Systems That Learn

The MIT Press Series in Learning, Development, and Conceptual Change
Lila Gleitman, Susan Carey, Elissa Newport, and Elizabeth Spelke, editors

Names for Things: A Study in Human Learning, by John Macnamara, 1982

Conceptual Change in Childhood, by Susan Carey, 1985

"Gavagai!" or the Future History of the Animal Language Controversy, by David Premack, 1985

Systems That Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists, by Daniel N. Osherson, Michael Stob, and Scott Weinstein, 1986

Systems That Learn
An Introduction to Learning Theory for Cognitive and Computer Scientists

Daniel N. Osherson
Michael Stob
Scott Weinstein

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

Second printing, 1986

© 1986 by The Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

This book was set in Times New Roman by Asco Trade Typesetting Ltd., Hong Kong, and printed and bound by Halliday Lithograph in the United States of America

Library of Congress Cataloging-in-Publication Data

Osherson, Daniel N.
Systems that learn.

(The MIT Press series in learning, development, and conceptual change)
"A Bradford book."
Bibliography: p.
Includes indexes.
1. Learning—Mathematical aspects. 2. Learning, Psychology of. 3. Human information processing—Mathematical models. I. Stob, Michael. II. Weinstein, Scott. III. Title. IV. Series.
BF318.083 1986 153.1'5'015113 85-19759
ISBN 0-262-15030-1

ISBN 0-262-65024-X (Paperback)

MIT Press
0262150301
OSHERSON
SYST THAT LEARN

Contents

	Series Foreword	xi
	Preface	xiii
	Acknowledgments	xv
	How to Use This Book	xvii
	Introduction	1
I	**IDENTIFICATION**	5
1	**Fundamentals of Learning Theory**	7
1.1	Learning Paradigms	7
1.2	Background Material	8
	1.2.1 Functions and Recursive Functions	8
	1.2.2 Recursively Enumerable Sets	10
1.3	Identification: Basic Concepts	11
	1.3.1 Languages	12
	1.3.2 Hypotheses	13
	1.3.3 Environments	13
	1.3.4 Learners	14
1.4	Identification: Criterion of Success	17
	1.4.1 Identifying Texts	17
	1.4.2 Identifying Languages	18
	1.4.3 Identifying Collections of Languages	19
1.5	Identification as a Limiting Process	22
	1.5.1 Epistemology of Convergence	22
	*1.5.2 Self-Monitoring Learning Functions	23
2	**Central Theorems on Identification**	25
2.1	Locking Sequences	25
2.2	Some Unidentifiable Collections of Languages	27
2.3	A Comprehensive, Identifiable Collection of Languages	29
2.4	Identifiable Collections Characterized	30
2.5	Identifiability of Single-Valued Languages	31

3	**Learning Theory and Natural Language**	34
3.1	Comparative Grammar	34
3.2	Learning Theory and Linguistic Development	36
	3.2.1 How Many Grammars for the Young Child?	36
	3.2.2 Are the Child's Conjectures a Function of Linguistic Input?	37
	3.2.3 What Is a Natural Language?	38
	3.2.4 Idealization	40
II	**IDENTIFICATION GENERALIZED**	43
4	**Strategies**	45
4.1	Strategies as Sets of Learning Functions	45
4.2	Computational Constraints	47
	4.2.1 Computability	47
	4.2.2 Time Bounds	50
	4.2.3 On the Interest of Nonrecursive Learning Functions	53
4.3	Constraints on Potential Conjectures	53
	4.3.1 Totality	53
	4.3.2 Nontriviality	54
	4.3.3 Consistency	56
	4.3.4 Prudence and r.e. Boundedness	59
	4.3.5 Accountability	61
	*4.3.6 Simplicity	63
4.4	**Constraints on the Information Available to a Learning Function**	66
	4.4.1 Memory-Limitation	66
	***4.4.2 Set-Driven Learning Functions**	73
4.5	**Constraints on the Relation between Conjectures**	74
	4.5.1 Conservatism	75
	4.5.2 Gradualism	77
	4.5.3 Induction by Enumeration	78
	***4.5.4 Caution**	79
	***4.5.5 Decisiveness**	80

4.6	Constraints on Convergence	82
	4.6.1 Reliability	82
	4.6.2 Confidence	86
	4.6.3 Order Independence	88
*4.7	Local and Nonlocal Strategies	92
5	**Environments**	96
5.1	Order and Content in Natural Environments	96
5.2	Texts with Blanks	96
5.3	Evidential Relations	98
5.4	Texts with Imperfect Content	100
	5.4.1 Noisy Text	100
	5.4.2 Incomplete Text	103
	*5.4.3 Imperfect Text	105
5.5	Constraints on Order	106
	5.5.1 Ascending Text	106
	5.5.2 Recursive Text	107
	*5.5.3 Nonrecursive Text	109
	*5.5.4 Fat Text	110
5.6	Informants	113
	5.6.1 Informants and Characteristic Functions	113
	5.6.2 Identification on Informant	115
	*5.6.3 Memory-Limited Identification on Informant	116
5.7	A Note on "Reactive" Environments	118
6	**Criteria of Learning**	119
6.1	Convergence Generalized	119
	6.1.1 Convergence Criteria	119
	6.1.2 Identification Relativized	120
6.2	Finite Difference, Intensional Identification	123
	6.2.1 FINT-Identification on Text	123
	6.2.2 FINT-Identification on Imperfect Text	125
	6.2.3 FINT-Identification in RE_{svt}	126
6.3	Extensional Identification	129

6.3.1 EXT-Identification in RE 130
6.3.2 EXT-Identification in RE_{svt} 132

*6.3.3 Finite Difference, Extensional Identification 133
*6.4 Bounded Extensional Identification 134
6.4.1 BEXT-Identification in RE 135
6.4.2 BEXT-Identification in RE_{svt} 137
6.4.3 Bounded Finite Difference Extensional Identification 138
*6.5 Finite Difference Identification 139
6.5.1 FD-Identification in RE 139
6.5.2 FD-Identification in RE_{svt} 140
6.5.3 Bounded Finite Difference Identification 141
*6.6 Simple Identification 143
6.7 Summary 144
6.8 Characteristic Index Identification 145
6.8.1 CI-Convergence 145
6.8.2 CI-Identification on Text and on Informant 147
*6.8.3 Variants of CI-Identification 149

7 Exact Learning 152

7.1 Paradigms of Exact Learning 152
*7.2 A Characterization of $[\mathscr{F}^{rec}, \text{text}, \text{INT}]^{ex}$ 155
7.3 Earlier Paradigms Considered in the Context of Exact Learning 157
7.3.1 Strategies and Exact Learning 157
*7.3.2 Environments and Exact Learning 159
7.3.3 Convergence Criteria and Exact Learning 160
*7.4 Very Exact Learning 160
7.5 Exact Learning in Generalized Identification Paradigms 162

III OTHER PARADIGMS OF LEARNING 165

8 Efficient Learning 167

8.1 Text-Efficiency 167
8.2 Text-Efficient Identification 170

8.2.1 Text-Efficient Identification in the Context of $\mathscr{F}$ 170
8.2.2 Text-Efficient Identification and Rational Strategies 171
8.2.3 Text-Efficient Identification in the Context of $\mathscr{F}^{rec}$ 172
8.2.4 Text-Efficiency and Induction by Enumeration 174
*8.2.5 Text-Efficiency and Simple Identification 175
8.3 Efficient Identification 176

9 Sufficient Input for Learning 178
9.1 Locking Sequences as Sufficient Input 178
9.2 Recursive Enumerability of LS_φ 179
*9.3 Predictability in Other Learning Paradigms 180

***10 Topological Perspective on Learning** 182
10.1 Identification and the Baire Space 182
10.2 Continuity of Learning Functions 183
10.3 Another Proof of Proposition 2.1A 184
10.4 Locking Texts 185
10.5 Measure One Learning 186
10.5.1 Measures on Classes of Texts 186
10.5.2 Measure One Identifiability 187
10.5.3 Uniform Measures 188
10.6 Probabilistic Learning 190

Bibliography 195
List of Symbols 198
Name Index 201
Subject Index 203

Series Foreword

This series in learning, development, and conceptual change will include state-of-the-art reference works, seminal book-length monographs, and texts on the development of concepts and mental structures. It will span learning in all domains of knowledge, from syntax to geometry to the social world, and will be concerned with all phases of development, from infancy through adulthood.

The series intends to engage such fundamental questions as

The nature and limits of learning and maturation: the influence of the environment, of initial structures, and of maturational changes in the nervous system on human development; learnability theory; the problem of induction; domain specific constraints on development.

The nature of conceptual change: conceptual organization and conceptual change in child development, in the acquisition of expertise, and in the history of science.

Lila Gleitman
Susan Carey
Elissa Newport
Elizabeth Spelke

Preface

It is a familiar observation that an organism's genotype may be conceived as a function that maps potential environments into potential phenotypes. Relativizing this conception to cognitive science allows human intellectual endowment to be construed as a particular function mapping early experience into mature cognitive competence. The function might be called "human nature relative to cognition." Learning theory is a mathematical tool for the study of this function. This book attempts to acquaint the reader with the use of this tool.

Less cryptically, learning theory is the study of systems that map evidence into hypotheses. Of special interest are the circumstances under which these hypotheses stabilize to an accurate representation of the environment from which the evidence is drawn. Such stability and accuracy are conceived as the hallmarks of learning. Within learning theory, the concepts "evidence," "stabilization," "accuracy," and so on, give way to precise definitions.

As developed in this book, learning theory is a collection of theorems about certain kinds of number-theoretic functions. We have discussed the application of such theorems to cognitive science and epistemology in a variety of places (e.g., Osherson, Stob, and Weinstein, 1984, 1985, 1985a; Osherson and Weinstein, 1982a, 1984, 1985). In contrast, the present work centers on the mathematical development of learning theory rather than on empirical hypotheses about human learning. As an aid to intuition, however, we have attempted to concretize the formal developments in this book through extended discussion of first language acquisition.

We have not tried to survey the immense field of machine inductive inference. Rather, we have selected for presentation just those results that seem to us to clarify questions relevant to human intellectual development. Several otherwise fascinating topics in machine learning have thus been left aside. Our choices no doubt reflect tacit theoretical commitments not universally shared. An excellent review of many topics passed over here is provided by Angluin and Smith (1982). Our own previously published work in the technical development of learning theory (e.g., Osherson and Weinstein, 1982, 1982a; Osherson, Stob, and Weinstein, 1982, 1982a, 1985) is entirely integrated herein.

Our concern in the present work for the mathematical development of learning theory has resulted in rigorous exposition. Less formal introductions to the central concepts and topics of learning theory are available in Osherson and Weinstein (1984) and Osherson, Stob, and Weinstein (1984).

We would be pleased to receive from our readers comments and corrections, as well as word of new results.

Acknowledgments

Our principal intellectual debts are to the works of E. Mark Gold and Noam Chomsky. Gold (1967) established the formal framework within which learning theory has developed. Chomsky's writings have revealed the intimate connection between the projection problem and human intelligence. In addition, we have been greatly influenced by the research of Blum and Blum (1975), Angluin (1980), Case and Smith (1983), and Wexler and Culicover (1980). Numerous conversations with Lila Gleitman and with Steven Pinker have helped us to appreciate the bearing of learning theory on empirical studies of first language acquisition, and conversely the bearing of first language acquisition studies on learning theory. We thank them for their patient explanations.

Preparation of the manuscript was facilitated by a grant to Osherson from the Fyssen Foundation for 1983–84 and by National Science Foundation Grants MCS 80-02937 and 82-00032 to Stob. We thank these agencies for their support.

How to Use This Book

Mathematical prerequisites for this text include elementary set theory and an intuitive understanding of the concept "computable function." Lewis and Papadimitriou (1981) provide an excellent introduction to this material. Acquaintance with the elementary portion of recursion theory is also advisable. We recommend Machtey and Young (1978).

Starred material in the text is of more advanced character and may be omitted without loss of continuity. We have relegated considerable exposition to the exercises, which should be at least attempted.

Definitions, examples, lemmas, propositions, open questions, and exercises are numbered independently within the section or subsection in which they appear. Thus proposition 4.4.1B refers to the second proposition of section 4.4.1; it appears before lemma 4.4.1A, the first lemma of the same section. Symbol, subject, and name indexes may be found at the end of the book.

We use standard set-theoretic notation and recursion-theoretic notation drawn from Rogers (1967) throughout. Note that $\subset$ denotes proper inclusion, whereas $\subseteq$ denotes (possibly improper) inclusion.

Systems That Learn

Introduction

Let us play a game.

We have selected a set of numbers, and you must guess the set that we have in mind. The set consists of every positive integer with a sole exception. Thus the set might be $\{2, 3, 4, 5, \ldots\}$ or $\{1, 3, 4, 5, \ldots\}$ or $\{1, 2, 4, 5, \ldots\}$, etc. We will give you an unlimited number of clues about the set, and you are to guess after each clue. We will never tell you whether you are right.

First clue: The set contains the number 1.

Please guess the set we have in mind. Would you like to guess the set $\{2, 3, 4, 5, \ldots\}$? (That would be unwise.)

Second clue: The set contains the number 3.

Please make another guess. How about $\{1, 2, 3, 4, 5, 6, 8, 9, 10, \ldots\}$ or does that seem arbitrary to you?

Third clue: The set contains the number 4.

Go ahead and guess.

Fourth clue: The set contains the number 2.

Does the fourth clue surprise you? Guess again.

Fifth clue: The set contains the number 6.

Guess.

Sixth clue: The set contains the number 7.

Guess.

Seventh clue: The set contains the number 8.

Guess.

We interrupt the game at this point because we would like to ask you some questions about it.

First question: Are you confident about your seventh guess? Give an example of an eighth clue that would lead you to repeat your last guess. Give an example of an eighth clue that would lead you to change your guess.

Second question: Let us say that a "guessing rule" is a list of instructions for converting the clues received up to a given point into a guess about the set we have in mind. Were your guesses chosen according to some guessing rule, and if so, which one?

Third question: What should count as winning the game? Consider the following criterion: You win just in case at least one of your guesses is right. This criterion makes winning the game too easy. Say why.

Fourth question: We advocate the following criterion: You win just in case you eventually make the right guess and subsequently never change your mind regardless of the new clues you receive. In this case let us say

that you "win in the limit." Is it possible to win the game in the limit even though you make one hundred wrong guesses? Is there any number of wrong guesses that is logically incompatible with winning the game in the limit?

Fifth question: Suppose that all the clues we give you are of the form: The set contains the number n. Suppose furthermore that for every positive integer i, we eventually give you a clue of this form if and only if i is in fact contained in the set we have in mind. (So for every number i in our set, you are eventually told that the set contains i; also you receive no false information about the set.) Do not suppose anything about the order in which you will get all these clues. We will order them any way we please. (Recall how we surprised you with the fourth clue.) Now let us call a guessing rule "winning" just in case the following is true. If you use the rule to choose your guesses, then no matter which of the sets we have in mind, you are guaranteed to win the game in the limit. Specify a winning guessing rule for our game.

Sixth question: We make the game harder. This time we are allowed to select any of the sets that are legal in the original game, but we may also select the set $\{1, 2, 3, 4, 5, 6, \ldots\}$ of all positive integers. The rules about clues are the same as given in question 5. Play this new game with a friend, and then think about the following question. Is there a winning guessing rule for the new game?

Seventh question: Let us make the last game easier. The choice of sets is the same as in the last game, but we now agree to order our clues in a certain way. For all positive integers i and j, if both i and j are included in the set we have in mind, and if i is less than j, then you will receive the clue "The set contains i" before you receive the clue "The set contains j." Can you specify a winning guessing rule for this version of the game?

Eighth question: Here is another variant. We select a set from the original collection (thus the set $\{1, 2, 3, 4, 5, \ldots\}$ of all positive integers is no longer allowed). Clues can be given in any order we please. You get only one guess. You may wait to see as many clues as you like, but your first guess is definitive. Play this game with a friend. Then show that no matter what rule you use to make your guess, you are not guaranteed to be right. Think about what happens if you are allowed two guesses in the game.

The games we have been playing resemble the process of scientific discovery. Nature plays our role, selecting a certain pattern that is imposed on the world. The scientist plays your role, examining an endless series

of clues about this pattern. In response to the clues, the scientist emits guesses. Nature never says whether the guesses are correct. Scientific success consists of eventually offering a correct guess and never deviating from it thereafter. Language acquisition by children can also be construed in terms of our game. The child's parents have a certain language in mind (the one they speak). They provide clues in the form of sentences. The child converts these clues into guesses about the parents' language. Acquisition is successful just in case the child eventually sticks with a correct guess.

The similarity of our game to these and other settings makes it worthy of more careful study. We would like to know which versions of the game are winnable and by what kinds of guessing rules. Research on these questions began in the 1960s by Putnam (1975), Solomonoff (1964), and Gold (1967). These initial investigations have given rise to a large literature in computer science, linguistics, philosophy, and psychology. This body of theoretical and applied results is generally known as *learning theory* because many kinds of learning (e.g., language acquisition) can be construed as successful performance in one of our games.

In this book we attempt to develop learning theory in systematic fashion, presupposing only basic notions from set theory and the theory of computation. Throughout our exposition, definitions and theorems are illustrated by consideration of language acquisition. However, no serious application of the theory is described.

The book is divided into three parts. Part I advances a fundamental model of learning due essentially to Gold (1967). Basic notation, terminology, and theorems are there presented, to be relied on in all subsequent discussion. In part II these initial definitions are generalized and varied in dozens of ways, giving rise to a multitude of learning models and theorems. We attempt to impose some order on these results through a system of notation and classification. Part III explores diverse issues in learning theory that do not fit neatly into the classification offered in part II.