TEXTS AND MONOGRAPHS IN COMPUTER SCIENCE

ADAPTIVE INFORMATION PROCESSING

An Introductory Survey

Jeffrey R. Sampson

Springer-Verlag
New York Heidelberg Berlin

Adaptive Information Processing

An Introductory Survey

Jeffrey R. Sampson

Springer-Verlag
New York Heidelberg Berlin
1976

Jeffrey R. Sampson

Department of Computing Science
The University of Alberta
Edmonton, Alberta T6G 2E1
Canada

editors

F. L. Bauer

Mathematisches Institut der
Technischen Hochschule
8000 München
Arcisstrasse 21
West Germany

David Gries

Cornell University
Department of Computer Science
Upson Hall
Ithaca, New York 14859
USA

AMS Subject Classification: 68-01, 68A25, 68A30, 68A35, 68A45, 68A55, 92A05
(C.R.) Computing Classifications: 3.12, 3.6, 5.22, 5.26, 5.6

Library of Congress Cataloging in Publication Data

Sampson, Jeffrey R. 1942–
Adaptive information processing.

(Texts and monographs in computer science)
"Began as a series of lecture notes for a course . . . at the University of Alberta . . . first taught in 1973."
Includes bibliographies.
1. Machine theory. 2. Information theory in biology. 3. Artificial intelligence.
I. Title.
QA267.S25 001.53 76-8470

Printed in the United States of America

ISBN 0-387-07739-1 Springer-Verlag New York

ISBN 3-540-07739-1 Springer-Verlag Berlin Heidelberg

Texts and Monographs in Computer Science

F. L. Bauer
David Gries
editors

Preface

This book began as a series of lecture notes for a course called Introduction to Adaptive Systems which I developed for undergraduate Computing Science majors at the University of Alberta and first taught in 1973. The objective of the course has been threefold: (1) to expose undergraduate computer scientists to a variety of subjects in the theory and application of computation, subjects which are too often postponed to the graduate level or never taught at all; (2) to provide undergraduates with a background sufficient to make them effective participants in graduate level courses in Automata Theory, Biological Information Processing, and Artificial Intelligence; and (3) to present a personal viewpoint which unifies the apparently diverse aspects of the subject matter covered.

All of these goals apply equally to this book, which is primarily designed for use in a one semester undergraduate computer science course. I assume the reader has a general knowledge of computers and programming, though not of particular machines or languages. His mathematical background should include basic concepts of number systems, set theory, elementary discrete probability, and logic.

As befits an introductory survey, on the other hand, I have kept the need for specialized knowledge both circumscribed and minimal. Probability is used only in Chapters 1 and 2, and logic only in Chapters 3 and 13, for example. Nowhere is the treatment so rigorous as to preclude use of the book by college students in other disciplines (mathematics, psychology, biology) or even by advanced high school students. The book could also serve as a useful adjunct to more specialized computer science courses, where it would introduce the student to material related to the study of, say, Artificial Intelligence.

With this last possibility in mind, I have kept each of the three major parts of the book as self contained as possible. Although each part begins with some prefatory remarks and a survey of the chapters to follow, I have also included an Introduction which provides a general and unifying overview of the book. A short list of references and annotated bibliographical items is appended to each chapter, as are a few suggested exercises.

Although fully responsible for the views expressed and the topics included in this book, I am deeply indebted to John Holland and others in the Department of Computer and Communication Sciences at the University of Michigan. My graduate training there remains the major source of ideas for how this book should be put together and what it should say.

Several of my colleagues at the University of Alberta, including I-Ngo Chen, Wayne Davis, and Kelly Wilson, have provided useful suggestions concerning the content and organization of the course out of which this book grew; Wilson also commented extensively on the manuscript. Ken Morgan of the Genetics Department read Part II and made many helpful suggestions, as did Dick Peddicord of the University of San Francisco. I am most deeply indebted to my Computing Science colleague Len Schubert, who tested the manuscript in the classroom and supplied numerous major improvements. Many sections of Part III rely heavily on Schubert's observations and unpublished notes.

Preparation of the manuscript was done using the *FMT system on the University of Alberta IBM 360/67 computer. The assistance of Glynis Dorey, Mireille Dubreuil, Anton Kritzinger, Robert Mercer, Fran Russell, Jennifer Semchuck, and Elaine Soetaert, and the support of the National Research Council of Canada, are gratefully acknowledged.

Reference

Sampson, J. R. An introductory adaptive systems course for undergraduate computer science majors. *ACM SIGSCE Fourth Technical Symposium on Computer Science Education*, *SIGSCE Bulletin*, February, 1974, *6*(1): 148–151.

Contents

Introduction

Machinery has always fascinated man. In both the workings of his world and the devices of his creation he marvels at intricacy of design and cleverness of function. He is perhaps most often impressed by a machine, natural or artificial, which is designed to modify its function to suit changing circumstances. A prerequisite to such adaptive behavior is the ability to process information—to receive, store, retrieve, use, and transmit it. Modern computers represent the ultimate artificial information processing machines so far created. In the realm of natural information processing machinery nothing rivals the brain of the computer's creator.

This book is about information processing machines. It is especially but not exclusively about machines which process information in order to work better. Our domain of discussion will be wide ranging, traversing abstract models of computation, biological information processing systems, and efforts to program intelligent behavior in computers. It is in the nature of our exploratory survey that we shall not dwell at length on individual topics or detailed examples. Rather I will sketch information processing features in broad strokes, highlighting instances of adaptive behavior.

This book is not about computers. At least it is not about electronic stored program digital computers. I assume you already know something about these machines and how they are programmed. In another sense this book is entirely about computers, if the term may embrace machines built of abstract components, of biochemical molecules, or of complex algorithms. I hope this book will induce you to appreciate the fundamental similarities which underlie so diverse a collection of information processing machines.

Each of the three parts which follow can be read more or less independently of the others. And each has its own introduction which previews the chapters that follow. For the moment then, let me loosely weave together the material of all fifteen chapters.

We begin with information, that abstract elusive stuff which must somehow be processed. We discover that information can be mathematically defined and measured, allowing us to find efficient and reliable ways of coding messages to be transmitted. Our first information processing machines are the highly abstract ones known as automata. We study the formal descriptions which allow us to characterize the information processing, or behavior, of any finite automaton. Turning to a kind of infinite automata known as Turing machines, we discover a form of behavior that can compute anything which is computationally possible. Such an ultimate computer might signal a happy ending to our search for adaptive machines, were it not for insurmountable practical limitations on the use of Turing machines. Another approach to universal computation behavior arises in the last type of abstract machine we study, the cellular automaton. Made up of relatively simple component machines or cells, these automata derive their information processing power from the interconnection of large arrays of components.

From the abstract cells of an automaton we turn in Part II to the biological cells of an organism. Here we encounter an overwhelming richness of adaptive behavior. The nature and expression of the long mysterious genetic code can now be almost completely explained. And we learn something about the ways in which biological information is transmitted from an organism to its offspring. Yet even the awesome information processing ability of a single cell is dwarfed by that of large intercommunicating networks of specialized nerve cells. In the brains of higher animals and man we find the ultimate in nervous system function. After a quick look at how such brains are put together we study the ways in which information is transmitted throughout the nervous system. Yet all this computing power would be useless for adaptive purposes if the organism could not acquire information from and influence its environment. So we look at some of the input–output systems with which the living computer has been endowed. Perhaps not surprisingly, both neural and genetic information processing techniques will give us cause to recall and apply the fundamental notions of information and coding with which we began. Finally we look at some examples of how computers have proved useful in simulation studies of natural adaptive systems.

When computers became powerful enough to simulate natural adaptation people began to consider ways in which a computer might be programmed to behave intelligently. Part III is concerned with aspects of Artificial Intelligence. To see what has been done on the information input side we look at various efforts in pattern recognition. Then we discuss how and with what success computers have been programmed to play games, prove mathematical theorems, solve problems, and converse in human language. We will be tempted to compare the artificial accomplishments with those of the natural systems already considered. And the artificial intelligentsia will usually disappoint us. But we must realize that the natural adaptive systems are the product of millions of years of evolution (itself a natural adaptive process).

At the end of each chapter you will find, in a section called Bibliography, a short list of some of the works that have most influenced my thinking about the subject of that chapter, with a comment or two about each item. Following the Bibliography will be a short set of Exercises. These range from solution of specific problems to suggestions for thought or essay, as appropriate to the subject matter of the chapter. In many cases the answers involve integration of material from two or more chapters. The exercises are mainly intended to stimulate your thinking and should not be regarded as thoroughly testing your knowledge of the contents of the chapter. By the same token, your experience with all these subjects will probably be far less rewarding if you do not at least attempt most of the exercises.

Part I
Information and automata

The five chapters which follow contain the most abstract material in the book. Yet it is appropriate to begin here, rather than with biological systems or intelligent programs, so that we can approach these latter subjects equipped with suitable formal tools and useful ways of thinking about information processing machines. We will not encounter much adaptive behavior in the automata of these chapters. But we will lay the foundations for a later understanding of how such machines can be used to model or simulate the behavior of other adaptive systems.

The first two chapters outline a formal approach to information. Chapter 1 treats the mathematical theory of communication in terms of the components of an idealized communication system. We learn how to measure the information contained in a message or set of messages. We review some truly surprising results concerning how accurately a given amount of information can be transmitted over a communication channel with a given capacity.

Chapter 2 first describes ways in which we can encode messages in order to obtain the fastest possible communication. For cases where accuracy is more important than speed, we study a group of special codes that can detect, and in some cases correct, errors in a transmitted message.

In Chapter 3 we meet our first information processing machines, finite automata, in three different but equivalent forms. We build automata to perform simple tasks, using networks, functions, and formal language expressions as components. And we learn some ways of transforming some of these representations into others.

The Turing machines in Chapter 4 are in many ways a natural extension of finite automata. Yet the computational ability of these infinite automata turns out to be essentially unlimited. We will quickly forsake the original formulation of Turing machines for a programmable version which is more natural for those with computer programming experience. We sketch a programmable form of Turing machine which can do anything that can be done by any other Turing machine. Such universal Turing machines will lead us into questions of just what classes of problems can be solved and whether there are questions that can never be answered. Finally we explore the relation between automata and formal grammars.

Chapter 5 recounts selected developments in the relatively new field of cellular automata. Of particular interest to those who work with these large arrays of simple machines are questions of the limits on computational power. What sort of cellular machine is required, for example, to emulate the behavior of universal Turing machines? How can self-reproducing machines be realized in cellular spaces? We look at two rather different kinds of answers to these questions.

1
Communication theory

In the late 1940's two wartime research efforts were troubled by the problem of faithfully reproducing a signal in the presence of interfering noise. At MIT, Norbert Wiener, better known as the inventor of cybernetics, was seeking reliable prediction for automatic fire control. Claude Shannon, working at the Bell Telephone Laboratories, wanted to make optimal use of communication channels for transmission of coded messages. These two men laid the foundations of what is now known as statistical or mathematical communication theory (or, more popularly, information theory). In this chapter we explore aspects of Shannon's contributions, as set forth in his 1948 paper entitled "The Mathematical Theory of Communication."

It is important to clarify in what sense the terms communication and information are to be used in our discussion, since they have a great variety of nontechnical interpretations. In his excellent essay on Shannon's work, Warren Weaver distinguishes three levels of communication problems: (1) the technical level, concerned with accurate transmission of the symbols of a message; (2) the semantic level, concerned with precise conveyance of the intended meaning of the message; and (3) the effectiveness level, concerned with appropriate impact of the message on the recipient's behavior. Communication theory deals only with technical problems. The meaning, significance, veracity, and effect of a message are of no concern to us in what follows. Messages are treated strictly in terms of their probabilities of occurrence. The numerical measure of information may be equal for two messages, one meaningful and moving, the other unintelligible nonsense.