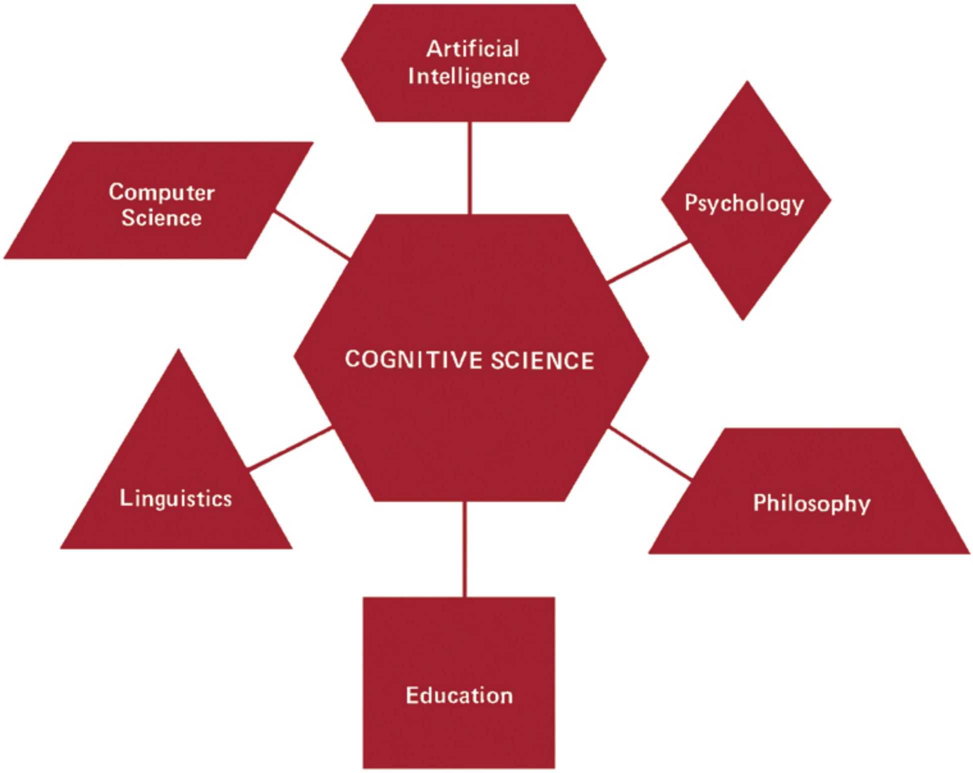
LANGUAGE, THOUGHT, AND CULTURE
Advances in the Study of Cognition

REPRESENTATION AND UNDERSTANDING

Studies in Cognitive Science



Edited by

Daniel G. Bobrow / Allan Collins

Academic Press, Inc. A Subsidiary of Harcourt Brace Jovanovich, Publishers

Representation and Understanding

Studies in Cognitive Science

LANGUAGE, THOUGHT, AND CULTURE: *Advances in the Study of Cognition*

Under the Editorship of: E. A. HAMMEL

DEPARTMENT OF ANTHROPOLOGY
UNIVERSITY OF CALIFORNIA
BERKELEY

Michael Agar, Ripping and Running: A Formal Ethnography of Urban Heroin Addicts

Brent Berlin, Dennis E. Breedlove, and Peter H. Raven, Principles of Tzeltal Plant Classification: An Introduction to the Botanical Ethnography of a Mayan-Speaking People of Highland Chiapas

Mary Sanches and Ben Blount, Sociocultural Dimensions of Language Use

Daniel G. Bobrow and Allan Collins, Representation and Understanding: Studies in Cognitive Science

Representation and Understanding

Studies in Cognitive Science

edited by

DANIEL G. BOBROW
Xerox Palo Alto Research Center
Palo Alto, California

and

ALLAN COLLINS
Bolt Beranek and Newman
Cambridge, Massachusetts

ACADEMIC PRESS, INC. New York San Francisco London 1975

A Subsidiary of Harcourt Brace Jovanovich, Publishers

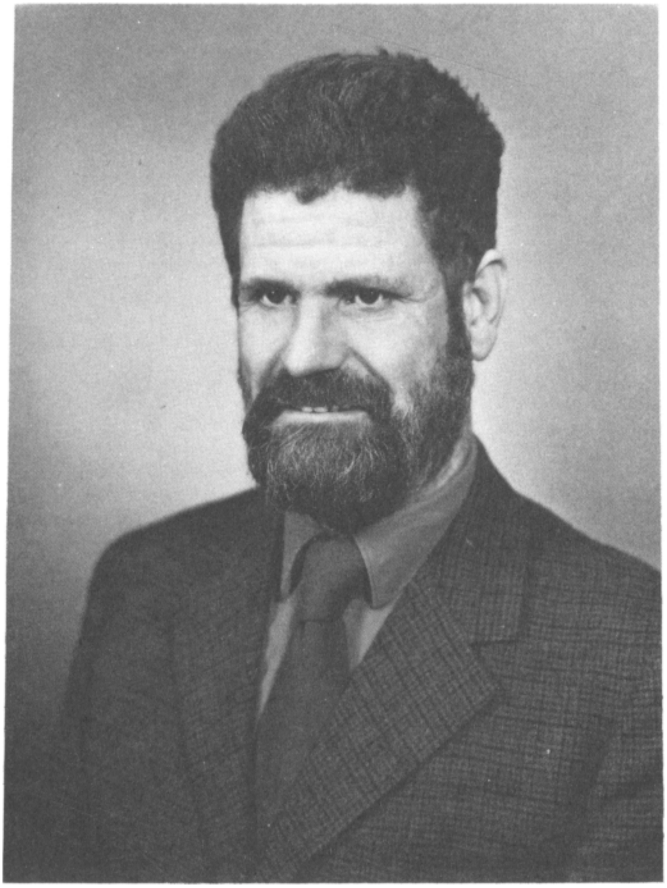
ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 75-21630

ISBN 0-12-108550-3

PRINTED IN THE UNITED STATES OF AMERICA



This book is affectionately dedicated to Jaime Carbonell
by his friends and colleagues.

This page intentionally left blank

CONTENTS

I. Theory of Representation

II. New Memory Models

III. Higher Level Structures

IV. Semantic Knowledge in Understander Systems

PREFACE

Jaime Carbonell was our friend and colleague. For many years he worked with us on problems in Artificial Intelligence, especially on the development of an intelligent instructional system. Jaime directed the Artificial Intelligence group at Bolt, Beranek, and Newman (in Cambridge, Massachusetts) until his death in 1973. Some of us who had worked with Jaime decided to hold a conference in his memory, a conference whose guiding principle would be that Jaime would have enjoyed it. This book is the result of that conference.

Jaime Carbonell's important contribution to cognitive science is best summarized in the title of one of his publications: *AI in CAI*. Jaime wanted to put principles of Artificial Intelligence into Computer-Assisted Instruction (CAI) systems. He dreamed of a system which had a data base of knowledge about a topic matter and general information about language and the principles of tutorial instruction. The system could then pursue a natural tutorial dialog with a student, sometimes following the student's initiative, sometimes taking its own intiative, but always generating its statements and responses in a natural way from its general knowledge. This system contrasts sharply with existing systems for Computer-Assisted Instruction in which a relatively fixed sequence of questions and possible responses have to be determined for each topic. Jaime did construct working versions of his dream--in a system which he called SCHOLAR. But he died before SCHOLAR reached the full realization of the dream.

It was a pleasure to work with Jaime. His kindness and his enthusiasm were infectious, and the discussions we had with him over the years were a great stimulus to our own thinking. Both as a friend and a colleague we miss him greatly.

Cognitive Science. This book contains studies in a new field we call *cognitive science*. Cognitive science includes elements of psychology, computer science, linguistics, philosophy, and education, but it is more than the intersection of these disciplines. Their integration has produced a new set of tools for dealing with a broad range of questions. In recent years, the interactions among the workers in these fields has led to exciting new developments in our understanding of intelligent systems and the development of a science of cognition. The group of workers has pursued problems that did not appear to be solvable from within any single discipline. It is too early to predict the future course of this new interaction, but the work to date has been stimulating and inspiring. It is our hope that this book can serve as an illustration of the type of problems that can be approached through interdisciplinary cooperation. The participants in this book (and at the conference) represent the fields of Artificial Intelligence, Linguistics, and Psychology, all of whom work on similar problems but with different viewpoints. The book focuses on the common problems, hopefully acting as a way of bringing these issues to the attention of all workers in those fields related to cognitive science.

Subject Matter. The book contains four sections. In the first section, **Theory of Representation**, general issues involved in building representations of knowledge are explored. Daniel G. Bobrow proposes that solutions to a set of design issues be used as dimensions for comparing different representations, and he examines different forms such solutions might take. William A. Woods explores problems in representing natural-language statements in semantic networks, illustrating difficult theoretical issues by examples. Joseph D. Becker is concerned with the representation one can infer for behavioral systems whose internal workings cannot be observed directly, and he considers the interconnection of useful concepts such as hierarchical organization, system goals, and resource conflicts. Robert J. Bobrow and John Seely Brown present a model for an expert understander which can take a collection of data describing some situation, synthesize a *contingent knowledge structure* which places the input data in the context of a larger structural organization, and which answers questions about the situation based only on the contingent knowledge structure.

Section two, **New Memory Models**, discusses the implications of the assumption that input information is always interpreted in terms of large structural units derived from experience. Daniel G. Bobrow and Donald A. Norman postulate active *schemata* in memory which refer to each other through use of *context-dependent descriptions,* and which respond both to input data and to hypotheses about structure. Benjamin J. Kuipers describes the concept of a *frame* as a structural organizing unit for data elements, and he discusses the use of these units in the context of a recognition system. Terry Winograd explores issues involved in the controversy on representing knowledge in declarative versus procedural form. Winograd uses the concept of a frame as a basis for the synthesis of the declarative and procedural approaches. The frame provides an organizing structure on which to attach both declarative and procedural information.

The third section, **Higher Level Structures,** focuses on the representation of plans, episodes, and stories within memory. David E. Rumelhart proposes a grammar for well-formed stories. His summarization rules for stories based on this grammar seem to provide reasonable predictions of human behavior. Roger C. Schank postulates that in understanding paragraphs, the reader fills in causal connections between propositions, and that such causally linked chains are the basis for most human memory organization. Robert P. Abelson defines a notation in which to describe the intended effects of plans, and to express the conditions necessary for achieving desired states.

The fourth section, **Semantic Knowledge in Understander Systems,** describes how knowledge has been used in existing systems. John Seely Brown and Richard R. Burton describe a system which uses multiple representations to achieve expertise in teaching a student about debugging electronic circuits. Bonnie Nash-Webber describes the role played by semantics in the understanding of continuous speech in a limited domain of discourse. Allan Collins, Eleanor H. Warnock, Nelleke Aiello, and Mark L. Miller describe a continuation of work on Jaime Carbonell's SCHOLAR system. They examine how humans use strategies to find reasonable answers to questions for which they do not have the knowledge to answer with certainty, and how people can be taught to reason this way.

Acknowledgments. We are grateful for the help of a large number of people who made the conference and this book possible. The conference participants, not all of whom are represented in this book, created an atmosphere in which interdisciplinary exploration became a joy. The people attending were:

From Bolt Beranek and Newman--Joe Becker, Rusty Bobrow, John Brown, Allan Collins, Bill Merriam, Bonnie Nash-Webber, Eleanor Warnock, and Bill Woods.

From Xerox Palo Alto Research Center--Dan Bobrow, Ron Kaplan, Sharon Kaufman, Julie Lustig, and Terry Winograd (also from Stanford University).

From the University of California, San Diego--Don Norman and Dave Rumelhart. From the University of Texas--Bob Simmons. From Yale University--Bob Abelson. From Uppsala University--Eric Sandewall.

Julie Lustig made all the arrangements for the conference at Pajaro Dunes, and was largely responsible for making it a comfortable atmosphere in which to discuss some very difficult technical issues. Carol Van Jepmond was responsible for typing, editing, and formatting the manuscripts to meet the specifications of the systems used in the production of this book. It is thanks to her skill and effort that the book looks as beautiful as it does. June Stein did the final copy editing, made general corrections, and gave many valuable suggestions on format and layout.

Photo-ready copy was produced with the aid of experimental formatting, illustration, and printing systems built at the Xerox Palo Alto Research Center. We would like to thank Matt Heiler, Ron Kaplan, Ben Kuipers, William Newman, Ron Rider, Bob Sproull, and Larry Tesler for their help in making photo-ready production of this book possible. We are grateful to the Computer Science Laboratory of the Xerox Palo Alto Research Center for making available the experimental facilities and for its continuing support.

Representation and Understanding

Studies in Cognitive Science

DIMENSIONS OF REPRESENTATION

Daniel G. Bobrow
Xerox Palo Alto Research Center
Palo Alto, California

I. INTRODUCTION

Workers in cognitive science have worried about what people know, and how to represent such knowledge within a theory.[1] Psychologists such as Paivio (1974) and Pylyshyn (1973) have argued, for example, over two alternative forms for visual memory in humans. The style of their arguments, which we return to at the end of this chapter, is to set up opposing characterizations and to argue about which one has more "natural" properties with respect to observed phenomena.

I claim that a more appropriate way of discussing the issues involved is to characterize each representation in terms of how it answers certain questions posed in this chapter. I pose these questions in terms of a set of design issues one would face in designing or analyzing an *understander system*--a system (human or computer) which could use the knowledge to achieve some goal. I propose a framework for viewing the problems of representation. In this framework each of the design issues defines a dimension of representation--a relatively independent way of looking at representations.

In this chapter I emphasize the structure of alternative solutions to the design issues. I illustrate the design options through three specific representations described here, and in examples from the literature and other chapters in this book. By considering representations along the separate dimensions, it often becomes apparent that a pair of seemingly disparate representations differ in very few significant features.

A. Representation and Mapping

I propose here a framework where representations are viewed as the result of a selective mapping of aspects of the world. Suppose we take a "snapshot" of the world in a particular state at some instant in time. Call this state *world-state-1*. Through some mapping M, a representation

---

[1] In the preface we describe cognitive science as a new field containing elements from psychology, linguistics, computer science, philosophy, education, and artificial intelligence.

(call it *knowledge-state-1*) is created which corresponds to world-state-1. This corresponds with world-state-1 in the sense that an understander has the alternative of answering questions about world-state-1 by directly observing the world state or by questioning the corresponding knowledge state (see Fig. 1).

$$\begin{tikzcd}
\text{World-state} \arrow[r, "M"'] \arrow[d, "\text{Make Observation}"] & \text{Knowledge-state} \arrow[d, "\text{Ask Question}"'] \\
\text{Result} \arrow[r, "M"] & \text{Answer}
\end{tikzcd}$$

Fig. 1. Mapping between world and knowledge states. Answering questions should correspond to making observations and mapping the result.

This implies, of course, the existence of a world-observation and knowledge-question function correspondence; simplicity of the mapping M, and simplicity of representing particular knowledge and questions must be considered in comparing representations of a world-state.

$$\begin{tikzcd}
\text{World-state-1} \arrow[r, "M"'] \arrow[d, "\text{Action}"] & \text{Knowledge-state-1} \arrow[d, "\text{Model Operation}"'] \\
\text{World-state-2} \arrow[r, "M"] & \text{Knowledge-state-2}
\end{tikzcd}$$

Fig. 2. A world-state can be changed by an action. An equivalent model operation should produce a change in the knowledge-state which corresponds to the changed world-state.

The world at a particular instant is static, and all the facts about the world reflect a single consistent state. If we now augment our simple view, and allow *actions* which change some properties of the world, then we must have *model operations* which make corresponding changes in the knowledge state. For a model to be consistent, an updated world-state-2 must correspond to the updated knowledge-state-2 (see Fig. 2).

In terms of this simple framework for viewing representation, we can now look at a number of different design issues. I pose these as a series of questions to be asked about any mapping and the resulting representation of the world:[2]

> *Domain and Range:* What is being represented? How do objects and relationships in the world correspond to units and relations in the model?
>
> *Operational Correspondence:* In what ways do the operations in the representation correspond to actions in the world?
>
> *Process of Mapping:* How can knowledge in the system be used in the process of mapping?
>
> *Inference:* How can facts be added to the knowledge state without further input from the world?
>
> *Access:* How are units and structures linked to provide access to appropriate facts?
>
> *Matching:* How are two structures compared for equality and similarity?
>
> *Self-awareness:* What knowledge does a system have explicitly about its own structure and operation?

B. Three Simple Visual Representations

To illustrate some options concretely on certain dimensions, I use three different specific representations for the same simple domain--two-dimensional black and white scenes. I describe how each represents a visual scene which contains a square rotated so that one diagonal is horizontal.

---

[2] In constructing this list of questions, I have been influenced by the dimensional analysis used by Moore & Newell (1973) in describing their system MERLIN.

Binary Matrix: Fig. 3a shows a two-dimensional binary matrix represention (MATRIX) of the spatial layout. A "1" is inserted in the matrix wherever the light intensity in the scene is below some threshold, and a "0" otherwise.

```
0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0
0 0 0 1 1 1 0 0 0 0
0 0 1 1 1 1 1 0 0 0
0 1 1 1 1 1 1 1 0 0
0 0 1 1 1 1 1 0 0 0
0 0 0 1 1 1 0 0 0 0
0 0 0 0 1 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
```

Fig. 3a. A binary matrix visual representation.
A 1 indicates a light intensity below a certain level.

A collection of connected 1s determines an object, with transitions between spaces containing 1s and 0s indicating the contours of an object.

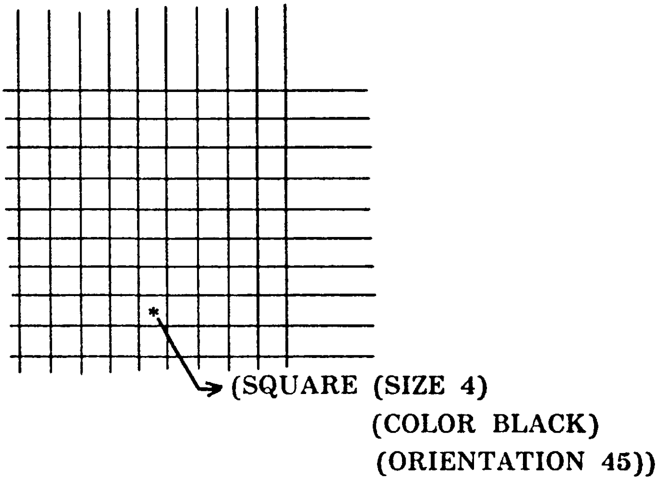


Fig. 3b. A Grid-positioned/feature oriented representation.

Grid-positioned feature: Fig. 3b shows what I call a grid-positioned feature representation (GRID) for a scene. An object is represented by a unit which specifies a set of features. The structure shown is of type SQUARE, with features specifying the size, color, and orientation of the square. The definition of SQUARE is not shown; it can be obtained given the type specification. The grid is used to locate objects in a scene. From a point on the grid corresponding to the location of the leftmost lowest point of the object, there is a link to the unit representing the object.

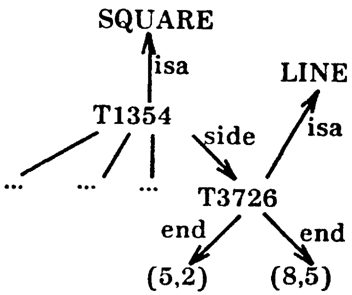


Fig. 3c. A semantic network representation.

Semantic Network: Fig. 3c shows a portion of a semantic network represention (NET) of the same visual scene. The units shown are a token of a square, tokens of sides of the square, and some number pairs representing the endpoints of the sides of the square. Only one of the endpoint sets are shown. Labelled links from one unit to another show the relations between the units.

II. DOMAIN AND RANGE

A. Units and Relations

The choice of units and structures reflects how one views the world one is modeling. A unit is something which can be used without knowing anything about its internal structure. This does not imply necessarily that it must not have any internal structure, just that there are occasions of use (e.g., inference rules) in which the existence of the unit is sufficient. In addition to its identity, response to a unit may be a function of its position in a larger containing structure, or special relation to other units, or to its internal form.

In choosing a representation for a particular world, some relationships can be stored explicitly and others need not be. For example, the size of the square is implicit in the matrix in MATRIX, as is the position of the square in GRID. These are reflections of what Hayes (1974) describes as the similarity between the medium of representation and the world, at least with respect to the relations being modeled.

Not all relations in a representation fully determine a portion of the world. For example, the relative position of two objects (A is left of B) may be implicit in locations represented in the model. Alternatively, this fact may be explicitly represented, with perhaps no absolute location information for either unit. How such "vague" predicates and partial information about the world are handled is an important characteristic of a representation. (See Chapter 2 by Woods for a more complete discussion of problems of vague predicates.)

B. Exhaustiveness

A representation is exhaustive with respect to a property if for any object, if it has that property, that fact is stored explicitly. Not only does the model represent the truth, it represents the whole truth. Thus in an exhaustive representation of the objects present on the surface of a table top, any object not explicitly noted as on top is not there, and if no object is associated with a location, then that location is guaranteed to be empty. In an exhaustive representation all objects that exist are represented explicitly, and any universal proposition can be verified by testing all elements of this set of objects. Exhaustiveness is a second aspect of what Hayes (1974) refers to as similarity of structure of the medium.

One way for a visual representation to maintain the property of exhaustiveness is for the mapping to have the property of extracting a uniform degree of detail. An aerial photographer does map terrain this way whereas a cartographer may not. In the photo, it is guaranteed that no object within the field of view and larger than the resolution of the lens will be missing. The whim of the mapmaker determines the objects and features represented on a map. MATRIX is by nature exhaustive; GRID and NET can be made so by design. Human visual memory does not seem to have this property of uniform extraction of detail, or of exhaustiveness.

C. Verbal Mediation

Instead of mapping the world directly, people have constructed systems which map the world using natural language descriptions. There are many issues involved in building adequate representations of English language statements. Woods (Chapter 2) points, for example, to the subtle problems involved in representing relative clauses and verbal restrictions within a semantic network. In this chapter, I focus only on the issues involving selection of basic units to represent linguistic information.

Word-Senses: An obvious choice for a unit is a single word. The relations chosen often are the case relations for verbs (Fillmore, 1968). This simplifies the mapping process by focusing on the obvious units and their grammatical relationships. A problem with this choice is that words are often ambiguous. Some systems finesse this issue by assuming that each word will have only one meaning within the domain of interest. Other systems face the issue by allowing individual words a number of different senses.

Several problems must be considered in systems which use word-senses. There is the obvious potential error in ignoring concepts for which there are no single words (or for which the user knows none), such as a single word to describe "those small orange cones used to divert traffic". A word-sense system must allow compound constructs to be used as well as atomic units.