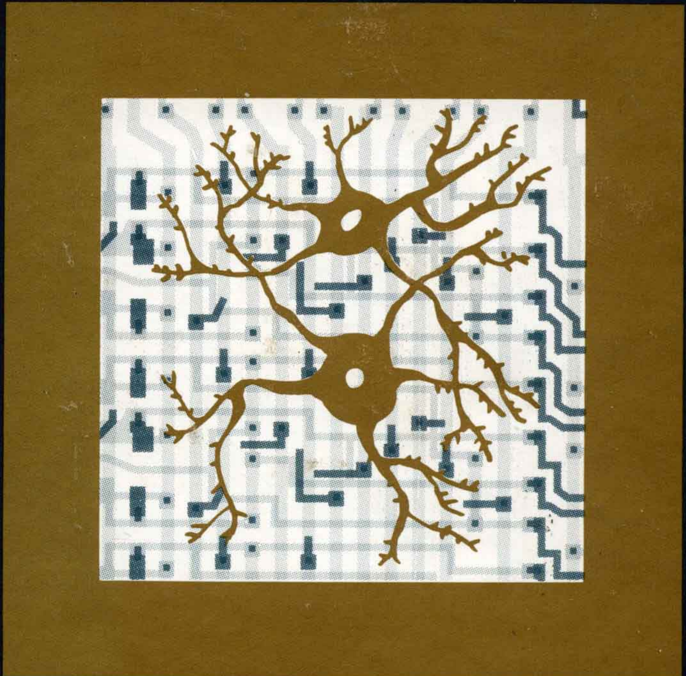
MACHINE LEARNING

An Artificial Intelligence Approach

Volume III



Yves Kodratoff
Ryszard Michalski

Editor and Publisher: *Michael B. Morgan*
Production Manager: *Shirley Jowell*
Project Management: *Jennifer Ballentine*
Cover Design: *Andrea Hendricks*
Electronic Composition: *Ocean View Technical Publications*

The publisher gratefully acknowledges permission to reproduce the following figures: 20–4, 20–5, 20–6 reprinted with permission from "Learning Internal Representations by Error Propagation," *Nature*, Vol. 323, pp. 533–536. Copyright © 1986, MacMillan Magazines, Ltd.; Figure 23–3, adapted from *The Little Prince*, by Antoine de Saint-Exupéry, © 1943 and renewed 1971, by Harcourt, Brace, Jovanovich, Inc., reprinted with permission of the publisher.

Library of Congress Cataloging-in-Publication Data
(Revised for vol. 3)

Machine learning.

Vol. [1] previously published: Palo Alto, Calif. : Tioga Pub. Co., c1983.
Vol. 3- edited by Yves Kodratoff and Ryszard Michalski.
Includes bibliographies and indexes.
1. Machine learning. 2. Artificial intelligence.
I. Anderson, John R. (John Robert), 1947–
II. Michalski, Ryszard Stanislaw, 1937–
III. Carbonell, Jaime G. (Jaime Guillermo)
IV. Mitchell, Tom M. (Tom Mitchell), 1951–
Q325.M32 1983b 006.3'1 86-2953
ISBN 0-934613-09-5 (v. 1)
Vol. 3 ISBN 1-55860-119-8

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

MORGAN KAUFMANN PUBLISHERS, INC.

Editorial Office
2929 Campus Drive, Suite 260
San Mateo, CA 94403
415-578-9911

Order From
P.O. Box 50490
Palo Alto, CA 94303-9953
415-965-4081

PREFACE

As the field of machine learning enjoys unprecedented growth and attracts many new researchers, there is a need for regular summaries and comprehensive reviews of its progress. This volume is a sequel to the previous volumes of the same title: Volume I appeared in 1983, Volume II in 1986. Volume III presents a sample of machine learning research representative of the period between 1986 and 1989.

One noteworthy characteristic of that period is that a much larger portion of research has been done outside of the United States, particularly in Europe. To reflect his, Volume III contains a significant number of non-U.S.A contributions. In addition, this volume covers topics not covered at all or covered only sparsely by the previous volumes, such as connectionist learning methods, genetic algorithms, and computational learning theory.

To provide a comprehensive representation of research, this volume has drawn on several sources. Most of the chapters are directly invited contributions by leading researchers in the field. Several chapters are updated and extended versions of invited presentations at the International Meeting on Advances in Learning (IMAL) held in Les Arcs, France in July 1986. These chapters are accompanied by commentaries prepared by the discussants at the meeting. Finally, few chapters are based on papers selected from among those presented at the 4th and 5th International Machine Learning conferences, held at the University of California at Irvine in June 1987 and the University of Michigan at Ann Arbor in June 1988, respectively.

The bibliography at the end of the book provides a comprehensive guide to these and related publications. It contains over 1000 entries and refers to publications in all major ML subareas for the period 1985–1989. All the entries are indexed, using a classification of ML publications into 17 categories.

For more complete coverage of the progress of the field, the reader is referred to relevant journals, in particular, *Machine Learning*, *Artificial Intelligence*, and *AI Magazine*, and to the proceedings of various conferences. Among the most relevant conferences are international machine learning conferences [T87, T88, T89], the meetings of the American Association for Artificial Intelligence [AAAI T86, T87 and T88], workshops on computational learning theory [COLT T88 and T89], the workshop on explanation-based learning [T88], international conferences on genetic algorithms [T87 and T89], conferences on neural nets, the European Working Sessions on Learning [EWSL T87, T88 and T89], the European congresses on artificial intelligence [ECAI T86 and T88], the International Joint Conferences on Artificial Intelligence [IJCAI T87 and T89], and International Workshop on Tools for Artificial Intelligence (1989).

It is the editors pleasant duty to thank all those who helped in the preparation of this book. Our deep gratitude goes to all the contributors for their efforts to write the chapters in a highly comprehensive and easy-to-read manner. We are very grateful to the reviewers, whose help was indispensible. We wish to thank Shirley Jowell, Production Manager for Morgan Kaufmann, for her contribution to this book.

Special thanks go to DIGITAL-EUROPE and the London office of the U.S. Army. These organizations sponsored IMAL, which gave the first impetus this volume. The editors also acknowledge the help and technical support extended to them by the faculty, staff, and research assistants of the Center for Artificial Intelligence and the Department of Computer Science at George Mason University and by the French National Research Center (CNRS).

Ryszard S. Michalski
George Mason University

Yves Kodratoff
French National Research Center (CNRS)
and George Mason University

CONTENTS

PART ONE

GENERAL ISSUES

1

RESEARCH IN MACHINE LEARNING:

Recent Progress, Classification of Methods, and Future Directions

Ryszard S. Michalski
(George Mason University)

Yves Kodratoff
(CNRS, Université de Paris-Sud and George Mason University)

Abstract

The last few years have produced a remarkable expansion of research in machine learning. The field has gained an unprecedented popularity, several new areas have developed, and some previously established areas have gained new momentum. While symbolic methods, both empirical and knowledge intensive (in particular, inductive concept learning and explanation-based methods), continued to be exceedingly active (see Parts Two and Three of the book, respectively), subsymbolic approaches, especially neural networks, have experienced tremendous growth (Part Five). Unlike past efforts that concentrated on single learning strategies, the new trends have been to integrate different strategies and to develop cognitive learning architectures (Part Four). There has been an increasing interest in experimental comparisons of various methods, and in theoretical analyses of learning algorithms. Researchers have been sharing the same data sets and have applied their techniques to the same problems in order to understand the relative merits of different methods. Theoretical investigations have brought new insights into the complexity of learning processes (Part Six).

This chapter gives a brief account of the recent progress and prospective research directions in the field, attempts to clarify some basic concepts, proposes a

multicriteria classification of learning methods, and concludes with a brief description of each chapter.

1.1 INTRODUCTION

One of the most striking differences between how people and computers work is that humans, while performing any kind of activity, usually simultaneously expend efforts to improve the way they perform it. This is to say that human performance of any task is inseparably intertwined with a learning process, while current computers are typically only executors of procedures supplied to them. They may execute very efficiently, but they do not self-improve with experience.

Research in machine learning has been concerned with building computer programs able to construct new knowledge or to improve already possessed knowledge by using input information. So far, this input information (examples, facts, descriptions, etc.) has been typically typed in by a human instructor. Future machine learning programs will undoubtedly be able to receive inputs directly from the environment through a variety of sensory devices.

The great appeal of this field to its practitioners is that machine learning offers an immense diversity of research tasks and testing grounds. This diversity is due to the fact that learning can accompany any kind of problem solving or process, and thus it can be studied in many different contexts, such as decision making, classification, sensory signal recognition, problem solving, task execution, control, or planning.

This continual appeal of the field has been enhanced recently by the fact that progress in machine learning has become central to the development of the field of artificial intelligence (AI) as a whole and affects almost all of its subareas. In particular, the work in machine learning has importance for expert systems development, problem solving, computer vision, speech understanding, autonomous robotics, conceptual analysis of databases, and intelligent tutoring systems. Consequently, the development of powerful learning systems may ultimately open an unprecedented range of new applications (e.g., [Michalski, 1986]).

Research on building learning programs goes back almost to the beginning of the computer area. After the first significant burst of research on perceptrons and self-organizing systems in the 1950s and the first few years of the 1960s, the field has been growing slowly but steadily. Some early successes include, for example, the Samuel's checkers program [Samuel, 1959], Winston's program for learning structural descriptions [Winston, 1970; 1975], the Meta-DENDRAL program for heuristic rule formation [Buchanan, Feigenbaum, and Sridharan, 1972], the AM and EURISKO discovery programs [Lenat, 1977; 1983], AQ11 for diagnostic rule learning [Michalski and Chilausky, 1980], LEX for learning symbolic integration [Mitchell, Utgoff, and Banerji, 1983], and CLUSTER for conceptual clustering [Michalski and Stepp, 1983].

These successes and the ever-present challenge to build powerful learning systems, have exerted strong pressure to expand the activities in this field. The first machine learning workshop was held at Carnegie Mellon University (CMU) in 1980. This workshop and the publication in 1983 of the first volume of *Machine Learning* [Michalski, Carbonell, and Mitchell, 1983] marked a breaking point. These two events gave the field a clear identity and a sense of direction, which in turn stimulated the rapid growth that has continued unabated since then.

There have been subsequent workshops and conferences: at the University of Illinois at Urbana-Champaign in 1983, at Rutgers University in 1985, at the University of California at Irvine in 1987, at the University of Michigan in 1988, and at Cornell University in 1989. In 1986, *Machine Learning, Volume II* appeared [Michalski, Carbonell, and Mitchell, 1986]. In response to the growing need for an adequate forum for presenting research progress, *Machine Learning* journal was established in 1986.

There have also been numerous workshops and meetings on special topics, such as computational learning theory (COLT 88 and 89), explanation-based learning (AAAI workshop at Stanford University, 1988), connectionist models of learning (e.g., summer schools at CMU in 1986 and 1988 and a number of international conferences), and knowledge discovery in databases (IJCAI-89 workshop in Detroit).

In parallel, there has been a rapid increase of interest in machine learning in Europe, as signified by many activities, meetings, and conferences. Among the most noteworthy were the European Working Sessions on Learning (Orsay 1986, Bled 1987, Glasgow 1988, and Montpellier 1989), the International Meeting on Advances in Learning in Les Arcs in 1986, the workshop on Knowledge Representation and Organization in Machine Learning (KROML 1987), Workshop on Machine Learning, Metareasoning and Logic (Sesimbra 1988), and Summer Schools in Machine Learning (Les Arcs 1988 and Urbino 1989), International Schools for the Synthesis of Knowledge (ISSEK 1987 and 1989). To reflect these activities, this volume includes a significant number of non-U.S. contributions.

1.2 RECENT DEVELOPMENTS

The last few years have witnessed both a continuation of the major traditional research approaches and a rapid increase of interest in several new methodologies. The most active research area in recent years has continued to be symbolic empirical learning (SEL). This area is concerned with creating and/or modifying general symbolic descriptions, whose structure is unknown *a priori*. This type of learning can be contrasted with, e.g., learning weights assigned to connections in a given neural net, or coefficients of equations in a predefined form. The descriptions are created on the basis of examples or specific facts. The word "empirical" signifies the fact that the learning process does not require much prior knowledge of the learner (if the process relies on a large amount of explicitly stated prior knowledge, then we have knowledge-intensive symbolic learning).

An important criterion underlying SEL methods is that knowledge created by a learning program is supposed to be easy for humans to interpret and comprehend. This means that there is a concern to make knowledge representations simple in terms of the structures used and the number of operators involved. It also means that the concepts employed in the descriptions should directly correspond to those used by human experts. This criterion is sometimes called the *comprehensibility principle* [Michalski, 1983]. Typical knowledge structures employed in the SEL systems include commonly used symbolic representations, such as logic-based descriptions, rules, decision trees, semantic networks, equations, frames, and grammars. Due to the comprehensibily criterion, the SEL systems can be particularly useful in applications in which people need to fully comprehend the results of learning—for example, in technical, medical, or agricultural diagnosis; decision making, planning, economic or political analysis; discovery of knowledge in databases, prediction, etc.

The most common topic in SEL is developing concept descriptions from concept examples. The machine learning bibliography (MLB; the last chapter of this book), which contains 1050 entries covering the period 1985—89, lists about 190 publications on this topic. Other major topics in SEL include qualitative discovery, conceptual clustering, and empirical sequence prediction. The MLB lists another 130 publications on these topics; thus together, there are about 320 papers listed in the MLB on symbolic empirical learning.

As mentioned above, empirical methods typically use relatively little background knowledge, by which we mean the relevant domain-dependent knowledge, such as facts or rules characterizing the application domain, and domain-independent knowledge, such as general definitional knowledge, commonsense knowledge, and explicit rules of inference, which the learner can bring to bear in the process of learning. In SEL systems, the background knowledge may include merely information about the value sets and types of attributes or terms (descriptors) used, the constraints on the attributes, preference criteria or biases for judging candidate solutions, etc. The domain-dependent information can be introduced to a program when it is applied to a particular problem, and therefore it is relatively easy to develop a general-purpose empirical learning program. The AQ family of rule learning programs (e.g., [Michalski, 1973; and Chapter 3, this volume], and the ID3-type decision tree learning programs [Quinlan, 1979; and Chapter 5, this volume] are examples of such general-purpose SEL systems. The AQ programs generate rules by manipulating knowledge structures according to rules of inference and knowledge transformation. The ID3-type systems create a decision tree by a recursive selection of attributes from a given set. The attribute selection is based on statistical considerations (e.g., the *minimum entropy rule*), rather than on explicit rules of inference.

The primary inference type used by SEL methods is *empirical induction*. This form of induction (as other forms, such as constructive induction and abduction, see the next section) is a falsity-preserving, rather than truth-preserving inference. Therefore, the results of SEL methods are generally hypotheses, which need to be validated by further experiments. This is often viewed as an important weakness of the empirical methods. It reflects the intrinsic uncertainty of any process of creating new knowledge about the world, and therefore is unavoidable in principle. The only way to circumvent it is to restrict the learning process either to copying existing knowledge or to strict deductions from knowledge that has been tested and assumed to be true. Such an *a priori* knowledge has to be encoded into the system before any learning can occur (see analytic methods in the next section).

Another weakness of the SEL methods is that the knowledge learned by them represents relations expressed merely in terms of attributes or concepts either directly specified in the input data, or closely related to them (an empirical program may include procedures for transforming the initial description space). Because the methods rely primarily on the input information, rather than on background knowledge, they cannot discover complex relationships or causal dependencies, which require high-level terms or concepts, not provided by the input.

The fact that symbolic empirical methods do not use/require much background knowledge is appealing to many researchers. Examples or observations are often easily available from existing databases or can be measured by sensors. There is no need for debugging and handcrafting large amounts of knowledge into the system. Consequently, empirical learning systems are readily applicable to a wide spectrum of practical problems. In addition, because the results are usually easy to interpret (in contrast to subsymbolic systems; see below), the methods are particularly attractive in application areas where understandability of the results is an important factor. A selection of research in symbolic empirical learning is presented in Part Two, Chapters 3 through 9.

In recent years, there have been various efforts to extend the capabilities of conventional SEL systems. A considerable amount of work has been done on learning concepts from imperfect inputs, e.g., learning from examples with noise (see Chapter 5). Related efforts have been concerned with learning concepts that lack precise definition and/or are context dependent (see Chapter 3).

Another major extension of empirical methods addresses the problem of employing more background knowledge in the process of inductively creating concept descriptions. The motivation is that people, due to their prior knowledge, can often create plausible inductive hypotheses from a few, or just one, instances. For example, if one sees a single window of a particular style in a tall skyscraper, then one does not need to look at other windows to hypothesize that all the windows in that building are of this style. The reason is that we know that windows in a building are typically made in the same style. As another example, consider a person who deceptively misinforms others about something really important. Usually, it would not take more than one such instance to cause others to never trust that person in the future. Again, this is because of a common belief, that if a person lied once, it is likely that this person may continue such behavior, and trusting such a person would carry a very high risk. Thus, by involving prior knowledge, one can create plausible inductive hypotheses from very little input information, contrary to some beliefs about inductive learning.

Also, in many applications, it is important to discover relationships that go beyond associations between inputs and outputs. In such applications, it is important to search for relationships that involve higher level concepts than those defined in the inputs, to generate and employ abstract attributes and relations, and/or to determine causal explanations of the observations. Any process of theory formation requires much background knowledge in addition to observational data.

To this end, some researchers started to work on *constructive induction*, which is a term for characterizing inductive processes that engage significant amount of background knowledge ([Michalski, 1983]; see also [Muggleton and Buntine, 1988; Rouveirol and Puget, 1989]). Such background knowledge may be in the form of expert-given domain knowledge rules, logical implications and equivalences, abstract concept definitions, heuristic procedures for generating new concepts, goal-oriented criteria for evaluating the importance of created knowledge, and others. This knowledge may be used in the conventional, deductive manner, thus, constructive induction typically includes a large component of deductive inference. Equipped with appropriate background knowledge, constructive inductive systems can change the representation of the problem or invent new attributes or concepts. As described in ([Michalski, 1990]; see also Section 1.3), constructive induction involves "reverse reasoning" or "tracing backward" of certain implicative rules, which are either *domain-independent* (tautological implications) or *domain-dependent* (representing domain knowledge). When domain-independent rules are primarily involved (specifically, the falsity-preserving generalization rules), then constructive induction reduces to empirical induction. When certain domain-dependent implicative relationships are "traced backward," then such induction becomes abduction (see next section). There are over 50 publications listed in the MLB in the area of constructive induction, abduction, and representation change.

Other classes of empirical learning systems include parametric and heterogeneous systems. In parametric systems, the learning process involves a modification of certain parameters or weights associated with predefined structures (networks, equations, production rules, etc.). Learning in heterogeneous systems involves both a direct modification of knowledge structures and a modification of the parameters associated with these structures. The most popular and important representative of parametric systems are neural nets and connectionist systems. In those systems, the learning process typically involves a modification of the strength of the connections between units in a statically or dynamically defined network. All units often perform the same general transformation, and therefore it is easy to build very large networks of that kind. It is important to note that a modification of the strengths of connections in a neural network can lead to a change in the knowledge structure. This structural change, however, is indirect and implicit, rather than direct and explicit, like in symbolic systems. The most explored subsets of heterogeneous systems are genetic algorithms and classifier systems. In those systems, the modification of the structures is done either by random changes (mutation), or semirandom changes (crossover), rather than by explicit rules of inference, like in typical symbolic systems. The weights assigned to individual production rules represent their importance or effectiveness in performing the assigned task.

Recent years have witnessed a remarkable renaissance of research on learning in neural networks. There is rapidly growing interest in exploring their properties and potential applications. Since these systems employ a general and uniform knowledge representation, and typically use little background knowledge, it is easy to implement them and apply them experimentally to a wide spectrum of problems. As they require very little guidance from a teacher, they are very appealing to many researchers.

A major limitation of neural networks and genetic algorithms is the difficulty of introducing large amounts of domain specific knowledge to them, and explicitly exploiting that knowledge or any feedback information in the learning process. To explain the latter, suppose that a neural network or genetic algorithm gets feedback that some example was incorrectly classified. To take care of the mistake, the system modifies its knowledge representation by stepwise corrections, rather than by an explicit analysis of the reasons for the mistake. This seems to explain why such systems tend to exhibit relatively slow rates of learning. Another weakness is the lack of transparency of the results of learning. The knowledge acquired by neural networks or conventional genetic algorithms is not in the form that people can easily understand. The comprehensibility principle has not been viewed as a major issue in implementing such systems. For that reason, they are sometimes called *subsymbolic learning systems.*

The lack of transparency is not necessarily always a problem. There are many application domains that do not require that the knowledge learned be easy to understand. For example, it is not important to understand the control algorithm of a robot, as long as it can move its hand to the given destination and within a defined space. This weakness is only a problem in areas where people need to understand the knowledge underlying the system's behavior; e.g., in diagnostic, advisory, or planning systems. It can be pointed out that a genetic algorithm could potentially be applied with a high-level symbolic knowledge representation (such a method would