COMPUTER PATTERN RECOGNITION AND HUMAN CONCEPT FORMATION

By

IVAN BRIAN KESSEL

Bachelor of Commerce

University of the Witwatersrand

Johannesburg, South Africa

1963

Submitted to the faculty of the Graduate College of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF BUSINESS ADMINISTRATION
July, 1966

COMPUTER PATTERN RECOGNITION AND

HUMAN CONCEPT FORMATION

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

PREFACE

This paper is a study of human concept formation and machine pattern recognition. The first three chapters deal with definition of terms, a review of the research being undertaken in the area of computer simulation, and the purpose of such research. In the following chapters a machine model of pattern recognition is constructed and a model of human concept formation is afterwards based on this. The hypotheses set forth in the human model are empirically tested using human subjects and the findings of this experiment conclude the report.

I wish to acknowledge my indebtedness to my adviser, Dr. W. A. Meinhart, Associate Professor of Management, for his guidance and for his assistance in the preparation of this report. The valuable suggestions of Dr. L. T. Brown of the Psychology Department, Oklahoma State University, pertaining to the preparation of the psychological test, and the willing assistance of Mr. R. R. Weeks and Dr. C. Roman in the procurement of subjects for this test are also greatly appreciated.

I am grateful to Mrs. Sara Millican for her excellent typing of this paper.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Learning is an arboriform stratification of guesses about the world. New guesses or new concepts (new patterns) are essentially simple combinations of words which form the patterns that have already been learned or that are inherent. In a large sense, learning is the evolution of patterns and the first steps are the hardest.[1]

Patterns are amazingly complicated things to come to grips with. Even a consensus of what we mean by the word "pattern" is lacking, but a growing number of people are beginning to feel that many of the central problems of behavior, intelligence and information processing are problems that involve patterns.[2]

The study of patterns is the study of complexes--of structures, interactions, grammars and syndromes. We find patterns not only in visual and other sensory stimuli but also in language and other symbols, in assessments and diagnoses and, in general, in descriptions of any complex domain.

A pattern is equivalent to a set of rules for recognizing it, and the pattern is determined by these rules rather than vice versa; however, the redundancy of the world is not always tailor-made to the lan-

---

[1] Oliver G. Selfridge, "Pattern Recognition and Learning," Information Theory, ed. Colin Cherry (New York, 1956), p. 349.

[2] Leonard Uhr, Pattern Recognition (New York, 1966), p. vii.

guage in which the rules must be stated.[3]

Now what about the term "recognition"? Hake (1957) says, "About the process of recognition itself little is known. We know little about the stimulus conditions which influence accuracy of recognition or evoke recognition responses."[4] However, to be a bit more specific, recognition can refer to tasks in which the subject indicates whether or not he has seen the stimulus before.

From the above it should be realized that there is no clear-cut understanding, let alone definition of pattern recognition *per se*. However, it can be concluded that by pattern recognition is meant the extraction of the significant features from a background of irrelevant detail, or classifying a set of data into the learned categories whereby "learning" is meant acquiring feasible operational definitions of the categories. Thus learning and pattern recognition are complementary.[5]

The word "significant" in the above is very important. Significance is a function of first, context, and second, experience. Now, of course, context is a function of experience, and the shape of a pattern is recognized with the help of its context. But more than this, experience alone affects the kind of thing we regard as significant.[6] In this way, the whole process of pattern recognition is inevitably tied up with

---

[3] Selfridge, p. 345.

[4] Leonard Uhr, "The Development of Perception and Language Simulated Models," *Computer Simulation of Personality*, eds. S. Tomkins and S. Messick (New York, 1963), p. 232.

[5] Selfridge, p. 345.

[6] Oliver G. Selfridge, "Pattern Recognition and Modern Computers," *Proceedings of the Western Joint Computer Conference* (Los Angeles, 1955), pp. 91-93.

ways of determining significance. This, I might add, is normally the distinction made between man and machine: that man can learn by experience to extract and deal with the significant things, and a machine cannot. Selfridge (1955) does not believe it is a valid distinction, since the machine must extract the essential or significant characters of the configuration in order to recognize the pattern to which it belongs and, this involves a certain amount of learning or experience on the part of the machine program.

Human readers rely heavily on the redundancy of the language, and therefore on the use of context. This operates, in the case of script, at the letter, word, and sentence level, and frequently beyond. In addition, the individual characters of a given script style are evaluated and used in the reading process. Any recognition system should therefore take account of context though necessarily at a simple level.[7]

Pattern recognizers are concerned with discriminating one out of a large number of specified signals and such recognizers are designed to function in a context where the set of alternatives is known--the letters of the alphabet, the phonemes of English speech--and the machine's job is to categorize each particular input as one of the known alternatives.[8]

There may well be an infinite number of different particular examples of a chair or face. The different examples may be similar to one another in certain well-defined respects, as in the case of an infinite set of straight lines, or fairly similar to one another in certain

---

[7] L. S. Frishkopf and L. D. Harmon, "Machine Reading of Cursive Script," Information Theory, ed. Colin Cherry (Washington, 1961), p. 300.

[8] Bert F. Green, Digital Computers in Research (New York, 1963), p. 210.

ill-defined respects, as in the case of the capital letter A as written by different people.

The interesting problems come when the procedure that tells the recognizer how to decide whether a particular example is a member of one or another of the alternative possible classes is not known. That is, the interesting problem is a problem of induction hypothesis--formation, learning and concept attainment. Almost all natural patterns are not even describable. The pattern recognizer must be able to learn; in fact, its basic job is to learn general concepts on the basis of specific examples--to perform inductions and even to form the hypotheses that the inductive evidence provides from specific examples.[9]

Pattern recognition studies have often been closely associated with concept formation experiments, and in most respects they are alike. The pre-code of the stimuli is the main difference in these experiments. Pattern recognizers usually accept stimuli coded into projections on a grid. The result is a string of bits, each bit representing the presence or absence of illumination on some part of the grid.

It is difficult to define what a concept is if one wants it to be more than mere generalization learning. One possible definition is the following. A person has a concept if he has disposition on the basis of which he can make nominal classification statements or responses. It is assumed that the disposition is learned from a number of instances which vary among themselves; and that the response can also be made to instances other than those contained in the set on which the concept was learned. It is also assumed that the classifications response is

---

[9] Uhr, Pattern Recognition, p. 3.

not the only possible one. If all this is included in the definition, then "having a concept" implies that the one has more than one concept, it implies a conceptual system.[10]

In perception, we are becoming aware that we not only need to keep track of the probability of correct recognition but should also pay particular attention to the kinds of errors which are made. Perceptual learning is an unsolved problem, in both psychology and "artificial intelligence," and it is also a reasonable place to begin when one hopes to develop "higher" processes in a more complex system.[11]

Such attention will become increasingly important when someone makes a serious attempt to simulate human pattern recognition on a computer. The experimentation and theory building and modification which is being undertaken today is rapidly building what appears to be the first relatively firm and meaningful theoretical structure for the science of higher mental processes, i.e., for pattern or form perception.[12]

This paper will approach the problem of pattern recognition and concept formation from two sides. First, we shall examine a pattern recognition model designed for a machine and after this build a human model based on this computer model and, subsequently, test the hypothesis set forth. These hypotheses are that people recognize patterns in a manner similar to that suggested by the computer model, that is, on the

---

[10] John P. Van de Geer and Joseph M. Jaspars, "Cognitive Functions," Annual Review of Psychology, 1966, eds. Farnsworth, McNemar and McNemar (1966), p. 149.

[11] Uhr, Computer Simulation of Personality, p. 265.

[12] Leonard Uhr and Charles Vossler, "A Pattern Recognition Program that Generates, Evaluates, and Adjusts its Own Operators," Computers and Thought, eds. Feigenbaum and Feldman (New York, 1963), p. 268.

basis of probability, and the combination of distinguishable features in the pattern. This will be tested with people as subjects using the patterns adopted in the human model in Chapter V. However, before such an attempt is made the following chapter will discuss current research in this area, and the chapter after this will examine the purpose and the implications of this research.

CHAPTER II

REVIEW OF THE RESEARCH AND THE LITERATURE

The large body of pattern recognition research that has arisen in the past ten years in the inter-disciplinary area between psychology, psychiatry, mathematics, engineering, and physiology that is variously called "cybernetics," "artificial intelligence," "system sciences," "communication sciences," and "information-processing sciences," among other names, has been largely concerned with a particular simplified version of the general problem of perception and pattern recognition.[13]

The bulk of the research has been on recognition of letters of the alphabet and, occasionally, other visual patterns. Most of the rest of the work has been on the recognition of spoken words or phonemes. A scattering of studies has examined recognition of other arrays including Morse code and diagnostic symptoms.

Virtually all of this research handles the problem of naming a static, isolated matrix whose primitive symbols are discrete and clearly distinguishable. This primitive set of symbols usually contains only the two values "black" and "white"--("0" and "1") in the case of visual patterns or a small range of intensities in the case of auditory patterns.[14]

Several models for pattern or shape recognition have been described

---

[13] Uhr, Pattern Recognition, p. 366.

[14] Ibid.

without using a programmed computer as the actual physical embodiment of the pattern recognizing model. Most of these were prior to the arrival of the computer, because computers have been sufficiently large to embody such a model only since about 1957. Some of these have served as sources for programmed models either peripherally--Hebb (1949), and McCulloch and Pitts (1947)--or quite directly, as in the case of Deutche's (1955) papers.

A great many models for pattern recognition have been programmed--mostly since 1959. They embody a wide range of approaches to the problem--from detailed idealized templates and other representations of images, and from statistical sampling, factor analytic, information theory, to Gestalt-like integrative attempts to isolate and examine meaningful properties and characteristics and the interrelations between them.

All these studies were focused on the recognition of letters of the alphabet, and often no explicit statement was made as to the relevance of the program as a model of human recognition. This has led many people to assume that there is a rigid line between simulated modeling of psychological processes on the one hand and "artificial intelligence" on the other. But this is really a very superficial dichotomy based chiefly on whether the individual and his society choose to call him a "behavioral scientist" or an "engineer," "mathematician" or "computer scientist."[15]

Attneave and Arnoult (1956) were among the first to pose the problem of the recognition of sensory patterns in terms clear and precise enough to point toward experimental examination. Their paper was written just

---

[15] Ibid., p. 291.

as the results of the first computer programs were being published and they were among the first to recognize the significance of such work for psychology. Although they did not specify completely and program their own ideas on the subject, they gave in this paper a number of stimulating suggestions for mechanisms for pattern recognition.[16]

Actually some of the literature on cognition, specially as it relates to concept formation, reorganization of information, and the discovery and learning of structures (for example, grammar) is at least pertinent to the problem of pattern recognition as is much of the literature on perception. Indeed, it turns out that the concept learning experiments of Hovland (1960), Bruner (1956) and others, when modeled via digital computer programs by Kochen (1961), Hunt and Denton (1962) are formally identical to pattern recognition. This type of concept formation asks the subject to examine strings of about 5-8 symbols that can really be thought of as strings of binary numbers, made up of the two symbols "0" and "1."

Pattern recognition, on the other hand, asks the subjects to do exactly the same thing, except that the string of symbols is much larger--on the order of hundreds or thousands. The symbols are arranged in a matrix, and the amount of detail that the subject must notice is correspondingly smaller. Consequently the models of Uhr and Vossler (1961), Roberts (1960), and Selfridge (1959), to name only a few, are sufficiently similar to be satisfactory representatives of concept formation models. Here, in fact, is a very striking example of the ability

---

[16] Fred Attneave and Malcolm D. Arnoult, "The Quantitative Study of Shape and Pattern Recognition," Pattern Recognition, ed. Leonard Uhr (New York, 1966), pp. 123-142.

of the precise computer model to pinpoint similarities between disparate processes, and to allow for models of much greater generality than psychology has been able to achieve in the past.

Generally speaking, most of the simulations that have been made can be put in one of four categories: (1) neural nets, (2) pattern recognizers, (3) problem solvers, and (4) language processers. Pattern recognition itself can be further classified by the terms template matching, analytic, and random.

The problem of reading printed characters is a clear-cut instance of a situation in which the classification is based ultimately on a fixed set of "prototypes." The variables involved are either distortions--systematic changes in size, position, orientation, or "noise" distortions--blurring, grain, low contrast and so on. If this noise is not too severe, we may be able to manage the identification by a normalization and template-matching process.[17]

The first step is to remove any differences attributed to size and position--this is called normalization. One way to do this is to transform the figure to obtain a certain fixed center of gravity and a unit second central moment.

Once normalized the figure can be compared with templates for the prototypes and by means of some measure of matching the best fitting template can be chosen.

The above scheme with its normalization and direct comparison and matching criterion is just too limited in conception to be of much use in more difficult problems. The template system in fact has negligible

---

[17] Marvin Minsky, "Steps Towards Artificial Intelligence," Computers and Thought, eds. Feigenbaum and Feldman (New York, 1963), p. 413.

descriptive power.[18]

The name "analytic" has been given to methods for perception that show the most immediate promise. These might best be described as the methods for abstracting from combinations of basic units particular features or qualities of importance, for identifying rather more abstract and complex things like edges, ends, curves, angles and slopes--in general for analyzing rather than matching.[19]

In the experiments of Uhr the pattern recognizer follows line segments while under the control of an assessing sub-routine, until it has identified a complete element as one of the 9-bits worth of possible elements. It then identifies the next element, stores the relative location at which elements touch, and continues until it has completed the figure. This description by elements seems almost a "natural" way of describing figures--specially man-made figures such as letters.[20] For example an "A" equals a vertical left and a vertical right touching at the top, with a horizontal line joining their middles. A "B" is a vertical left with top and bottom curved loops, both closed. An "R" differs from a "B" in that there is no bottom curved loop. Even handwriting when reduced to this sort of element should give standard characterization.

The letters to be "recognized" are stored in the computer's memory, along with their characterization lists--literally, what they look like. The computer first searches these lists to find out what to look for--

---

[18] Ibid., p. 415.

[19] Leonard Uhr, "Intelligence in Computers: The Psychology of Perception in People and in Machines," Behavioral Science, April 1961, p. 179.

[20] Ibid., p. 180.

what types of elements to expect--in the figure it is trying to identify. The lists determine the directions in which the computer will search.

Selfridge in 1955, together with Dineen, published a series of papers showing how the very simple functionings such as those that find edges, angles and connectivity could be programmed on the computer. In 1958 Selfridge published another paper giving a general outline for the organization of a parallel type of model.[21] This stimulated several programs that gave specific embodiment of such a model, that is, the Morse code program of Gold and Doyle's (1959) letter recognition program. Also the recent work of Neisser (1959) on parallel versus serial processing in humans and of Lettvin, Maturana, McCulloch and Pitts (1959) on detection functioning in frogs were influenced by Selfridge's model.

This program which Selfridge calls Pandemonium uses the property approach to pattern recognition. The property-list scheme, while not a very general form of pattern recognition does lend itself nicely to some rather straightforward inference schemes. One can treat the separate properties as more or less independent evidence for the defined categories.

A "property" is defined as a two-valued function which divides figures into two classes; a figure is said to have or not to have the property according to whether the value of the function is "1" or "0."[22] Each cognitive "demon," as Selfridge calls it, is connected to several computational demons that examine the properties of the input pattern. Each demon examines the data to determine whether his particular property

[21] Oliver G. Selfridge, "Pandemonium: A Paradigm for Learning," Proceedings of a Symposium on Mechanization of Thought Processes (London, 1959), pp. 513-526.

[22] Minsky, p. 415.

exists or not. If it does, that demon produces an output to all the cognitive demons to which it is connected. With all the demons reporting at the same time pandemonium results, hence the name.

However, the pandemonium model is very general with no feedback since all the properties are examined at the same time, in parallel. This approach can be contrasted with the sequential system of decision, in which the output of one demon determines which other demons get to see the input next.

> It seems certain that an adequate model for human pattern recognition will have to include both parallel and sequential processing with feedback at several levels so that any decision initially made can be undone if further results clearly show that an error has been made.[23]

A variation of the above approaches is the random net system. A random net is a large set of similar and simply-acting elements whose attributes and interactive connections may be randomly established.[24] Some of the units are usually designated input, and some output units. The units themselves are termed neurons or cells. The underlying interest in random nets is the belief that if the "right" responses are rewarded by some "reinforcement" and "wrong" ones discouraged, then the net as a whole will organize itself so as to tend to make only correct responses, even when they are very complicated.

The program of Uhr and Vossler (1963) is based on the above and it attempts to make as much use as possible of methods for discovery and learning by trying to recognize specific examples of pattern sets.

---

[23] Green, p. 211.

[24] Marvin Minsky and Oliver Selfridge, "Learning and Random Nets," Information Theory, ed. Colin Cherry (Washington, 1961), p. 335.

It "learns" to do this through experience with specific examples, along with their correct names; and then it is tested on its ability to recognize different unknown examples that are given unnamed. The measurers, or operators, that the program can discover and develop are restricted to those that could be performed by nets of neurons and by higher-level combinations of these nets. The program attempts to develop its own set of operators, rather than having these built-in. The program learns by accumulating weightings and success counts for each measurer, developing higher level measurers and throwing away bad measurers. Because this model did not have a specific set of functions given to it a priori it was able to learn to recognize a large variety of different types of patterns, including pictures of simple objects, stick figures, cartoons and photographs of faces.

Such learning methods have also been investigated by Kamentsky and Liu (1963) and Prather (1964) and in the related areas of concept formation and verbal learning by Kochen (1961), Hunt (1962) and Feigenbaum (1959).[25]

Two other researchers, Bledsoe and Browning (1959), embodied a conceptually very simple, logical scheme in a pattern recognition program that exhibited a considerable amount of power and, at least in certain limited senses, certain abilities to learn to recognize, no matter what type of pattern. Pattern recognition of the "whole," that is, Gestalt recognition forms the basis of their program.[26] The very essence of the

---

[25] Uhr, Pattern Recognition, p. 294.

[26] W. W. Bledsoe and I. Browning, "Pattern Recognition by Machine," Proceedings of the Eastern Joint Computer Conference (Boston, 1959), pp. 225-232.

concept of a "pattern" is that it cannot be chopped up, that it is a structure of individual parts whose meaning lies not at all in these separate parts but rather in the fact of their structure.

A program of a very different sort was presented by Grimsdale, Sumner, Tunis and Kilburn (1959).[27] Their model assumes that a pattern is a collection of basic strokes (essentially lines of different length, slope and curvature), and attempts to build up a description of any pattern presented in terms of these basics. They are using characterizers quite similar to those that would be suggested by intuition and common sense and, indeed, have recently been shown to exist in animals, in the experiments of Hubel and Wiesel (1962), and of Lettvin, Maturana, McCulloch and Pitts (1959). Several other programs have been coded or suggested to embody the same general type of model--Bomba (1959), Prather and Uhr (1964), Marrill (1963), Eden and Halle (1961), and Narasimhan (1963).

Roberts (1960) started his research in pattern recognition by examining the ability of the "Perceptron" of Rosenblatt (1958). Perceptron is a class name for a group of pattern recognition machines which can "learn" to discriminate several categories. The Perceptron has three distinguishing sets of functional units: (1) sensory, (2) association, (3) response units. It is first and foremost a brain model, and the program is mainly concerned with investigating the physical structure and neuro-dynamic principles which underlie natural intelligence.[28] It

---

[27] R. L. Grimsdale et al., "A System for the Automatic Recognition of Patterns," Proceedings of the Institute of Electrical Engineers, Vol. 106, Part B, No. 26, March 1959, pp. 210-221.

[28] Alice Mary Hilton, Logic, Computing Machines and Automation (Washington, 1963), p. 38.

is a very general, almost completely unstructured, random collection of elements, similar in some respects to neurons or simple switches.[29]

Theoretical investigation has resulted in some existence proofs as to the possibility that there would be some particular random configuration among the large sets of possible "perceptrons" that would, given enough training, become capable of recognizing interestingly difficult patterns.[30] Roberts found in his actual computer runs that this was a very slow task. Nevertheless, using the same input as Doyle he achieved 94 percent correct identification.

---

[29]Uhr, Pattern Recognition, p. 293.

[30]Ibid.

CHAPTER III

THE IMPORTANCE OF THE CURRENT RESEARCH

The notion that a digital computer program could be a model of human behavior is a natural outgrowth of a field of engineering endeavor that has come to be called "artificial intelligence."[31] However, "intelligence" is a slippery concept.

Minsky has offered the following definitions of intelligence: "You regard an action as intelligent until you understand it. In explaining, you explain away."[32] "It denotes little more than the complex of performances which we happen to respect, but do not fully understand."[33]

Simulation of learning is one of the most interesting potential applications of computer simulation techniques, since the ability to learn is one of the clear-cut differences between human and machine performance.[34] As indicated in the previous chapter, a number of different types of learning procedures are being simulated, one of which is the type of learning involved in recognizing patterns imbedded in a complex stimulus. It seems a simple thing for a human to respond to a

---

[31]Bert F. Green, "Computer Models of Cognitive Processes," Psychometrika, March 1961, p. 86.

[32]Paul Armer, "Attitudes Toward Intelligent Machines," Computers and Thought, eds. Feigenbaum and Feldman (New York, 1963), p. 391.

[33]Minsky, Computers and Thought, p. 447.

[34]Carl I. Hovland, "Computer Simulation of Thinking," The American Psychologist, November 1960, p. 689.

triangle as a triangle, whether it is large or small, short or tall, tilted or upright, and to distinguish it clearly from a square. But to specify rigorously the criteria in such a way that a machine can learn to recognize it invariably is quite a job, and the difficulty clearly indicates that there is a lot we do not understand about the phenomenon, even at the human level where we take the procedure for granted.[35]

Extensive tool building is a necessary prerequisite for developing programs which actually are intended to simulate human behavior in a precise way. Most of the work done in the area of computers is allocated to more tool building. In one sense, then, computers serve the same purpose as mathematical models--in both cases we have conceptual tools which do not permit any vagueness and mercilessly bring to light any weakness of conceptualization which otherwise might have passed unnoticed.[36]

Today psychology has amassed a great number of particular facts as to the interactions of the many factors involved in even the simplest perceptual acts. But it has not developed anything in the way of a coherent theory of how the crucial recognition, toward which the entire perceptual process leads, actually takes place.[37]

Because of the complexity of their problems, the psychologist and the neurophysiologist have studied the levels of behavior and neuron pathways quite separately and have made few serious attempts to bridge the gap between them. When psychological processes are recast into the language of information processing as it flows through a computer model,

---

[35] Ibid.

[36] Van de Geer and Jaspars, p. 163.

[37] Uhr, Pattern Recognition, p. 366.

an important link has been made. For the computer, from one point of view, is basically a switching network, and a switching network is the most degenerate form of nerve net, that is, the process by which a computer simulates a transformation of information--no matter how complex--boils down to a tremendously large set of simple, switch-like steps. Nerve nets in the brain almost certainly do something far more complex than this, but the particular complex processing they perform can themselves be modeled as a set of switch-like processes. If we assume that the switching net is a fair first model for a nerve net, there will be a fairly straightforward interpretation of a computer program for the psychological process of pattern recognition at the physiological level of nerve nets.[38]

The computer is not restricted to a simple nerve-net interpretation. Rather the computer is a very general device that is capable of manipulating anything that is described or presented to it. So any statement as to how to go about pattern recognition--i.e., any theory of recognition at either the psychological or physiological level--could be written as a computer program.[39]

Computer simulated models are merely theoretical models of the classical sort. The computer is simply the vehicle on which the model is written. The program must be as completely and precisely stated as a mathematical theorem. This given them the added virtues of clarity and consistency--these are two virtues that are far too frequently absent in more traditional theorizing in psychology and the behavioral

---

[38]Ibid., p. 4.

[39]Ibid.

sciences.[40]

The value of analysis for the psychologist is that often it helps to clarify his thinking. If he can express the problem in terms of a machine model, then he feels that he has defined his terms scientifically, and has eliminated any probable mysticism or magic from his discussion. Since psychologists are often unable to make a precise mathematical statement they resort to the next best approach to precision--the conceptual model. A model is certainly superior to a purely verbal argument, although inferior to a complete mathematical theory.

If we are eventually to understand the capability of high organisms for perceptual recognition, generalization, recalling and thinking, we must have the answers to three fundamental questions:[41] (1) How is information about the physical world sensed by the biological system? (2) In what form is information stored, or remembered? (3) How does information contained in storage, or in memory, influence recognition and behavior?

The first question, which is in the area of sensory psychology, is the only one for which appreciable understanding has been achieved. It is, therefore, of the utmost importance for the theorists to utilize the computer program to gain further insight in the latter two unanswered questions. The precisely specified and programmed model can also throw a great deal of light on what we really must mean by some of the ill-formulated concepts which we use to paste together our verbal

---

[40] Ibid., p. 5.

[41] Frank Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," Psychological Review, November 1958, p. 386.

theorizing.[42]

Pattern recognition, along with most other higher mental processes, is surely complex. The computer program allows us to begin to build more complex models and complexity is necessary in a model of a complex phenomenon.

There are at least three reasons for studying pattern recognition:[43] (1) To improve our understanding of perception and conjunction. (2) To develop more interesting "artificial" systems for studying intelligence and learning. (3) To develop technologically practicable solutions to specific problems.

The important point is that the machine can be used as a subject when it is desirable to vary a factor that cannot be tampered with in people (for example, previous experience).

> The concept of comparing the behavior of man and machine in an n-dimensional continuum recognizes differences as well as similarities. A common argument against machine intelligence is that the brain is a living thing--the machine is not. In our continuum we simply recognize the dimension of living and note that machines and men occupy different positions in this dimension. It is hoped that the definition of research on "artificial intelligence" as an effort to push machines further out in the continuum of intelligent behavior will reduce some of the semantic difficulties surrounding discussion of such research.[44]

---

[42] Uhr, *Pattern Recognition*, p. 5.

[43] Leonard Uhr and Charles Vossler, "The Search to Recognize," *Optical Character Recognition*, eds. G. L. Fischer et al. (Washington, 1962), p. 326.

[44] Armer, p. 405.

CHAPTER IV

A MACHINE PATTERN RECOGNITION PROGRAM

A pattern recognition model has been tested by Worthie Doyle of the Lincoln Laboratory, Massachusetts Institute of Technology.[45] The procedure and method used in these experiments will form the basis for the human concept formation model in the following chapter. This pattern recognition scheme is intended to handle noisy and highly distorted data. The input data is subjected to a series of tests, after which a decision is reached. The scheme incorporates two notions--a decision is reserved until all test results are in, and discrimination is based on "learning" from real data. The input data was hand-printed English capitals, which can obviously be varied considerably. The program is in fact a Pandemonium one with a limited environment of 10 letters, the features and attributes of which are not memorized as combinations but rather as individual attributes. Our most important aim in each problem is the selection of useful items or tests. In each case no single item is very effective, while combined information from many items can be very effective.

Doyle split his data into two parts, using one part during the learning phase and the other part to test the results of the recognition

---

[45] Worthie Doyle, "Recognition of Sloppy, Hand-Printed Characters," Proceedings of the Western Joint Computer Conference (San Francisco, 1960), pp. 133-142.

process. There are six levels to the program: (1) input, (2) clean-up, (3) inspection of attributes, (4) comparison with learned-feature distribution, (5) computation of probabilities, and (6) decision.

The input stage involves the presentation of a letter to the machine which uses a photo cell device to scan the figure and transform it into a series of dots on a 32 X 32 matrix. All the cells through which the lines of the letter pass are filled in ("one" cells), the remainder remain blank ("zero" cells.)

Thirty attributes of each pattern were examined. Attributes included were vertical straight lines, horizontal lines, left, right, upper and lower concavities. Others were the maximum number of intersections of a vertical or horizontal straight line with the pattern. The computer draws the lines through every horizontal row in the matrix and recognizes intersections as sequences of "one" cells separated by sequences of "zero" cells. The maximum number of intersections being recorded, some attributes had two or three possible values while others had as many as ten or twenty.

During the learning phase, the computer determined the values of each of the thirty attributes and tallied these results in a set of distributions for that character.

In the table on the following page, the "vertical line" column records the maximum number of intersections with the horizontal line. Out of five D's, all five have a maximum of two intersections with the straight lines, only one B with three intersections and four with two intersections. The probability for each character is in parentheses.

The probability of a letter having a maximum number of two intersections being a D is 5/19 or .263, and a B is 4/19 or .211.

TABLE I

FEATURES AND PROBABILITIES FOR MACHINE RECOGNITION

Letter	Vertical Line 2 intersections	Vertical Line 3	Horizontal Line 1	Horizontal Line 2	Horizontal Line 3	Upper Concavity	Lower Concavity
B	4 (.211)	1 (1.00)	0	1 (.084)	4 (1.00)	0	0
D	5 (.263)	0	0	5 (.416)	0	0	0
V	5 (.263)	0	4 (1.00)	1 (.084)	0	5 (.50)	0
X	5 (.263)	0	0	5 (.416)	0	5 (.50)	5 (1.00)
Total	19	1	4	12	4	10	5

After all the samples have been fed into the machine during this learning stage, an "unknown" sample is projected onto the matrix. The same procedure is followed with horizontal and vertical lines being drawn to test for intersections and subsequently the 30 features.

The test for features is carried out; the maximum number of intersections are noted. Next the machine consults the property distribution to determine the likelihood of each letter; entering the estimated probabilities in a table. Then the total probabilities are computed for each letter. The letter with the highest probability is decided upon.

Doyle's program achieved about 87 percent success. Human identification of the same samples yielded a slightly higher level of success, 97 percent.

We might regard this result with a certain amount of apprehension, but if we compare the computer's ability to recognize unknown patterns with a person who after being totally blind all his life has his sight restored, then the machine's achievement is all the more resounding. It has been found that such persons have experienced tremendous difficulty recognizing relatively simple patterns such as the different features on two faces. We must regard the computer as being in a very similar position when initially confronted with unknown data.

Doyle's program can incorporate any attributes that will provide additional discriminations.

> The program is based upon a catch-as-catch-can philosophy. Nevertheless, this does not rule it out as a basis for a model of perception. Very likely human perceivers are able to use whatever information they can get about a pattern in a catch-as-catch-can fashion.[46]

---

[46] Green, Digital Computers in Research, p. 213.

CHAPTER V

A MODEL OF HUMAN CONCEPT FORMATION

Before attempting to construct a human model for concept formation our approach needs clarification, and terms need re-defining. Concepts can either be static or dynamic. We shall focus our attention on the static concepts, since these seem to be the earliest concepts formed in the human.[47] In other words we are interested in concrete rather than abstract concepts.

Although for a realistic simulation of typical concept learning we should include situations where there is characteristic human fallibility in memory, we shall avoid such a complication if for no other reason than for the absence of it in Doyle's program.[48]

To agree with our previous approach to machine concept formation it would be appropriate to start off with a human brain without any formed concepts, that is, a newborn child. When a child is born he comes into the world without any idea of what will confront him or what the "picture" before him will look like. Owing to a lack of communication, all concepts formed by the child will have to be necessarily internal. He will have to do all the "thinking for himself." This is the learning stage. He will form his own "picture" of a certain pattern and will associate cer-

---

[47] B. Berelson and G. Steiner, Human Behavior (New York, 1964), p. 198.

[48] Hugh E. Cahill and Carl I. Hovland, "The Role of Memory in Acquisition of Concepts," Journal of Experimental Psychology, March 1960, p. 137.

tain features with this pattern, until he is motivated to change this relationship.

"There are two important psychological processes operating in learning concepts--generalization and discrimination. Those factors that encourage discrimination facilitate concept formation."[49] Until a person is motivated to discriminate, he shall not be able to form any concepts. Generalization tends to restrict the person's ability to extend his range of concepts. For example a child having been conditioned to fear a white rat will generalize when confronted with a white cloth. Once he is able to discriminate between the two on the basis of some feature, a white cloth will no longer frighten him.

Let us now take the case of the child faced with a dog and a cat. No doubt the dog has been associated with certain "doglike" features such as the sound he makes. When confronted with a cat for the first time, the child will probably recognize it as a dog since it has very many "doglike" features--four legs, fur coat, two eyes, etc. All these characteristics build up a "doglike" picture in his mind. Confusion between dog and cat will result until he is able to find a specific feature common to only one--such as the sound either makes.

What if he is confronted with a third pattern--a bird? This has only two legs and no ears and is clearly neither dog nor cat since it has very few "doglike" or "catlike" features. He will associate this pattern with "birdness." He will now be able to recognize the difference between dog and bird or cat and bird, but he will still have trouble separating cat from dog--the only really distinguishing feature at this

[49]H. Kendler, Basic Psychology (New York, 1963), p. 25.

early stage being the noise each makes.

The manner in which these "pictures" are constructed can be summarized by the following:

> All information received by the brain must be coded into electric impulses in our sensory nerves, and then be reconstructed into a pattern in our brain. The "picture" in the brain consists of patterns of neural connections. The picture we reconstruct inside our heads of the things outside are the concepts of those things, which may or may not be accurate reconstructions.[50]

By the process of feature and characteristic inspection the child is able to build up concepts of his environment. There is constant comparison and checking of the features with information stored. If one feature is unrecognizable he will conclude that on the basis of all the other features it "must" be a pattern that he has already seen. As explained previously, if the child, having learned the features of a cat, is confronted with a rabbit he will conclude that on the basis of all the features observed it can only be a "funny looking" cat. The more cats seen and recognized as such, the higher will be the probability that the rabbit is indeed a cat. (Provided that no rabbits are seen.)

Expanding this probability notion further, that is, that probability plays an essential role in the recognition of patterns, we can create a model by assuming that at any fixed moment there are a certain number of animal features stored in his brain; for example, legs, eyes, ears, coat, shape, size, etc. Recognition of the pattern is thus based on these features.

For our purpose let us assume that he sees each of the following different animals five times: dog, cat, bird, rabbit and squirrel. A

---

[50]Kuhn, p. 36.

number of features can be associated with each. For simplicity we assume that there are only four features to be recognized that will account for the identification of the animal, that is, legs--either two or four; ears--either pointed or round; tail--either short or long; face--either flat or pointed.

Having been presented with each animal on five different occasions he is able to compute a probability list of how many times each of the four features occurred.

In Table II out of five dogs seen each had four legs; two had pointed ears and three had round ears; three had long tails and two had short tails; and four had flat faces while one had a pointed face. Each animal is classified in the above manner. A probability table is then made (figures in parentheses) on the basis of the occurrence of each feature in each animal, against the total number of features observed in all the animals. For example, the probability that an animal with pointed ears is a dog is 2/16 or .126; a rabbit 5/16 or .312, and so on. For each feature a probability is worked out. The above is the learning stage.

We are now in a position to present the child with a completely new animal (one of the above five) and he must decide what it is. This "unseen" animal has four legs, round ears, long tail and a flat face. On the basis of previous information and the comparison of features with estimated probabilities, Table III results. After the four-feature test has been carried out, each outcome is "looked up" in the "stored" probability table. The estimated probabilities are then entered into a table. The probability of each feature occurring is multiplied, and the animal with the highest result is chosen. This is necessary; otherwise if added

TABLE II

FEATURE AND PROBABILITY

Animal	LEGS		EARS		TAIL		FACE	
	2	4	Pointed	Round	Long	Short	Flat	Pointed
Dog	0	5 (.250)	2 (.126)	3 (.750)	3 (.177)	2 (.250)	4 (.211)	1 (.166)
Cat	0	5 (.250)	4 (.250)	1 (.250)*	4 (.235)	1 (.125)	5 (.263)	0
Bird	5 (1.00)	0	0	0	5 (.294)	0	0	5 (.834)
Squirrel	0	5 (.250)	5 (.312)	0	5 (.294)	0	5 (.263)	0
Rabbit	0	5 (.250)	5 (.312)	0	0	5 (.625)	5 (.263)	0
Total	5	20	16	4	17	8	19	0

(Probability of each feature occurring in parentheses.)

*Assuming one cat has had his ears clipped.

TABLE III

TOTAL PROBABILITY

Feature	Dog	Cat	Bird	Squirrel	Rabbit
4 legs	.250	.250	0	.250	.250
Round ears	.750	.250	0	0	0
Long tail	.177	.235	.294	.294	0
Flat face	.211	.263	0	.263	.263
Total	1.388	.998	.294	.807	.513
Multiplication	.007002	.000798	0	0	0

the outcome would result in a probability greater than one, and this is not possible. Furthermore, identification would be impossible if more than one animal had a probability of greater than one. Naturally, an animal is immediately eliminated if the probability of any feature occurring is zero. However, if this occurs for all the animals and since a choice must be made, the total probability for each animal must be divided by four, and the animal with the highest probability is chosen (for example, 1.388 divided by four). In the above case, the animal with the highest "score," is the dog.

The above might seem very machine-like and cumbersome, but possibly the brain does identify patterns on similar lines. The whole operation would occur instantaneously, and thus it might seem feasible.

As more features are learned, the brain is able to discriminate more easily, and the probability of a correct choice being made will be more certain. If there is a feature pertaining only to one particular pattern, then the probability for that pattern in the feature test will be one; and on the basis of this the decision will be certain, since the probabilities for this feature in the other patterns will be zero.

CHAPTER VI

METHOD, PROCEDURE AND RESULTS

The underlying hypothesis is that people recognize patterns on the basis of past experience. The more patterns of a certain category seen and identified correctly, the more likely an unknown pattern will be identified as this pattern rather than as another.

A second hypothesis is that it is the combination of features rather than individual features that enable correct identification. In other words, the hypotheses are that people recognize patterns in a similar fashion to our computer program, and that they will adopt the method for simple animal recognition suggested in the human model of the last chapter.

Materials

The stimulus objects were drawings of animals on an 8 X 11 inch card. The animals that were used were the same as in the previous model, that is, dog, cat, bird, rabbit and squirrel. Each drawing contained the four features which would enable recognition. These were: legs--either two or four; ears--either pointed or round; face--either pointed or flat; and tail--either long or short. Apart from these four variables, all other features were identical in all the drawings.

There were five drawings of each animal, and these drawings contained the combination of features as set out in Table II. In other words, all five dogs were shown to have four legs each; two had pointed ears, and three had round ears; three had long tails, and two had short tails; and four had a flat face while one had a pointed face. The five cats had four legs each; four had pointed and one had round ears; four had long tails and one had a short tail; and all five had a flat face. The five birds had identical features: two legs, no ears, long tail, and pointed face. The five squirrels and five rabbits also had identical features except for the tail, which was long in the former and short in the latter drawings. Apart from this difference both had four legs, pointed ears and a flat face.

In addition to these twenty-five drawings, there were two more: the "unknown" pattern and a drawing used for example purposes. The unknown sample had the same features as in the previous model, i.e., round ears, four legs, long tail and a flat face. The example drawing was that of a horse, the four critical features being clearly distinguishable.

Selection of the Subjects

The Ss were twenty undergraduate student volunteers from the school of business, Oklahoma State University. Each student was tested individually for half an hour. The testing interviews were spread over two consecutive days and were conducted under "suitable" testing conditions.

Apparatus

The sequence of stimuli were arranged as a deck of cards and were presented manually by E to S. The plastic covered cards were placed on a reading stand before each S with each card being replaced by another every twenty seconds.

Procedure

Prior to conducting the test each S was informed of the procedure and the instructions to be followed. These were read by E to S and thus the content and presentation of these were uniform for all Ss. S was told what four features to look for and during this indoctrination the example drawing was shown which gave an indication to S what type of stimuli to look for.

Each class of animal was presented five times, in other words, twenty-five drawings were presented during the learning stage, with the sequence of presentation being varied for each S. The name of each animal was written on the presentation card except on the unknown sample. S was told that he could write down any information he wished.

After the "learning" stage S was asked to identify the unknown sample. This he was to do on the basis of what he had already seen, taking into account only the four variable features. In order to eliminate pure guessing, which would have yielded a 50 percent "correct" response,[51] S was asked to explain his method of arriving at his decision and also to indicate why the alternative choice was not made. It was made clear that the sequence of presentation was irrelevant and it was also stressed that the four features were the only relevant ones.

Results

Table IV exhibits the over-all results obtained from the interviews. Of the 20 Ss tested, all suggested that the unknown sample was either

---

[51]Prior to the test the unknown sample was presented to a different group of people who were asked to recognize it; at first glance 70 percent claimed it was a cat and 30 percent, a dog.

TABLE IV

RESULTS

Recognition		Reason for Choice			Method of Decision Making			Certainty of Decision
		Ears	Face	All Features	Probability	Elimination	Other	
Dog	13	5	--	8	5	5	3	13
Cat	7	--	6	1	1	1	5	1
Subjects tested	20							

dog or cat. According to the human model the correct response should have been dog. The results obtained indicate that there was a 65 percent correct identification. Of these thirteen, five Ss based their decision to a large extent on one feature, i.e., ears. Of the five dogs seen, three had round ears whereas only one cat with round ears was seen. This feature was significant to those Ss and they did indicate that this was the overriding feature in their decision although they obviously should have taken the other two features--tail and face--into consideration. However, eight did take these two features into account and they did compare all the four features and even suggested that although the two features of the cat--long tail and flat face--could have influenced their decision, the round ears were the major factor for not choosing the cat.

As indicated, seven Ss decided upon the cat to represent the unknown sample. Of these six claimed the face was the predominant factor in their thinking and, contrary to the dog "choosers," they maintained that the face above all else, even the round ears, gave indication that it was a cat.

It is interesting to note that one S decided upon the cat after examining all the features. The reasoning was that of the five dogs only three had long tails; whereas four cats had long tails, the S then gave one point in favor of the cat. The cat gained another point when the flat face feature was examined; all five cats had a flat face and only four dogs had this feature. However the round ears feature--three dogs and one cat--enabled the dog to draw level with the cat. Since each animal had four legs this feature did not enter the scrutiny. The final choice, although admitted to be a guess by S, did strongly suggest that possibly with more time a probability approach, taking into account the features of the bird, rabbit and squirrel (as in our human model), would result.

The method of arriving at a decision was particularly interesting. Of those who decided upon the dog, five stated directly that the sample was not a cat because the probability of it being a dog was higher. The one S who chose a cat on a probability basis must have made an error in recognizing a feature on a dog drawing. Leaving aside the latter instance, 25 percent of the Ss made the correct response (according to our human model) using probability as the basis for their decision.

The other major method was the process of elimination. This involved eliminating an animal on the basis of not having one of the features that the unknown sample had. The Ss checked each feature systematically against the five categories of animals. The bird was eliminated immediately since it has only two legs. The tail scrutiny eliminated the rabbit, and finally the ears examination put an end to the squirrel's chances. This left only the dog and cat, and each feature of the unknown animal was possessed by at least one dog and one cat seen in the learning phase. The significant difference between this group of Ss as opposed to those who adopted the probability approach is that they stopped at this point and made up their mind on the basis of the ears feature or combination of features. Only one of the Ss who chose the cat used the elimination system and subsequently based the final decision on the face feature.

Those Ss recorded under "other" clearly did not make their final choice between dog or cat by the elimination of feature process. They disregarded the other three animal categories simply because they did not resemble or "look like" the unknown sample. These Ss made their final decision on the basis of either the ears or face feature.

Of the Ss who gave the dog response, all thirteen were certain of their decision on the basis of "more likely." On the other hand only one, or 14 percent who chose the cat, was certain of his decision on the basis of probability. This means that the Ss who gave the "correct" response were more sure of their choice than those who made the "incorrect" response. Furthermore, this indicates that they did think in terms of probability when investigating the ears feature and also that the other features, tail and face which were clearly favorable for the cat decision, were taken into account and weighed against the overwhelming ears feature.

Interpretation of Results

From the study we can draw many conclusions as to whether or not the concept formation process undertaken by individuals follows our human model.

The information obtained indicates that the probability hypothesis does not seem very significant; only 25 percent explicitly stated that they used this approach, although 65 percent did indicate that they were "more sure of their choice." However, the process of elimination used by 60 percent of the Ss does resemble the method in our human model. In this model, the multiplication of probabilities automatically eliminated the animal with one of the features missing.

A further conclusion is that 55 percent of the Ss based their decision on only one feature, all the features were obviously not given the same weighting. This, I think, might have resulted from past experience. The six Ss who relied on the ears feature might have regarded it as inconceivable that a cat, although they had seen one, could have round ears. On the other hand those Ss who relied on the face feature made their decision on the basis that all cats have flat faces whereas only some dogs have a flat face. Obviously the more dogs with flat faces and the more cats with round ears seen, the less will be the emphasis placed on such unique features. Or to say it in another way, the probability that it is a dog or cat will be far less.

We can also draw the conclusion that in the case of very simple stimuli with few features, the combination of features is not as important as the individual features themselves. However, as the stimuli become more complex the distinction between various features might not be easily distinguishable, and in such cases it will be the combination of features that will form the concept.

It seems that all these various characteristics are linked together in forming the concept and schemata by which objects are identified. If we were to utilize fully our capacity for registering the differences in things and to respond to each event encountered as unique, we would soon be overwhelmed by the complexity of our environment. The existence of discriminating capacities would, if fully used, make us slaves to the particular--the resolution is achieved by man's capacity to categorize. To categorize is to render distinguishably different things equivalent, to group the objects and events and people around us into classes, and to respond to them in terms of their class membership rather than their uniqueness. Our refined discriminatory activity is reserved only for those segments of the environment with which we are specially concerned. For the rest we respond by rather crude forms of categorical placement.[52] The learning and utilization of categories represents one of the most elementary and general forms of cognition by which man adjusts to his environment.[53] It is of immense importance to the individual to have a reliable method of differentiating the objects which do or do not vary from time to time and, in the case of the latter, of the manner in which they may be expected to vary.

In the case of the computer model, the machine has no way of knowing whether a pattern is simple or complex. Consequently, it must establish a routine that necessitates the checking of every feature and deciding upon an answer on the combination of features. The computer is unable to determine which feature carries the most weight unless it has checked all of them and refers back to its memory. But it is important to note that every feature does carry some weight in the final decision. The manner in which some of the Ss arrived at a decision does to some (albeit not very much) extent resemble pattern recognition by machine.

The result obtained, although not "significant" in the usual sense, are to some limited degree generalizable to people other than those used as Ss, but the degree to which they are generalizable to new stimuli remains a matter of conjecture.

---

[52] J. S. Bruner, J. J. Goodnow, and G. A. Austin, <u>A Study of Thinking</u> (New York, 1956), p. 1.

[53] Ibid., p. 2.

CHAPTER VII

CONCLUSION

The purpose of this paper has been to determine whether human beings recognize patterns in the same manner as suggested by a specific computer model. This computer model was constructed by Worthie Doyle at the Lincoln Laboratory, Massachusetts Institute of Technology. The program identified hand-written English letters on the basis of feature examination and probability distributions. During the learning stage, named patterns were presented to the machine and a memory was constructed embracing the features that would account for identification of an unknown pattern in the future. A probability distribution was associated with each feature, and it was the combination of these probabilities that enabled the machine to make the "most likely" choice when presented with an unknown sample.

In order to empirically test the hypotheses that people recognize patterns in the same way as the machine model, that is, by probability and feature combination, a human model was constructed using the patterns of animals instead of letters of the alphabet. Of the five animals presented each had four variable features associated with it. During the learning phase, a probability distribution for each feature in each animal was computed. When the unknown sample was presented, the examination of the four critical characters, followed by a probability tabulation, resulted in the "most likely" choice being made.

This method and procedure of pattern recognition was tested with a group of human subjects. The results obtained suggested that when recognizing patterns the various characters are linked together by the individual enabling concepts of the pattern to be formed. The actual method of final recognition was not determined with any degree of significance, however, the results do indicate that the notion of probability does enter into the decision-making process, since 100 percent of the Ss who made the "correct" response were certain of their choice on the basis of likelihood, as against only 14 percent of the Ss who made the "incorrect" response.

The study revealed a number of limitations that became evident during the testing of the Ss. Probably the most important was the influence of past experience. Although the Ss were told that for recognition they were only to use the data on the drawings, the responses of many of the Ss did indicate that previous concepts and knowledge of these animals influenced their decision.

Another limitation was that by verbally questioning the Ss at the conclusion of the test, E was unable to penetrate to any great depth to extract precise and unambiguous information. On a few occasions Ss became suspicious of persistent questioning, consequently their responses to the questions were restricted.

Suitable recommendations to overcome these drawbacks can be made. To eliminate past experience biasedness, that is, patterns seen before in everyday life, patterns of a "nonsense" nature should be substituted for the better known patterns. These patterns would, of course, contain variable characteristics as in any sensible pattern, but the pattern as such would be meaningless to the individual. Furthermore, to eliminate name association with the pattern, nonsensical names should be attached to each of these patterns. In this way the testing environment would be the only relevant one and past experience and learning would be nullified.

A further recommendation is that a printed questionnaire be used to elicit S responses. The questionnaire should, of necessity, be objective in nature and consist of only a two-choice response. In this way the S would be forced to answer questions in a specific and unambiguous manner.

Our most precise knowledge of perception is in those areas which have yielded to psycho-physical analysis (for example, the perception of size, color and pitch) but there is virtually no psychophysics of shape or pattern. The arrangement, differentiation and categorization of shapes into "configurations" have been the subject of the extensive research of the Gestalt psychologists.[54]

Clearly, we have not yet reached the stage where we can lay claim to know everything about the recognition process, consequently by using computer simulations some light may be thrown on the problem. Computer simulations make possible the study of systems of a complexity far too great to be handled by analytic methods, and thus open up for psychologists the possibility of studying the brain without first hopelessly simplifying it.[55]

As has been emphasized, attempts at computer simulation serve an important function in the clarification of theories in requiring detail-representation of psychological processes in a rigorous language. Once

---

[54] Attneave and Arnoult, p. 123.

[55] Leonard Uhr and Charles Vossler, "The Search to Recognize," Optical Character Recognition, eds. G. L. Fischer et al., (Washington, 1962), p. 326.

formulated in these terms the actual running of the program on a computer can provide information about further implications of the theory and about the possible inconsistencies in it.[56]

Once we have managed to perfect a pattern recognition program, many previously untestable hypotheses about the human brain will become testable using computer models. Consequently, a better understanding of the functioning of the human brain will result. It is much easier and more precise to test hypotheses on a machine, in which case the environment can be rigidly controlled and variables restricted. In this manner scientists will be able to break down the field of concept formation in the human brain into segments, and once these analyses have been carried out, they can then piece the puzzle together.

It is hoped that the research done in this field will help us to understand how humans learn patterns by providing additional clues to the fundamental mechanisms of human perceptual and conceptual learning.[57] Undoubtedly the combination of man and computer is capable of accomplishing things that neither of them can do alone.[58]

---

[56] Uhr, Computer Simulation of Personality, p. 3.

[57] Ibid., p. 232.

[58] W. A. Rosenblith, "What Computers Should Be Doing," Management and the Computer of the Future, ed. Martin Greenberger (1962), p. 313.

A SELECTED BIBLIOGRAPHY

Armer, Paul. "Attitudes Toward Intelligent Machines." Computers and Thought. Ed. Feigenbaum and Feldman. New York: McGraw-Hill, 1963, pp. 389-405.

Attneave, Fred, and Malcolm D. Arnoult. "The Quantitative Study of Shape and Pattern Recognition." Pattern Recognition. Ed. Leonard Uhr. New York: Wiley, 1966, pp. 123-142.

Berelson, B., and G. Steiner. Human Behavior. New York: Harcourt, 1964.

Bledsoe, W. W., and I. Browning. "Pattern Recognition by Machine." Proceedings of the Eastern Joint Computer Conference, Boston, December 1959. pp. 225-232.

Bruner, Jerome S., Jacqueline J. Goodnow, and George A. Austin. A Study of Thinking. New York: Wiley, 1956.

Cahill, Hugh E., and Carl I. Hoveland. "The Role of Memory in Acquisition of Concepts." Journal of Experimental Psychology (March 1960), pp. 137-144.

Doyle, Worthie. "Recognition of Sloppy, Hand-Printed Characters." Proceedings of the Western Joint Computer Conference, San Francisco, May 1960, pp. 133-142.

Frishkopf, L. S., and L. D. Harmon. "Machine Reading of Cursive Script." Information Theory. Ed. Colin Cherry. Washington: Butterworths, 1961, pp. 300-316.

Green, Bert F. "Computer Models of Cognitive Processes." Psychometrika (March 1961), pp. 85-91.

_______. Digital Computers in Research. New York: McGraw-Hill, 1963.

Grimsdale, R. L. et al. "A System for the Automatic Recognition of Patterns." Proceedings of the Institute of Electrical Engineers, Vol. 106, Part B, No. 26 (March 1959), pp. 210-221.

Hilton, Alice Mary. Logic, Computing Machines and Automation. Washington, D.C.: Spartan, 1963.

Hoveland, Carl I. "Computer Simulation of Thinking." The American Psychologist (November 1960), pp. 687-693.

Kendler, H. Basic Psychology. New York: Meredith, 1963.

Kendler, Tracy S. "Concept Formation." Annual Review of Psychology, 1961. Eds. Farnsworth and McNemar. California: Annual Reviews, 1961, pp. 447-472.

Kuhn, Alfred. The Study of Society: A Unified Approach. Illinois: Irwin, 1963.

Minsky, Marvin. "Steps Toward Artificial Intelligence." Computers and Thought. Eds. Feigenbaum and Feldman. New York: McGraw-Hill, 1963, pp. 406-450.

_______, and Oliver Selfridge. "Learning and Random Nets." Information Theory. Ed. Colin Cherry. Washington: Butterworths, 1961, pp. 335-347.

Rosenblatt, Frank. "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." Psychological Review (November 1958), pp. 386-408.

Rosenblith, W. A. "What Computers Should Be Doing." Management and the Computer of the Future. Ed. Martin Greenberger. Houghton, Michigan: M.I.T. Press, 1962, pp. 290-325.

Selfridge, Oliver. "Pandemonium: A Pradign for Learning." Proceedings of a Symposium on Mechanization of Thought Processes. London: H. M. Stationery Office, 1959, pp. 513-526.

_______. "Pattern Recognition and Learning." Information Theory. Ed. Colin Cherry. New York: Academic, 1956, pp. 345-353.

_______. "Pattern Recognition and Modern Computers." Proceedings of the Western Joint Computer Conference, Los Angeles, March 1955, pp. 91-93.

Uhr, Leonard. "Intelligence in Computers: The Psychology of Perception in People and Machines." Behavioral Science (April 1960), pp. 177-182.

_______. Pattern Recognition. New York: Wiley, 1966.

_______. "The Development of Perception and Language Simulated Models." Computer Simulation of Personality. Ed. Tomkins and Messick. New York: Wiley, 1963, pp. 231-266.

_______, and Charles Vossler. "A Pattern Recognition Program that Generates, Evaluates, and Adjusts its own Operators." Computers and Thought. Eds. Feigenbaum and Feldman. New York: McGraw-Hill, 1963, pp. 251-268.

_______, and Charles Vossler. "The Search to Recognize." Optical Character Recognition. Eds. Fischer et al. Washington, D.C.: Spartan, 1962, pp. 319-329.

_______, Charles Vossler, and James Uleman. "Pattern Recognition over Distortions by Human Subjects and by a Computer Simulation of a Model for Human Form Perception." Journal of Experimental Psychology (March 1962), pp. 227-234.

Van de Geer, John P., and Joseph M. Jaspars. "Cognitive Functions." Annual Review of Psychology, 1966. Eds. Farnsworth, McNemar and McNemar. California: Annual Reviews, 1966, pp. 145-176.

White, B. W. "Computer Applications to Psychological Research: Studies in Perception." Behavioral Science (July 1962), pp. 396-401.

VITA

Ivan Brian Kessel

Candidate for the Degree of

Master of Business Administration

Thesis: COMPUTER PATTERN RECOGNITION AND HUMAN CONCEPT FORMATION

Major Field: Business Administration

Biographical:

Personal Data: Born in Pretoria, South Africa, June 22, 1941, the son of Fabius and Sadie Kessel.

Education: Attended primary school in Johannesburg, South Africa; matriculated from Parktown Boy's High School, Johannesburg, in 1958; received the Bachelor of Commerce degree from the University of the Witwatersrand, Johannesburg, in March 1963; completed requirements for the Master of Business Administration degree in July, 1966.

Professional experience: Employed by Olivetti (Africa) Pty. Ltd., in Johannesburg as a sales representative from March 1963 to December 1964.