MACHINE LEARNING RESEARCH IN THE ARTIFICIAL INTELLIGENCE LABORATORY AT ILLINOIS

by

R. S. Michalski

Machine Learning Research
in the Artificial Intelligence Laboratory at Illinois

Ryszard S. Michalski
University of Illinois at Urbana-Champaign

This note provides a brief account of major projects on Machine Learning done in
the Artificial Intelligence Laboratory at the Department of Computer Science of
the University of Illinois. Some of the projects are done in collaboration with
research groups from other organizations. A few projects mentioned here are
described in more detail as separate reports in the book. Research conducted by
Professor G. DeJong and his group at the Coordinated Science Laboratory is
reported separately. References contain publications describing details on projects
described below and closely related work. Due to the space limitations these
references include only a selection of papers from those written in the last year.

Our research has several objectives: to develop unifying principles of learning, to
investigate relationships in learning and inference methods employed by people
and machines, to develop experimental learning systems, and to test the methods
developed on selected practical problems in various domains. Within the spectrum
of scientific approaches to machine learning one can distinguish two extremes:
general-purpose knowledge-free, and *domain-specific knowledge-intensive*. The
approach taken by our research group represents a middle range between the two;
it can be called *domain-generic knowledge-modular*. Within the large space of
different learning processes one searches for *generic domains*, which represent
classes of conceptually related tasks. A research goal is then to develop methods
and programs for chosen generic domains. As examples of generic domains consider
learning structure-based descriptions from examples, conjunctive conceptual
clustering or predicting sequence of events characterized by attributes. Solving
any specific task from a generic domain requires that the program for this
domain be supplied with a module of knowledge relevant to the given task. For
different tasks different modules are developed. The next research stage is to
develop an integrated learning system that includes various learning subsystems,
each of which serving as an expert on problems within a given generic domain.
Several projects conducted by our group relate to selected generic domains. Below
is a brief description of major current projects.

Inference-based Theory of Learning

A learning process can be viewed as a transformation of information provided by
a source of information (a teacher or environment) into a more efficient and/or
more general form in which it is stored for future use. This transformation
involves the background knowledge of the learner (that includes goals of learning,
concepts learned from the past experience, hypothesis evaluation criteria, etc.), and

uses various types of inference. The type of inference employed in any given act of learning depends on the source of information and the goal of learning.

The inference type defines the *learning strategy*. For example, when learning a concept from examples, explaining the behavior of a system, or discovering an equation that binds specific numeric values, one employs inductive inference. Therefore these are forms of *inductive learning*. The inductive learning strategy is complex and inherently fallible, but it frees the teacher from spoon-feeding the student. The other extreme is rote learning which requires little or no inference from the student, but requires that the teacher provide all the information in a form directly absorbable by the student. This project investigates the trade-offs between various learning strategies, and attempts to develop a unifying theory of learning.

Inductive Learning of Concepts, Rules and Procedures

Following earlier work on the AQ11 and INDUCE concept learning programs, this project is concerned with the development of advanced multi-purpose learning systems for various classes of problems, and their application to different practical problems. The knowledge representation formalism is based on various forms of *variable-valued logic* and the *annotated predicate calculus* (a typed predicate calculus with additional operators plus annotations). Recently studied topics include methods for incremental learning and refinement of rule bases, automated generation of relevant attributes and terms, learning from incomplete and/or noisy examples, learning rules with exceptions, and inferring causal models from facts. Selected results are described in more detail in separate reports in this book: Igor Mozetic describes how one of our inductive learning programs refined and simplified by two orders of magnitude a knowledge base for diagnosing cardiac arrhythmias, and Kaihu Chen describes the work on learning procedures from examples. Results on learning with exceptions and from noisy data are described by Becker. Other results are described in reports listed in references.

Probabilistic Learning

This project, directed by Professor Larry Rendell, investigates probabilistic learning systems capable of forming concepts and optimizing problem solving strategies using a form of conceptual clustering. Such systems can automatically split a problem solving state space into equivalence classes, which greatly improves the efficiency of a search process. They can handle noisy data and improve incrementally. Current research includes development of general principles and models for efficient induction, and the investigation of the role of probabilistic methods in learning. More details on this work are in a separate report by Rendell.

Conceptual Clustering and Automated Design of Classifications

This project, co-directed by Professor Robert Stepp and the author is concerned with problems of dividing a collection of observations into separate classes via *conceptual clustering*. Classes are created not because objects are similar according to some predefined numerical measure of similarity, but because they relate to each other in such a way that together they constitute (or can be explained by) a simple concept. We have developed two conceptual clustering programs, CLUSTER/2 for building classifications using attribute-based descriptions, and CLUSTER/S for building classifications using structure-based descriptions. Current research involves problems of building goal-oriented classifications and automated generation of relevant descriptors. A separate report by Stepp gives details on some of the recent work.

Learning to Program by Analogy

Learning by analogy can be viewed as inductive and deductive learning combined. Through induction the system finds a common substructure in the entities being compared, and then by deduction uses this substructure for analogical mapping. This project, led by Professor Dershowitz, uses analogy as a tool in automatic programming. To determine a program for a new problem, the specification of the problem is compared with the specification of an existing program. Through analogy, the existing program is modified to meet the new specification.

Qualitative Prediction

This research aims at developing methods for discovering rules and relationships characterizing discrete processes and then using these rules and relationships to predict the continuation of the process. As examples of such processes consider a sequence of states of a patient under treatment, or stages in the development of a crop disease in a field. This research started with a successful development of a system SPARC/E (*sequential pattern recognition*) for discovering rules in the card game Eleusis. This game provides a microcosm modeling the process of scientific discovery. Recent work involves the development of a program SPARC/G for solving a general class of non-deterministic prediction problems, and an application of the program to various domains, such as robotics and biology. More details on this project are provided in the references. Some aspects this of research are described by Heedong Ko and Kaihu Chen in separate reports in the book.

Quantitative Discovery Systems

In many areas of science, there occurs the problem of discovering a numerical law explaining or characterizing a given set of observations. Attempts to automate some parts of this process were made in pioneering work by Simon, Langley and collaborators on the development of various versions of the BACON system. Our research strives to extend the work on BACON and combine it with our concept learning systems. The system ABACUS, recently developed by Falkenhainer in his M.S. thesis, can not only to discover qualitative laws that characterize given sets of observations, but also formulate the areas of applicability of these laws. When different subsets of data are governed by different laws, the program fomulates the laws relevant for each subset, and provides conceptual descriptions defining preconditions for application of each law. A new project, done in collaboration with Nayel El-Shafei from MIT, attempts to develop a system integrating methods for quantitative and qualitative discovery.

Knowledge Acquisition for Expert Systems through Inductive Learning

In the context of research on the meta-expert system ADVISE, directed by Baskin, Michalski and Stepp, we are investigating methods for building expert systems with inductive learning abilities. ADVISE is an integrated set of tools for designing and experimenting with expert systems with multiple control strategies, different methods for propagating uncertainties, several knowledge representations, and learning capabilities. A module of ADVISE, called QUIN (for *qu*ery and *inf*erence), provides a set of learning and inference operators that apply to data in relational table format. These operators can generate decision rules from examples, formulate classifications, select relevant attributes or determine representative learning events.

Variable-Precision Logic

The concept of variable-precision logic was introduced by the author and Patrick Winston from MIT's Artificial Intelligence Laboratory. This logic is concerned with problems of reasoning that represents different trade-offs between certainty of conclusions, their specificity, and computational resources involved in deriving these conclusions. A paper finished recently on this subject uses *censored production rules* as a representational and computational mechanism for handling some of these trade-offs and learning with exceptions. Some aspects of VPL were recently implemented by Becker in his M.S. thesis work.

Human Plausible Reasoning and Concept Formation

This research involves two projects, one in collaboration with Alan Collins from Bolt, Beranek and Newman, Inc., and the other with Doug Medin from the Department of Psychology at the University of Illinois. The goal of this research is to develop a theory explaining how humans conduct plausible reasoning, how they learn concepts, and how they generate classifications. A paper describing the core of a *perturbation theory* of human plausible reasoning is in its final stages. Recent research involved experiments comparing human and machine performance on selected problems of concept formation and classification development. A paper describing the results is in preparation.

REFERENCES

A Selection of Papers Written in 1985

J. BECKER, "AQ-PROLOG: A Prolog Implementation of an Attribute–Based Inductive Learning System," ISG 85-1, UIUCDCS-F-85-930, Department of Computer Science, University of Illinois, Urbana, January 1985.

P. HOFFMAN, "MEL – A Learning Program that Improves by Experience in Playing the Game of MILL," ISG 85-2, UIUCDCS-F-85-931, Department of Computer Science, University of Illinois, Urbana, January 1985.

J. BECKER, "Topics in Incremental Learning of Discriminant Descriptions," ISG 85-5, UIUCDCS-F-85-935, Department of Computer Science, University of Illinois, Urbana, February 1985.

T. DIETTERICH and R. S. MICHALSKI, "Learning to Predict Sequences," ISG 85-9, UIUCDCS-F-85-938, Department of Computer Science, University of Illinois, Urbana, February 1985, also in *Artificial Intelligence Journal*, Vol 25, No 2, pp 187-232, February 1985.

R. S. MICHALSKI, H. KO and K. CHEN, "SPARC/E(V.2), An Eleusis Rule Generator and Game Player," ISG 85-11, UIUCDCS-F-85-941, Department of Computer Science, University of Illinois, Urbana, February 1985.

I. MOZETIC, "Compression of the ECG Knowledge–base Using the AQ Inductive Learning Algorithm," ISG 85-13, UIUCDCS-F-85-943, Department of Computer Science, University of Illinois, Urbana, March 1985.

R. E. REINKE and R. S. MICHALSKI, "Incremental Learning of Decision Rules: A Method and Experimentals Results," presented at the *Machine Intelligence Workshop 11*, Ross Priory, University of Strathclyde, March 1985.

R. S. MICHALSKI, "Knowledge Repair Mechanisms: Evolution vs Revolution," ISG 85-14, UIUCDCS-F-85-946, Department of Computer Science, University of Illionis, Urbana, July 1985.

L. RENDELL, "Genetic Plans and the Probabilistic Learning System: Synthesis and Results," *Proceedings of the International Conference on Genetic Algorithms and their Applications*, Carnegie–Mellon University, July 1985.

J. BECKER, "Inductive Learning of Decision Rules with Exceptions: Methodology and Experimentation," ISG 85-14, UIUCDCS-F-85-945, Master of Science Thesis, Department of Computer Science, University of Illinois, Urbana, August 1985.

L. RENDELL, "Substantial Constructive Induction Using Layered Information Compression: Tractable Feature Formation in Search," *Proceedings of the Ninth International Joint Conference on Artificial Intellignece*, University of California at Los Angeles, August 1985.

B. FALKENHAINER, "Quantitative Empirical Learning: An Analysis and Methodology," ISG 85-16, UIUCDCS-F-947, Master of Science Thesis, Department of Computer Science, University of Illinois, Urbana, August 1985.

R. S. MICHALSKI and P. WINSTON, "Variable Precision Logic," Artificial Intelligence Memo #857, Massachusetts Institute of Technology, Cambridge, MA, 1985.

R. STEPP and R. S. MICHALSKI, "Conceptual Clustering of Structured Objects: A Goal-Oriented Approach," accepted for *AI Journal*, Winter 1986 (forthcoming).

R. S. MICHALSKI, "Understanding the Nature of Learning: Issues and Research Directions," chapter in *Machine Learning: An Artificial Intelligence Approach*, Vol. II, Morgan–Kaufman Publishers, 1986 (forthcoming).

N. DERSHOWITZ, "Programming by Analogy," chapter in *Machine Learning: An Artificial Intelligence Approach*, Vol. II, Morgan–Kaufman Publishers, 1986 (forthcoming).

D. L. MEDIN, Wm D. WATTENMAKER and R. S. MICHALSKI, "Constraints in Inductive Learning: An Experimental Study Comparing Human and Machine Performance," in preparation, 1985.

A. COLLINS and R. S. MICHALSKI, "Logic of Plausible Reasoning," in preparation, 1985.