BIBLIOTHEQUE DU CERIST

SYMBOLIC COMPUTATION

# Machine Learning

## An Artificial Intelligence Approach

Edited by
Ryszard S. Michalski Jaime G. Carbonell
Tom M. Mitchell

Springer-Verlag
Berlin Heidelberg New York Tokyo

C 1022

# Machine Learning

## An Artificial Intelligence Approach

Edited by

R. S. Michalski J. G. Carbonell T. M. Mitchell

With Contributions by

J. Anderson R. Banerji G. Bradshaw J. Carbonell
T. Dietterich N. Haas F. Hayes-Roth G. Hendrix
P. Langley D. Lenat R. Michalski T. Mitchell J. Mostow
B. Nudel M. Rychener R. Quinlan H. Simon
D. Sleeman R. Stepp P. Utgoff

Centre de Recherche sur l'Information Scientifique et Technique
Bibliothèque

Springer-Verlag
Berlin Heidelberg NewYork Tokyo 1984

BIBLIOTHEQUE DU CERIST

# PREFACE

BIBLIOTHEQUE DU CERIST

The ability to learn is one of the most fundamental attributes of intelligent behavior. Consequently, progress in the theory and computer modeling of learning processes is of great significance to fields concerned with understanding intelligence. Such fields include cognitive science, artificial intelligence, information science, pattern recognition, psychology, education, epistemology, philosophy, and related disciplines.

The recent observance of the silver anniversary of artificial intelligence has been heralded by a surge of interest in machine learning—both in building models of human learning and in understanding how machines might be endowed with the ability to learn. This renewed interest has spawned many new research projects and resulted in an increase in related scientific activities. In the summer of 1980, the First Machine Learning Workshop was held at Carnegie-Mellon University in Pittsburgh. In the same year, three consecutive issues of the *International Journal of Policy Analysis and Information Systems* were specially devoted to machine learning (No. 2, 3 and 4, 1980). In the spring of 1981, a special issue of the *SIGART Newsletter* No. 76 reviewed current research projects in the field.

This book contains tutorial overviews and research papers representative of contemporary trends in the area of machine learning as viewed from an artificial intelligence perspective. As the first available text on this subject, it is intended to fulfill several needs. For researchers in artificial intelligence, computer science, and cognitive psychology, it provides an easily accessible collection of state-of-the-art papers presenting current results, which will hopefully spur further research. For students in artificial intelligence and related disciplines, this volume may serve as a supplementary textbook for a course in artificial intelligence, or as a primary text for a specialized course devoted to machine learning. Finally, due to the potential impact of machine learning on a variety of disciplines, this book may be of interest to a diverse range of readers, including computer scientists, robotics experts, knowledge engineers, educators, philosophers, data analysts, psychologists and electronic engineers.

The major contemporary research directions in machine learning covered in this book include: learning from examples, modeling human learning strategies, knowledge acquisition for expert systems, learning heuristics, learning from instruction, learning by analogy, discovery systems, and conceptual data analysis. A glossary of selected terminology and an extensive up-to-date bibliography are provided to facilitate instruction and suggest further reading.

*—Ryszard S. Michalski, Jaime G. Carbonell, Tom M. Mitchell*

## ACKNOWLEDGMENTS

The editors wish to express a deep gratitude to all the authors whose efforts made this book possible, and to the reviewers whose ideas and criticism were indispensable in refereeing and improving the contributions. They are most grateful to the Office of Naval Research for supporting the First Machine Learning Workshop, where the idea of publishing this book originated.

The following people provided invaluable assistance in preparing the book: Margorie Ast, Trina and Nathaniel Borenstein, Kitty Fischer, Edward Frank, Jo Ann Gabinelli, Joan Mitchell, June Wingler and Jan Zubkoff. Their hard work and perseverance are gratefully acknowledged. The editors also acknowledge the support and access to technical facilities provided by the Computer Science Departments of Carnegie-Mellon University, the University of Illinois at Urbana-Champaign, and Rutgers University.

BIBLIOTHEQUE DU CERIST

# CONTENTS

BIBLIOTHEQUE DU CERIST


| | | |
|---|---|---|
| **Preface** | | **v** |
| **PART ONE** | **GENERAL ISSUES IN MACHINE LEARNING** | **1** |
| **Chapter 1** | **An Overview of Machine Learning**<br>*Jaime G. Carbonell, Ryszard S. Michalski, and Tom M. Mitchell* | **3** |
| 1.1 | Introduction | 3 |
| 1.2 | The Objectives of Machine Learning | 3 |
| 1.3 | A Taxonomy of Machine Learning Research | 7 |
| 1.4 | An Historical Sketch of Machine Learning | 14 |
| 1.5 | A Brief Reader's Guide | 16 |
| **Chapter 2** | **Why Should Machines Learn?**<br>*Herbert A. Simon* | **25** |
| 2.1 | Introduction | 25 |
| 2.2 | Human Learning and Machine Learning | 25 |
| 2.3 | What is Learning? | 28 |
| 2.4 | Some Learning Programs | 30 |
| 2.5 | Growth of Knowledge in Large Systems | 32 |
| 2.6 | A Role for Learning | 34 |
| 2.7 | Concluding Remarks | 35 |
| **PART TWO** | **LEARNING FROM EXAMPLES** | **39** |
| **Chapter 3** | **A Comparative Review of Selected Methods for Learning from Examples**<br>*Thomas G. Dietterich and Ryszard S. Michalski* | **41** |
| 3.1 | Introduction | 41 |
| 3.2 | Comparative Review of Selected Methods | 49 |

| | | |
|---|---|---|
| 3.3 | Conclusion | 75 |
| **Chapter 4** | **A Theory and Methodology of Inductive Learning**<br>*Ryszard S. Michalski* | **83** |
| 4.1 | Introduction | 83 |
| 4.2 | Types of Inductive Learning | 87 |
| 4.3 | Description Language | 94 |
| 4.4 | Problem Background Knowledge | 96 |
| 4.5 | Generalization Rules | 103 |
| 4.6 | The Star Methodology | 112 |
| 4.7 | An Example | 116 |
| 4.8 | Conclusion | 123 |
| 4.A | Annotated Predicate Calculus (APC) | 130 |
| **PART THREE** | **LEARNING IN PROBLEM-SOLVING AND PLANNING** | **135** |
| **Chapter 5** | **Learning by Analogy: Formulating and Generalizing Plans from Past Experience**<br>*Jaime G. Carbonell* | **137** |
| 5.1 | Introduction | 137 |
| 5.2 | Problem-Solving by Analogy | 139 |
| 5.3 | Evaluating the Analogical Reasoning Process | 149 |
| 5.4 | Learning Generalized Plans | 151 |
| 5.5 | Concluding Remark | 159 |
| **Chapter 6** | **Learning by Experimentation: Acquiring and Refining Problem-Solving Heuristics**<br>*Tom M. Mitchell, Paul E. Utgoff, and Ranan Banerji* | **163** |
| 6.1 | Introduction | 163 |
| 6.2 | The Problem | 164 |
| 6.3 | Design of LEX | 167 |
| 6.4 | New Directions: Adding Knowledge to Augment Learning | 180 |
| 6.5 | Summary | 189 |
| **Chapter 7** | **Acquisition of Proof Skills in Geometry**<br>*John R. Anderson* | **191** |
| 7.1 | Introduction | 191 |
| 7.2 | A Model of the Skill Underlying Proof Generation | 193 |
| 7.3 | Learning | 201 |
| 7.4 | Knowledge Compilation | 202 |


BIBLIOTHEQUE DU CERIST

| | | |
|---|---|---|
| 7.5 | Summary of Geometry Learning | 217 |
| **Chapter 8** | **Using Proofs and Refutations to Learn from Experience**<br>*Frederick Hayes-Roth* | **221** |
| 8.1 | Introduction | 221 |
| 8.2 | The Learning Cycle | 222 |
| 8.3 | Five Heuristics for Rectifying Refuted Theories | 225 |
| 8.4 | Computational Problems and Implementation Techniques | 234 |
| 8.5 | Conclusions | 238 |
| **PART FOUR** | **LEARNING FROM OBSERVATION AND DISCOVERY** | **241** |
| **Chapter 9** | **The Role of Heuristics in Learning by Discovery: Three Case Studies**<br>*Douglas B. Lenat* | **243** |
| 9.1 | Motivation | 243 |
| 9.2 | Overview | 245 |
| 9.3 | Case Study 1: The AM Program; Heuristics Used to Develop New Knowledge | 249 |
| 9.4 | A Theory of Heuristics | 263 |
| 9.5 | Case Study 2: The Eurisko Program; Heuristics Used to Develop New Heuristics | 276 |
| 9.6 | Heuristics Used to Develop New Representations | 282 |
| 9.7 | Case Study 3: Biological Evolution; Heuristics Used to Generate Plausible Mutations | 286 |
| 9.8 | Conclusions | 302 |
| **Chapter 10** | **Rediscovering Chemistry With the BACON System**<br>*Pat Langley, Gary L. Bradshaw, and Herbert A. Simon* | **307** |
| 10.1 | Introduction | 307 |
| 10.2 | An Overview of BACON.4 | 309 |
| 10.3 | The Discoveries of BACON.4 | 312 |
| 10.4 | Rediscovering Nineteenth Century Chemistry | 319 |
| 10.5 | Conclusions | 326 |


BIBLIOTHEQUE DU CERIST

| | | |
|---|---|---|
| **Chapter 11** | **Learning From Observation: Conceptual Clustering**<br>*Ryszard S. Michalski and Robert E. Stepp* | **331** |
| 11.1 | Introduction | 332 |
| 11.2 | Conceptual Cohesiveness | 333 |
| 11.3 | Terminology and Basic Operations of the Algorithm | 336 |
| 11.4 | A Criterion of Clustering Quality | 344 |
| 11.5 | Method and Implementation | 345 |
| 11.6 | An Example of a Practical Problem: Constructing a Classification Hierarchy of Spanish Folk Songs | 358 |
| 11.7 | Summary and Some Suggested Extensions of the Method | 360 |
| **PART FIVE** | **LEARNING FROM INSTRUCTION** | **365** |
| **Chapter 12** | **Machine Transformation of Advice into a Heuristic Search Procedure**<br>*David Jack Mostow* | **367** |
| 12.1 | Introduction | 367 |
| 12.2 | Kinds of Knowledge Used | 370 |
| 12.3 | A Slightly Non-Standard Definition of Heuristic Search | 374 |
| 12.4 | Instantiating the HSM Schema for a Given Problem | 378 |
| 12.5 | Refining HSM by Moving Constraints Between Control Components | 384 |
| 12.6 | Evaluation of Generality | 398 |
| 12.7 | Conclusion | 399 |
| 12.A | Index of Rules | 403 |
| **Chapter 13** | **Learning by Being Told: Acquiring Knowledge for Information Management**<br>*Norm Haas and Gary G. Hendrix* | **405** |
| 13.1 | Overview | 405 |
| 13.2 | Technical Approach: Experiments with the KLAUS Concept | 408 |
| 13.3 | More Technical Details | 413 |
| 13.4 | Conclusions and Directions for Future Work | 418 |
| 13.A | Training NANOKLAUS About Aircraft Carriers | 422 |


BIBLIOTHEQUE DU CERIST

**Chapter 14** **The Instructible Production System: A Retrospective Analysis** 429
*Michael D. Rychener*

14.1 The Instructible Production System Project 430
14.2 Essential Functional Components of Instructible Systems 436
14.3 Survey of Approaches 443
14.4 Discussion 453

**PART SIX** **APPLIED LEARNING SYSTEMS** 461

**Chapter 15** **Learning Efficient Classification Procedures and their Application to Chess End Games** 463
*J. Ross Quinlan*

15.1 Introduction 463
15.2 The Inductive Inference Machinery 465
15.3 The Lost N-ply Experiments 470
15.4 Approximate Classification Rules 474
15.5 Some Thoughts on Discovering Attributes 477
15.6 Conclusion 481

**Chapter 16** **Inferring Student Models for Intelligent Computer-Aided Instruction** 483
*Derek H. Sleeman*

16.1 Introduction 483
16.2 Generating a Complete and Non-redundant Set of Models 488
16.3 Processing Domain Knowledge 503
16.4 Summary 507
16.A An Example of the SELECTIVE Algorithm: LMS-I's Model Generation Algorithm 510

**Comprehensive Bibliography of Machine Learning** 511
*Paul E. Utgoff and Bernard Nudel*

**Glossary of Selected Terms In Machine Learning** 551

**About the Authors** 557

**Author Index** 563

**Subject Index** 567

BIBLIOTHEQUE DU CERIST