute assigned to some node cannot be measured, or there is a significant change in the costs of measuring attributes or in the frequency distribution of events from different decision classes. An attractive approach to resolving this problem is to learn and store knowledge in the form of decision rules, and to generate from them, whenever needed, a decision tree that is most suitable in a given situation.

An additional advantage of such an approach is that it facilitates building compact decision trees, which can be much simpler than the logically equivalent conventional decision trees (by compact trees are meant decision trees that may contain branches assigned a set of values, and nodes assigned derived attributes, i.e., attributes that are logical or mathematical functions of the original ones). The project describes an efficient method, AQDT-1, that takes decision rules generated by an AQ-type learning system (AQ15 or AQ17), and builds from them a decision tree optimizing a given optimality criterion.

The method can work in two modes: the standard mode, which produces conventional decision trees, and compact mode, which produces compact decision trees. The preliminary experiments with AQDT-1 have shown that the decision trees generated by it from decision rules (conventional and compact) have outperformed those generated from examples by the well-known C4.5 program both in terms of their simplicity and their predictive accuracy.

Selected References

Imam, I.F. and Michalski, R.S., "An Empirical Comparison Between Learning Decision Trees from Examples and from Decision Rules," *Proceedings of the Ninth International Symposium on Methodologies for Intelligent Systems (ISMIS-96)*, Zakopane, Poland, June 10-13, 1996.

Imam, I.F. and Michalski, R.S., "Learning Decision Trees from Decision Rules: A Method and Initial Results from a Comparative Study," *Reports of the Machine Learning and Inference Laboratory*, MLI 93-6, School of Information Technology and Engineering, George Mason University, May 1993.

Imam, I.F. and Michalski, R.S., "Should Decision Trees Be Learned from Examples or from Decision Rules?" *Lecture Notes in Artificial Intelligence*, Springer Verlag; *Proceedings of the 7th International Symposium on Methodologies for Intelligent Systems, ISMIS*, Trondheim, Norway, June 15-18, 1993.

Imam, I.F. and Michalski, R.S., "Learning Decision Trees from Decision Rules: A Method and Initial Results from a Comparative Study," *Journal of Intelligent Information Systems JIIS*, L. Kerschberg, Z. Ras and M. Zemankova (Eds.), Vol. 2, No. 3, pp. 279-304, Kluwer Academic, Boston, MA, 1993.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Intelligent Agents with Learning Capabilities

(Michalski, Kaufman, Imam, Ribeiro, Wnek)

Standard expert systems do not have learning capabilities. Their knowledge bases are built entirely by hand-encoding of an expert's knowledge. Such a process is time-consuming and prone to error. This project is concerned with the development of a PC-based expert system shell with learning capabilities. The system incorporates a knowledge base for storing rules and a data base for storing facts and examples. It has a learning program for rule acquisition, and a powerful inference mechanism.

The project is based on our earlier experience with ADVISE and AURORA systems. ADVISE is a large-scale inference system with rule learning capabilities and multiple control schemes. The system served as a laboratory for experimenting with methods for knowledge acquisition, multiple knowledge representation and machine learning. Aurora is a PC-based inference system, and an expert system shell that incorporates a program for incremental rule learning and improvement.

A related project concerns a method for discovering qualitative and quantitative models from data characterizing the behavior of a system. This method builds upon our experience with the ABACUS system for quantitative discovery. The current system is capable of determining a set of equations that fit a given set of datapoints, and a set of symbolic descriptions characterizing preconditions for the application of these equations. ABACUS integrates methods for data-driven quantitative discovery, concept learning from examples and conceptual clustering. This research has applications in building advanced expert systems and discovering quantitative and qualitative regularities in data. This project is being done in collaboration with AGH.

Selected References

Bloedorn, E. and Wnek, J., "Constructive Induction-based Learning Agents: An Architecture and Preliminary Experiments," *Proceedings of the First International Workshop on Intelligent Adaptive Systems (IAS-95)*, pp. 38-49, Melbourne Beach, FL, April 26, 1995.

Imam, I., "Intelligent Agents for Management of Learning: An Introduction and a Case Study," *Proceedings of the First International Workshop on Intelligent Adaptive Systems (IAS-95)*, pp. 95-106, Melbourne Beach, FL, April 26, 1995.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Multi-Level Image Sampling and Interpretation: MIST Methodology

(Michalski, Duric, Zhang, Maloof)

The goal of this project is to develop a system that can learn descriptions of visual objects (images, visual sources, visual scenes) and to use these descriptions to recognize unknown objects. We have developed a general methodology for this purpose, called multi-level image sampling and interpretation (MIST).

The basic idea under this project can be explained as follows. Given an image with labeled samples of different surfaces, the learning system determines a sequence of operators that transform the image to a "symbolic" image, in which picture elements are labels of corresponding surface areas. The sequence of operators that produces such a labeling serves as a surface description ("surface signature"). A surface description is a logical expression in disjunctive normal form associated with a decision class (here, a texture class). Each conjunction in this expression together with the associated decision class can be viewed as a single decision rule. The basic operator in the process of generating surface description is an application of a set of logic-style rules to transformed surface samples. The rules can be applied in parallel, and serve as "logical templates" that are applied to "events" (attribute vectors) representing surface samples. To recognize an unknown surface sample, the system matches it with all candidate surface descriptions. This is done by applying decision rules to the events in the sample. For each event, the class membership (surface class) is determined. The assignment of the sample to a given decision class (surface) is based on determining which of the candidate classes gets the majority (or) plurality of votes. Thus, even if some events in the sample are incorrectly recognized, the classification of the sample may be correct.

A series of experiments is conducted with gradually increased complexity of data, increased influence of noise, and under variety of other external conditions. The images used for experiments are divided into two groups: office environment images, and outdoor scene images.

Selected References

Michalski, R.S., Zhang, Q., Maloof, M.A. and Bloedorn. E., The MIST Methodology and its Application to Natural Scene Interpretation, *Proceedings of the Image Understanding Workshop*, Palm Springs, CA, Feburary, 1996.

Zhang, Q., Duric Z.,Maloof, M.A. and Michalski, R.S., "Target detection in SAR images using the MIST/AQ method" *Reports of Machine Learning and Inference Laboratory,* MLI 96-12, George Mason University, Fairfax, VA, 1996.

Bala, J. and Michalski, R.S., "Learning Texture Concepts Through Multilevel Symbolic Transformations," *Proceedings of the Third International Conference on Tools for Artificial Intelligence*, San Jose, CA, November 9-14, 1991.

Channic, T., "TEXPERT: An Application of Machine Learning to Texture Recognition," M.S. Thesis, University of Illinois, Urbana-Champaign, 1988.

For more references, see Publication section.

Machine Vision and Learning

(Michalski, Duric, Zhang, Maloof, Bloedorn)

The goal of this project is to develop methods and experimental vision systems that are capable of learning general visual concept descriptions from specific observed objects, and then use these descriptions to efficiently recognize new objects in a visual scene. It is assumed that the system should be able to recognize objects among other objects in a scene under a variety of conditions, such as changing viewpoints, changing illumination, object overlap, and in the presence of noise in the sensory data.

Our approach is based on a two-pronged architecture in which the first prong processes the surface information about objects, and the second prong processes the shape information. Learning unique surface characteristics ("surface signatures") involves problem-oriented transformations of the representation space, and an iterative application of an inductive learning program. The input to the system are classified samples of surfaces.

In the object recognition phase, the system applies the learned rules to identify the surface, and then uses this information to generate a set of candidate hypotheses about the object's identity.

These hypotheses are then employed to retrieve specific 3D structural models of the objects from a knowledge base. We use CAD/CAM descriptions of objects to discover their characteristic structural and symbolic features and feature relations, and to learn recognition strategies. These processes are also driven by vision tasks, such as localization, recognition and inspection. Learned models are then used to determine characteristics of objects sufficient to identify the object in the scene. These discriminatory characteristics are determined by a process called dynamic recognition.

The research on this project is conducted in collaboration with the Computer Vision Laboratory of the University of Maryland.

Our laboratory, in collaboration with the UMD Computer Vision Laboratory, organized the NSF/ARPA workshop on Machine Vision and Learning. The workshop was held in Harpers Ferry, WV, in October 15-17, 1992. It was the first workshop that brough together leading researchers in computer vision and machine learning. Below is a reference to the report that was partially based on this workshop and partially based on new material:

Selected References

Michalski R.S., Rosenfeld, A., Aloimonos Y., Duric, Z., Maloof M.A., Zhang Q., "Progress On Vision Through Learning: A Collaborative Effort of George Mason University and University of Maryland, *Proceedings of the Image Understanding Workshop*, Palm Springs, CA, Feburary, 1996.

Michalski, R.S., Zhang, Q., Maloof, M.A. and Bloedorn. E., The MIST Methodology and its Application to Natural Scene Interpretation, *Proceedings of the Image Understanding Workshop*, Palm Springs, CA, Feburary, 1996.

Zhang, Q., Duric Z.,Maloof, M.A. and Michalski, R.S., "Target detection in SAR images using the MIST/AQ method" *Reports of Machine Learning and Inference Laboratory,* MLI 96-12, George Mason University, Fairfax, VA, 1996.

Michalski, R.S., Rosenfeld, A., and Aloimonos, Y., "Machine Vision and Learning: Research Issues and Directions," *Reports of the Machine Learning and Inference Laboratory*, MLI 94-6, Machine Learning and Inference Laboratory, George Mason University, Fairfax, VA; Reports of the Center for Automation Research CAR-TR-739, CS-TR-3358, University of Maryland, College Park, MD, October 1994.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Multistrategy Learning in Vision: Integrating Symbolic and Neural Net Learning for Vision Tasks

(Michalski, Zhang)

The project concerns the development of a novel multistrategy learning methodology that is specifically oriented toward vision learning. The methodology combines symbolic rule learning and neural-based learning strategies in order to achieve high efficiency and accuracy in learning visual object descriptions, and in applying these descriptions to rapid object recognition.

The initially developed vision system has several advantages: it can be easily modified and applied to new problems (due to learning), its learning speed can be at least an order of magnitude faster than neural net learning (due to symbolic pre-structuring of the net), it has short recognition times (due to its parallel architecture), and its underlying recognition rules are easy to understand by a human operator (due to the symbolic knowledge representation of the basic decision rules). The developed system was experimentally applied to natural scene recognition.

The method works in two stages: 1) rule learning using the AQ algorithm. This phase generates rules that generally and approximately describe the training examples, 2) neural net learning to determine the final visual concept description. The network is structured according to the rules obtained in stage. Each node in the hidden-layer of the network corresponds to a single rule. The degree of match of an example to the rule represents node activation. This activation value is input to the sigmoid transfer function associated with each node. Weight values for the connections between nodes and outputs are obtained using backpropagation method.

Selected References

Michalski, R.S., Zhang, Q., Maloof, M.A. and Bloedorn. E., The MIST Methodology and its Application to Natural Scene Interpretation, *Proceedings of the Image Understanding Workshop*, Palm Springs, CA, Feburary, 1996.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

Learning to Recognize Shapes

(Michalski, Duric, Maloof)

The goal of this research is to apply inductive learning methods to problems of 2D shape recognition under highly variable perceptual conditions. The multilevel logical template (MLT) methodology is being used to detect blasting caps in x-ray images of luggage. An intelligent system capable of quickly and reliably performing this task could be used to assist airport security personnel in baggage screening.

We have acquired x-ray images of luggage containing blasting caps that appear at differing degrees of occlusion and at various orientations with respect to the x-ray source. Task-oriented image transformations are used to segment blasting caps and other objects, and to extract training events, which are vectors of attribute values. These training events serve as input to the learning process which induces descriptions of shape that are robust with respect to planar rotation and translation and partial occlusion. Induced shape descriptions can be used to recognize unknown objects.

Various symbolic, non-symbolic and statistical learning approaches are being investigated for acquiring descriptions of shape, including AQ15c, neural networks, and k-nn. These learning approaches are compared using predictive accuracy, and learning and recognition time. Experimental results have demonstrated strong advantages of AQ15c over neural networks and k-nn. AQ15c also has the advantage of producing comprehensible symbolic descriptions that can be optimized by either a human or by a machine process in post-learning phases.

Selected References

Maloof, M.A., Michalski, R.S., "Learning Symbolic Descriptions of Shape for Object Recognition In X-Ray Images," *Expert Systems with Applications,* 12(1), 11-20, 1997.

Maloof, M.A. and Michalski, R.S., "Learning Symbolic Descriptions of 2D Shapes for Object Recognition in X-ray Images," *Proceedings of the 8th International Symposium on Artificial Intelligence*, Monterrey, Mexico, October 17-20, 1995.

Maloof, M. and Michalski, R.S., "Learning Descriptions of 2D Shapes for Object Recognition and X-Ray Images," *Reports of the Machine Learning and Inference Laboratory*, MLI 94-4, Machine Learning and Inference Laboratory, George Mason University, Fairfax, VA, October 1994.

For more references, see Publication section.

Dynamic Recognition

(Michalski, Bloedorn)

Any recognition process involves making a connection between a concept representation stored in the system's memory and a stream of observational data. Present recognition systems attempt to recognize objects by matching descriptions with the data stream. If the input data satisfies rules characterizing an object, the object is recognized. To implement such a system for practical tasks, a very large number of rules may be required.

This aspect severely limits present recognition systems, as it prevents them from being applied to the recognition of a large number of objects. In contrast to this approach, humans can recognize objects from a great variety of different cues, without "matching" rules. For example, they can recognize a known person from seeing a face, a silhouette, the characteristic way of walking, hearing the person's voice, or even from observing the person's gesticulation or seeing his/her shoes.

The dynamic recognition (DR) approach (initially proposed by Michalski in 1986) overcomes this problem by using inductive inference to dynamically determine discriminant object descriptions from characteristic object descriptions, and this allows the system to avoid matching rules. Only one characteristic description per concept is stored in memory. Potentially, the DR method can efficiently handle a great variety of different practical recognition problems. An initial implementation of the system has strongly supported the theoretical expectations.

Selected References

Michalski, R, S,, "Dynamic Recognition: An Outline of Theory of How to Recognize Concepts without Matching Rules," *Reports of the Intelligent Systems Group*, ISG 86-16, UIUCDCS-F-86-965, Urbana, 1986.

Michalski, R. S., "A Variable-Valued Logic System as Applied to Picture Description and Recognition," Chapter in the book, *Graphic Languages*, F. Nake and A. Rosenfeld (Editors), North-Holland Publishing Co., 1972.

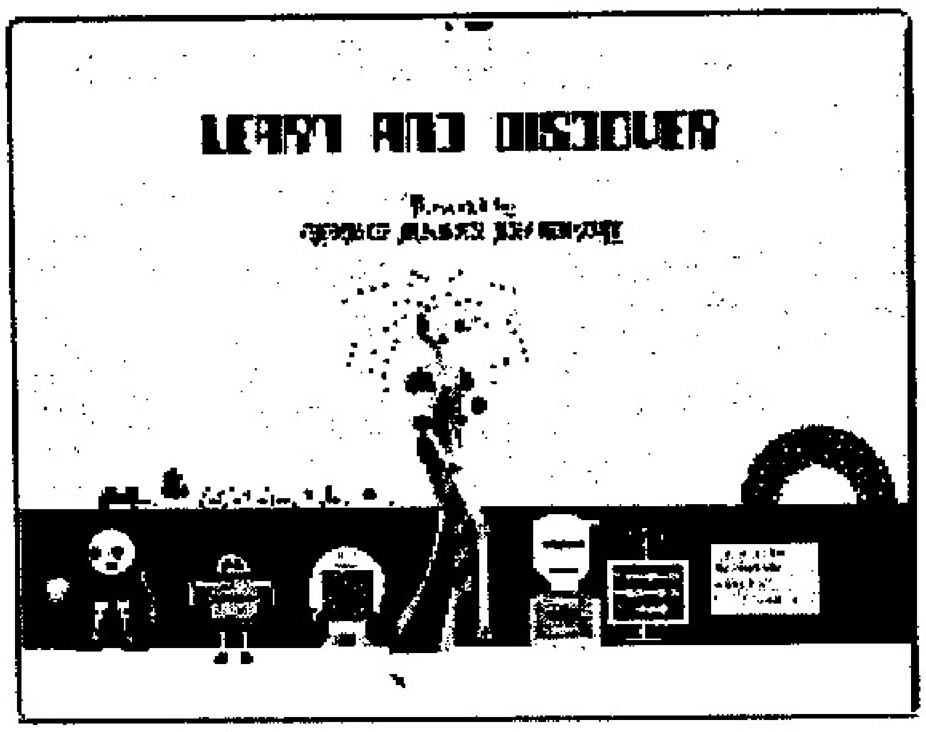
For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

EMERALD:

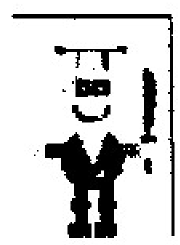
An Integrated Large-Scale Learning and Discovery System for Education and Research in Machine Learning

(Michalski, Kaufman, Lee, Bloedorn, De Jong, Schultz, Wnek)

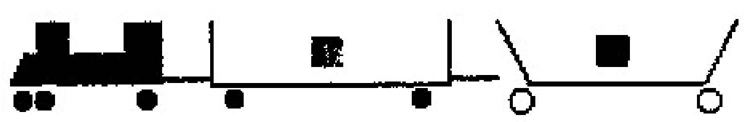


This project concerns the development and maintenance of an integrated system for machine learning and discovery, EMERALD, that serves as a tool for education and research in machine learning and cognitive modeling of learning processes. The system is regularly used in teaching the course INFT 811: "Principles of Machine Learning and Inference" and occasionally some other courses.

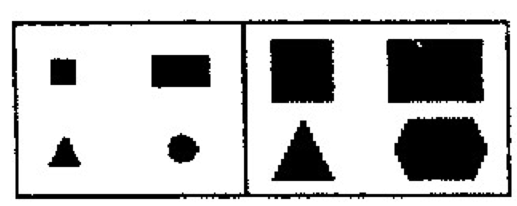
The EMERALD system (Experimental Machine Example-based Reasoning and Learning Disciple) consists of five modules ("robots"), each displaying a capability for some form of learning or discovery:



AQ learns general decision rules from examples of different classes of correct or incorrect decisions made by experts. An example of the program performance employs little robots like these:

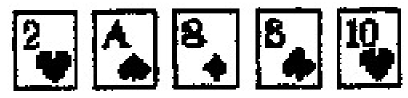


INDUCE learns structural descriptions of groups of objects, and determines important distinctions between the groups. An example of INDUCE performance uses little trains like these:



CLUSTER creates meaningful categories and classifications of given entities, and formulates descriptions of these created categories. One of the examples illustrating CLUSTER performance involves the clustering of various geometric objects into differenc classes as shown here.

SPARC predicts possible future objects or events by discovering rules characterizing the sequence of objects or events observed so far. One of the examples illustrating SPARC performance involves the prediction of a sequence of cards in the game ELEUSIS.



ABACUS conducts experiments, collects data, discovers mathematical and logical descriptions of data, and then uses these descriptions for predicting the behavior of some phenomenon. One of the experiments illustrating ABACUS involves the discovery of a law characterizing bodies falling through different media.

Each module is represented by a robot figure and employs a different voice (through a voice synthesizer) for communicating with the user.

An earlier and smaller version of the system, called ILLIAN, was a part of the exhibition "Robots and Beyond: The Age of Intelligent Machines," organized by a consortium of eight U.S. Museums of Science (Boston, Charlotte, Fort Worth, Los Angeles, Seattle, Chicago, Minneapolis and Columbus). Support for the development of the exhibit version was provided in part by the Boston Museum of Science, Digital Equipment Corporation, and the University of Illinois at Urbana-Champaign. The system was seen by few hundred thousand people.

EMERALD is the first system of its kind ever built, which integrates several learning capabilities with natural language processing, voice communication, and a highly user-oriented graphical interface. It enables users to experiment on-line with various learning and discovery programs under a unified control, and to use predefined objects to set different learning tasks for the system. EMERALD was developed under the direction of Professor R.S. Michalski in collaboration with his students and associates. The system has recently been adapted for SUN workstations, and used in teaching machine learning.

EMERALD has been distributed to a number of U.S. and European universities and organizations. **If you are interested in obtaining EMERALD, contact Dr. J. Wnek (jwnek@gmu.edu) or Ken Kaufman (kaufman@aic.gmu.edu).**

Although EMERALD modules are demonstrated in the context of certain predefined classes of problems, they are not specifically oriented toward these problems and objects. These modules are domain-independent programs that have already been used or have a potential to be used for concept learning and discovering regularities in such fields as medicine, agriculture, engineering, biology, chemistry, plant control, financial decisions, air traffic control, computer vision and intelligent robots.

Selected References

Kaufman K. A. and Michalski, R. S., "EMERALD 2: An Integrated System of Machine Learning and Discovery Programs for Education and Research, Programmer's Guide for the Sun Workstation

(Updated Edition), *Reports of the Machine Learning and Inference Laboratory*, MLI 97-9, George Mason University, Fairfax, VA, 1997.

Kaufman K. A. and Michalski, R. S., "EMERALD 2: An Integrated System of Machine Learning and Discovery Programs for Education and Research, User's Guide (Updated Edition), *Reports of the Machine Learning and Inference Laboratory*, MLI 97-8, George Mason University, Fairfax, VA, 1997.

Kaufman, K.A. and Michalski, R.S., "EMERALD: An Integrated System of Machine Learning and Discovery Programs to Support Education and Experimental Research," *Reports of the Machine Learning and Inference Laboratory*, MLI 93-10, School of Information Technology and Engineering, George Mason University, Fairfax, VA, September 1993.

For more references, see Publication section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

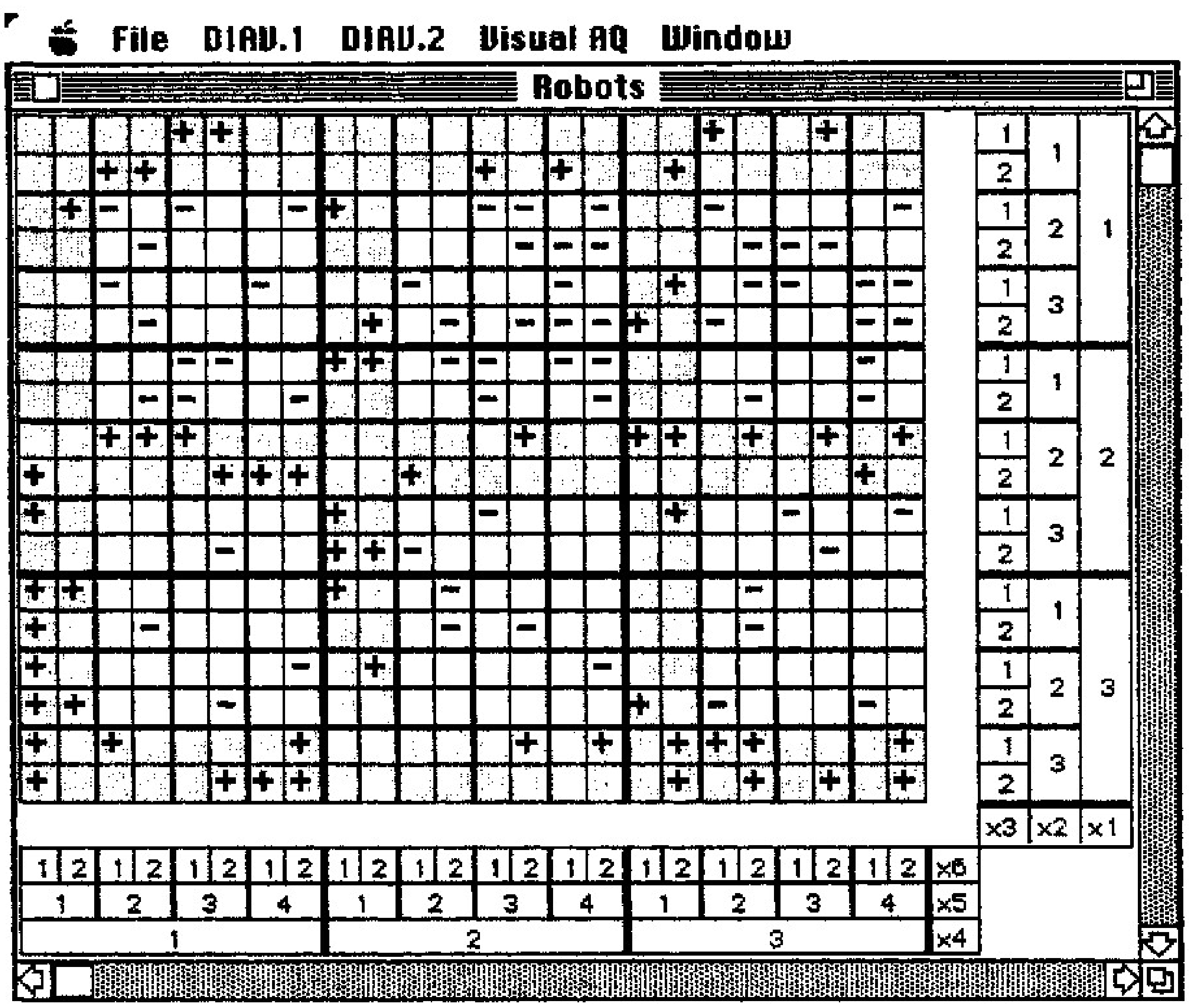
KNOWLEDGE VISUALIZER (KV)

A Diagrammatic Visualization of Data Mining and Machine Learning Processes

(Michalski, Zhang, Wnek)

The KV project concerns the development of a system for visualizing data mining, machine learning and knowledge discovery processes involving discrete multi-dimensional functions. It employs a planar model of a discrete multidimensional space, called *generalized logic diagram or GLD*, proposed by Michalski (1978). The diagram is spanned over a set of discrete attributes and consists of cells, each representing one unique combination of attribute values (a vector of attribute values). Thus, there are as many cells as there are possible vectors of attribute values. To determine the cell corresponding to a given vector, one seeks the intersection of the areas corresponding to the values of individual attributes.

For example, in the diagram below, the top-left cell represents the vector:
(x1 = 1, x2 = 1, x3 = 1, x4 = 1, x5 = 1, x6 = 1).



In the diagram above, Positive examples of a concept are visualized using "+", counter-examples of the concept are visualized using "-". A decision rule (a conjuction of conditions on attribute values) corresponds to regular arrangement of cells that can be easily recognized visually. A concept description is in the form of a collection of such decision rules (a ruleset). For example, the yellow area in the diagram represents a concept desciption described by the disjunction of two rules:

R1: [x5 = 1]
R2: [x1 = x2]

If the target and learned concepts are represented in the diagram, then their set-difference denotes errors in the learned concept ("error area").

The diagram can also illustrate results of any operation on the concept, such as generalization or specialization, or any change of the description space, such as adding or deleting attributes, or their values. Another interesting feature is that it can also visualize concepts acquired by non-symbolic systems, such as neural nets or genetic algorithms. Using the diagram one can directly express the learned concepts in the form of decision rules. Thus, the diagram allows one to evaluate both the quality and the complexity of the results of symbolic, as well as non-symbolic learning.

We have implemented two systems: DIAV-2 in Smalltalk, and KV in Java. These systems can display description spaces with up to one million events, i.e., spaces spanned over up to 20 binary variables (or a correspondingly smaller number of multiple-valued variables). The systems have proven to be very useful for analyzing behavior of learning algorithms. They are available to universities and industrial organizations.

Selected References

Zhang, Q. "Knowledge Visualizer User's Guide," *Reports of the Machine Learning and Inference Laboratory*, George Mason University, Fairfax, VA, 1997 (in preparation).

Wnek, J., "DIAV 2.0 User Manual: Specification and Guide through the Diagrammatic Visualization System," *Reports of the Machine Learning and Inference Laboratory*, MLI 95-5, George Mason University, Fairfax, VA, 1995.

Michalski, R.S. and Wnek, J., "Learning Hybrid Descriptions," *Proceedings of the 4th International Symposium on Intelligent Information Systems*, Augustow, Poland, June 5-9, 1995.

Wnek, J., Kaufman, K., Bloedorn, E. and Michalski, R.S., "Selective Induction Learning System AQ15c: The Method and User's Guide," *Reports of the Machine Learning and Inference Laboratory*, MLI 95-4, George Mason University, Fairfax, VA, March 1995.

Wnek, J. and Michalski, R.S., "Discovering Representation Space Transformations for Learning Concept Descriptions Combining DNF and M-of-N Rules," *Working Notes of the ML-COLT'94 Workshop on Constructive Induction and Change of Representation*, New Brunswick, NJ, July 1994.

Wnek, J. and Michalski, R.S., "Conceptual Transition from Logic to Arithmetic," *Reports of the Machine Learning and Inference Laboratory*, MLI 94-7, Center for Machine Learning and Inference, George Mason University, Fairfax, VA, December 1994.

Wnek, J. and Michalski, R.S., "Hypothesis-driven Constructive Induction in AQ17-HCI: A Method and Experiments," *Machine Learning*, Vol. 14, No. 2, pp. 139-168, 1994.

Wnek, J. and Michalski, R.S., "Comparing Symbolic and Subsymbolic Learning: Three Studies," in *Machine Learning: A Multistrategy Approach, Vol. 4.*, R.S. Michalski and G. Tecuci (Eds.), Morgan

Kaufmann, San Mateo, CA, 1994.

Wnek, J., *Hypothesis-driven Constructive Induction*, Ph.D. dissertation, School of Information Technology and Engineering, Reports of Machine Learning and Inference Laboratory, MLI 93-2, Center for Artificial Intelligence, George Mason University, (also published by University Microfilms Int., Ann Arbor, MI), March 1993.

Wnek, J., Sarma, J., Wahab, A. and Michalski, R.S., "Comparing Learning Paradigms via Diagrammatic Visualization: A Case Study in Concept Learning Using Symbolic, Neural Net and Genetic Algorithm Methods," *Proceedings of the 5th International Symposium on Methodologies for Intelligent Systems - ISMIS'90*, Knoxville, TN, pp. 428-437, October 1990.

Michalski, R.S., "A Planar Geometric Model for Representing Multi-dimensional Discrete Spaces and Multiple-valued Logic Functions," *Reports of Computer Science Department*, Report No. 897, University of Illinois, Urbana, IL, January 1978.

For more references, see the Publications section.

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]

People

Associated Faculty and Research Affiliates

Dr. Tomasz Arciszewski, Associate Professor, Urban Systems Engineering

Dr. Eric Bloedorn, Research Affiliate, MITRE

Dr. Hugo De Garis, Affiliated Research Scientist, Brain Builder Group, ATR, JAPAN

Dr. Kenneth De Jong, Professor of Computer Science

Dr. Zoran Duric, Assistant Professor of Computer Science

Dr. Ophir Frieder, Associate Professor of Computer Science

Dr. Ibrahim Imam, Research Scientist

Dr. Larry Kerschberg, Professor , Information and Software Systems Engineering

Dr. Yves Kodratoff, Research Director at the University of Paris-South and Adjunct Professor at the GMU Laboratory

Dr. Nikolai Lyashenko, Vice President of Empirical Inference Corporation, Schenectady, NY and a Visiting Professor since January, 1993

Dr. Marcus A. Maloof, Research Affiliate, Institute for the Study of Learning and Expertise

Dr. George Michaels, Associate Professor of Computational Biology and Deputy Director of the Institute for Computational Sciences and Informatics

Dr. Ryszard S. Michalski, PRC Chaired Professor of Computer Science and Systems Engineering, Director, Machine Learning and Inference Laboratory

Dr. David C. Rine, Professor of Computer Science and Software Systems Engineering

Dr. Vladimir N. Sazonov, Senior Member of the Technical Staff of Computer Sciences Corporation, and a member of the GMU Mathematical Science and Computer Science Departments

Dr. David Schum, Professor, Operations Research and Engineering

Dr. Tibor Vamos, Professor of the Technical University of Budapest and Distinguished Research Affiliate

Dr. Janusz Wnek, Affiliated Research Scientist

Visiting Professors

Dr. Adam Borkowski, Polish Academy of Sciences, Warsaw, Poland

Dr. Quin Chen, Dalian University of Technology, Dalian, People's Republic of China

Dr. Reuven Karni, Technion Institute, Israel

Dr. Daniele Nardi, University of Rome

Ph.D. Students

Nabil Alkharouf

Scott Fischthal

Ken Kaufman

Suet Chun Lee

Jim Mitchell

Larry Sklar

Tobias Ternstedt

Qi Zhang

Management

Dr. Ryszard S. Michalski, PRC Chaired Professor of Computer Science and Systems Engineering, Director of the Laboratory

Abhay Kasera, Research Manager

Ken Kaufman, Facilities Manager, EMERALD Manager

Qi Zhang, Assistant Editor of MLI Reports, Colloquium Coordinator, Software Manager, Webmaster

Jim Mitchell, Assistant Editor of MLI Reports

MLI Developed Software

Machine Learning, Inference and Discovery Systems developed in the Machine Learning and Inference Laboratory

The figure above is the opening screen of the EMERALD system developed in the Machine Learning and Inference Laboratory. The EMERALD system (Experimental Machine Example-based Reasoning and Learning Disciple) integrates five modules ("robots") each displaying a capability for some form of learning and discovery.

The Laboratory has developed a series of machine learning and Inference programs. Among these are: ABACUS, AQ15c, AQ16 (POSEIDON), AQ17-DCI, CLUSTER, EMERALD, INDUCE, and SPARC. These programs are described in more detail below.

- ABACUS 2 is a program for assisting a user for determining integrated quatitative and qualitative descriptions of data (a system for integrated "quatitative and qualitative discovery"). Given numerical and possibly also qualitative data describing some system or process, it generates mathematical equations characterizing the system or the process, and conditions funder which these equations apply. These equations can then be used for predicting the behavior of this system or process.

- AQ Family: All of the programs in the AQ family learn general decision rules from examples of decision classes. Here are standard features of the "base" version of the AQ program

The learned decision rules are optimized according to user-defined criteria or a default optimality criterion. The criteria refer to syntactic simplicity of the rules (measured by the number of rules, number of conditions in the rules, the simplicity of the conditions,or a combination of these factors), and/or the evaluation cost of the rules (the cost of measuring the attributes involved in the rules). Programs allow the user to generate different types of desriptions ("rulesets"), such as *discriminant* (that discriminate among given decision classes), or *characteristic* (that specify common features of the objects in the individual classes. The programs can also generate rulesets that have different relations among the rules -- intersecting (rules of different classes may logicaly intersect over areas not covering training examples), disjoint (rules or different classes are logically disjoint) or ordered (rules for each class are totally ordered and must be executed in the given order when applied to a given object). Learned rules are evaluated either by a strict match or by a flexible match. Individual versions of AQ programs have some additional features above the "base" version of the program.

AQ15c: The latest, most popular plain version of the AQ learning program (implemented in the ANSI C). This version is available for SunOS 4.1, MacOS 7.5 and DOS 6.x

AQ16 (POSEIDON): Plain AQ with mechanisms for optimizing rules by applying rule modification mechanisms. There are two mechanisms: TRUNC--that truncates insignificant rules (which corresponds to performing a form of ruleset specialization) or TRUNC/SG that modifies rules conditions and truncates insignificant rules (which corresponds to performing of both specialization and generalization of rules). Rules are evaluated either by a strict match or by a flexible match. These version is oriented toward learning concepts from noisy data or learning "flexible" concepts, that lack precise definition. The program applies som simple froms of "two-tiered" concept representation. A two-tiered represetnation consist of a base concept representation (BCR) that captures typical concept properties, and inferential concept representation that captures non-typical, variable, or exceptional concept properties. (See MLI papers on two-tiered concept representation). This version is available for SunOS 4.1

AQ17-DCI: AQ program with Data-driven constructive induction capabilities. These capabilities allow the program to autmatically modify the representation of the problem, e.g. adding or removing attributes or removing attribute-values. This version is available for SunOS 4.1.

AQ17-HCI: AQ program with Hypothesis-driven constructive induction capabilities. These capabilities allow the program to autmatically modify the representation of the problem, e.g. adding or removing attributes.

➲**CLUSTER** creates meaningful categories and classifications of given entities, and formulates descriptions of these created categories. Each class description is given in conjunctive form involving selected object attributes.CLUSTER has been applied to varied practical problems including classifying Spanish Folksongs, microcomputers, and reconstructing soybean disease categories.

➲**DIAV**: Diagrammatic Visualization of learning algorithms and discrete knowledge transmutations.

➲**EMERALD**: Integrated Learning Systems for Research and Education.

➲**INDUCE** learns structural descriptions of groups of objects, and determines important distinctions between the groups.

- **SPARC** predicts possible future objects or events by discovering rules characterizing the sequence of objects or events observed so far.

  Sparc/G: General purpose Sparc

  Sparc/E: Eleusis card playing version

To obtain a copy of any of these systems contact:

Qi Zhang MLI Laboratory Software Manager PHONE: USA 703 993 1716, FAX: USA 703 993 3729, Machine Learning and Inference Laboratory, George Mason University, 4400 University Dr. Fairfax, VA, 22030, USA

Recipients of MLI Software

ABB Corporate Research, Norway
Abo AKademi University, Finland
Austrian Research Institute for AI (ARIAI), Austria
Beckman Institute, USA
Bilkent University, Turkiye
Blue Cross / Blue Shield of Hawaii, USA
Bologne University, Italy
Brigham Young University, USA
British Columbia Cancer Agency, Canada
Brooklyn College of CUNY , USA
CAD, Philip Morris Research Center, USA
Carleton University, Canada
Deakin University, Australia
Centre dEstudis Avancats de Blanes, Spain
CSIRO, Australia
Dell Ingeneria Universita Di Modena, Italy
Department of Engergy, USA
Domaine Universitaire de St Jerome, France
Fakultaet fur Informatik, Germany
Elasis - Fiat, Italy
Foundation on Cardiac Surgery Development, Poland
European Computer-Industry, Germany
George Mason University, USA
GMD (The German Natl. Research Center for Comp. Science), Germany
Graz University of Technology, Austria
Hawaii Medical Service Association, USA

Information Technology Institute, Republic of Singapore
INPRO, Germany
Institute of Computer Science, Greece
Institute of Computer Science, Polish Academy of Sciences, Poland
Institute of Informatics, Bulgaria
Institute of Statistics, Poland
Instytut Informatyki, Poland
ISX Corp., USA
ITMI, France
KAIST, KOREA
Karlsruhe University, Germany
Kielce University of Technology, Poland
Landcare Research N2 Ltd., New Zealand
Landcare Research, New Zealand
Lawrence Berekely Laboratory., USA
Middle Tennessee State University, USA
MITRE Corp., USA
Mostra D'Oltremare, Italy
Monash University, Australia
National Insititute of Health, Japan
National Insititute of Health, USA
Norwegian Defense, Norway
Novell Inc., USA
Oklahoma State University, USA
Purdue University, USA
Sao Paulo University, Brazil
Sciences of Lisbon, Portugal
Silesian Technical University, Poland
Soong Sil University, Korea
Southwestern Bell Telephone, Co., USA
SRA, USA
Stanford University, USA
Syracuse University, USA
TASC, USA
Technion-Israel Institute of Technology, Israel
TECSIEL, Italy
The Norwegian Institute of Technology, Norway
The University of Manchaster, UK
Tsinghua University, P.R. China
United Technologies Research Center, USA
Universidade Nova de Lisboa, Portugal
Universidad de Granada, Spain
Universitat Politecnica, Spain
Universitat Rovira i Virgili, Spain
Universite de Valenciennes, France
Universite Paris VI, France
University of Aberdeen, UK
University of Arizona, USA

University of British Columbia, Canada
University of Buckingham, UK
University of Calgary, Canada
Univeristy of California at Berkley, USA
University of California at Irvine, USA
University of Dortmund, Germany
University of Florida, USA
University of Gliwice, Poland
University of Illinois-Urbana Champaign, USA
University of Karlsruhe, Germany
University of Keal, UK
University of Manchester, USA
University of Newcastle Upon Tyne, UK
University of North Carolina-Charlotte, USA
University of North Texas, USA
University of Otago, New Zealand
University of Ottawa, Canada
University of Portsmouth, United Kingdom
University of Siena, Italy
University of South Florida, USA
University of Stellenbosch, South Africa
University of Strathclyde, UK
University of Sydney, Australia
University of Toledo, USA
University of Waterloo, Canada
US NAVY, USA
Virginia Commonwelth Univ., USA
VP Technology at Eidetic Systems, USA
Yale University, USA

Previous Next

[MLI Home] [Mission] [People] [Research] [Software] [Publications] [Colloquia] [Events] [Guest Book]