1991

Ms. Pacrat: A feeling, thinking machine

Lorraine D'Ortona

Rochester Institute of Technology
School of Computer Science and Technology

Ms. Pacrat: A Feeling, Thinking Machine

by Lorraine D'Ortona

A thesis, submitted to
the Faculty of the Computer Science Department
in partial fulfillment of the requirements for the degree of
Master of Science in Computer Science

Approved by:

30 July 91

Prof. Peter Anderson, Computer Science Department

30 July '91

Prof. David Suits, Philosophy Department

7/30/91

Prof. Al Biles, Computer Science Department

7/30/91

Prof. Fereydoun Kazemian, Computer Science Department

Title of Thesis

Computers and the Emotions

I Lorrane D'Ortuna hereby grant permission to the Wallace Memorial Library, of R.I.T., to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Date: 10/12/91

ABSTRACT

Since before the time of the first digital computers, the workings of the mind have been compared to that of a machine. With the onset of the discipline of Artificial Intelligence a truly organized attempt has been made to build intelligent machines that model the mind. Many interesting programs have been built, but the legitimacy of their success is a matter of great controversy. None of the AI programs developed so far have come close to the true power and intelligence of the brain. Expert systems, for example, are the most successful commercial AI programs, and even they have shown to be brittle, and only able to deal with knowledge in very narrow domains.

I suggest that those interested in modeling the mind should explore the emotions. I propose that intelligence and the emotions have a dependent and critical relationship. This relationship suggests that attempts to model human intelligence should consider how the emotions effect our thinking, reasoning, problem-solving, and learning and incorporate this information into computer models. This thesis will review what has been done in the field of AI to build intelligent machines and will examine the relationship between emotions and intelligence. A computer model of emotions will be presented: MS. PACRAT - A Feeling Thinking Machine.

TABLE OF CONTENTS

1. INTRODUCTION 1

2. AI AND THE ATTEMPT TO BUILD INTELLIGENT MACHINES: A CRITICAL HISTORY 3

2.1. PRE - PRE-HISTORY 4

2.1.1. Blaise Pascal 5

2.1.2. Gottfried Leibniz 5

2.1.3. Babbage and the Analytical Engine 6

2.1.4. George Boole 6

2.1.5. Alfred North Whitehead and Bertrand Russell 7

2.2. PRE-HISTORY: 1930-1950 8

2.2.1. Claude Shannon and Information Theory 8

2.2.2. Alan Turing - The Grandfather Of AI 10

2.2.3. Warren McCulloch and Walter Pitts: The First Neural Nets 13

2.2.4. Norbert Weiner and Cybernetics 14

2.2.5. John von Neumann 14

2.2.6. Donald Hebb 16

2.3. DAWN: 1950-MID 1960 17

2.3.1. Arthur Samuel and His CHECKERS program 18

2.3.2. Frank Rosenblatt and the Perceptron 20

2.3.3. John McCarthy- The Father Of AI 21

2.3.4. Marvin Minsky 21

2.3.5. Newell, Simon, and Shaw Logic Theorist and GPS 22

2.3.6. Joseph Weizenbaum's ELIZA 25

2.4. DARK AGES: LATE 1960's 25

2.4.1. Minsky Again and Papert: The attack 26

2.4.2. Feigenbaum and Buchanan Knowledge Engineering 26

2.5. RENAISSANCE: 1970's 27

2.5.1. PROSPECTOR 28

2.5.2. HEARSAY II 28

2.5.3. MYCIN, PUFF, INTERNIST-1, GUIDION, CADEUS, and RECONSIDER 29

2.5.4. Terry Winograd and SHRDLU 31

2.5.5. Colby's PARRY 32

2.5.6. Minsky Again: Frames 33

2.5.7. Roger Shank: Scripts, CYRUS, and FRUMP 34

2.5.8. Ross Quillian and Semantic Nets 36

2.5.9. Doug Lenat: Automated Mathematics (AM) 36

2.6. NEW AGE: 1980's 37

2.6.1. Past Tense Of English Verbs: Rumelhart, Hinton and Williams 38

2.6.2. R. Pfeifer and D. Nicholas: A Model Of Anger 38

2.6.3. ART: Automated Reasoning Tool 42

2.6.4. Common Sense: Cyc by Doug Lenat 43

2.6.5. NETtalk, Rosenburg and Sejnowski 45

2.7. THE 1990's 45

2.7.1. Attila the Robot 45

3. EMOTIONS AND INTELLIGENCE 48

3.1. PSYCHOLOGICAL THEORIES OF EMOTIONS 50

3.1.1. Two Prominent Theories 52

3.1.2. Other Theories 53

3.1.2.1. Sigmund Freud 53

3.1.2.2. Charles Darwin 55

3.1.2.3. S.S.Tomkins 56

3.1.2.4. O.H.Mowrer 59

3.1.2.5. S.Schachter and J.Singer 61

3.1.2.6. Carroll Izard 64

3.1.2.7. Magda Arnold 67

3.1.3. Summary 69

3.2. MARVIN MINSKY ON THE EMOTIONS 70

3.2.1. Background Information 71

3.2.2. The Interaction Of Emotions and the Intellect 72

3.2.2.1. Emotion's role in motivation and self-control 72

3.2.2.2. The involvement of emotions in intellectual application and development 74

3.2.2.3. Emotions and Learning 75

3.2.2.4. Emotions and the Process of Transferring Goals and Values 76

3.2.2.5. Emotions as warning signs 76

3.2.2.6. Choices, comparisons, and Compromises in Problem Solving 77

3.2.2.7. Miscellaneous Thoughts 78

4. PACRAT: NON-LIVING INTELLIGENCE 80

4.1. THE IMPLEMENTATION OF SCANLON AND JOHNSON 80

4.1.1. Let's Meet Pacrat: An Introduction 80

4.1.2. What is Pacrat Really Like 82

4.1.2.1. Pacrat And His World 82

4.1.2.2. Pacrat's Anatomy 83

4.1.2.2.1. Details Of Pacrat's Brain 85

4.1.2.2.2. The Hypothalamus 85

4.1.2.2.3. The Cingulate Gyrus 85

4.1.2.2.4. The Hippocampus 86

4.1.2.2.5. The Thalamus 87

4.1.2.2.6. The Amygdala 88

4.1.2.2.7. The Reticular Ascending Substance 89

4.1.2.2.8. The Medial Forebrain Bundle 89

4.1.3. Why a 'Thinking, Feeling' Machine? 90

4.1.3.1. We Think With Feeling 90

4.1.3.2. Thinking and Feeling are not Rule-Based Or Logical 91

4.1.4. The Future of Pacrat 92

4.1.5. Final Remarks 92

4.2. MY IMPLEMENTATION OF PACRAT 93

4.2.1. Program Introduction 93

4.2.2. The Coding of Ms. Pacrat 93

4.2.3. Experiences with the Pacrat program 94

4.2.4. Program Results 95

4.2.4.1. Graph Specifics 96

4.2.4.2. Remaining Visual Displays 97

4.2.4.2.1. Figure 1 97

4.2.4.2.2. Figure 2 98

4.2.4.2.3. Figure 3 99

4.2.4.2.4. Figure 4 100

4.2.4.2.5. Figure 5 100

4.2.4.2.6. Figure 6 100

5. CONCLUDING REMARKS 101

6. APPENDIX: SOME INTERESTING QUOTES 102

7. BIBLIOGRAPHY 107

1. INTRODUCTION

As a result of their attempt to build intelligent machines, Artificial Intelligence researchers have opened up new avenues in the study of the nature and subsistence of the mind. The comparison of the computer to the human mind/brain had its beginnings, even before chips and neural networks. Turing, the great contributor to the study of computing in the 1930's, pondered and proposed answers to the study of the mind/brain question. Can a mind/brain be programmed? Can a machine really think? Artificial intelligence has emerged to join in the attempt to answer these questions. Besides AI there are actually a diversity of fields involved in this task. They include psychology, philosophy, cognitive science, neuroscience, mathematics, and computer science.

Even with all the impressive results so far in the field of AI, there have been certain characteristics, or processes of human thinking that have not been easily simulated using computers. These items may well contain the key to how we make decisions, solve problems, and survive. These processes include creativity, emotions, intuition, insight, introspection, and consciousness, among others. These processes certainly affect the way humans think, and have withstood most (or all?) attempts to program them.

As the main subject of this paper I will concentrate on one of these processes, the emotions. I propose that intelligent behavior and the emotions have a dependent, critical relationship and that this relationship should be examined.

This thesis is divided into three main sections. Section 1 is the introduction. Section 2 includes a critical history of the field of AI and its attempt to produce intelligent machines. Section 3 discusses the relationship between the emotions and intelligence. Section 4 presents MS. PACRAT, a computer model of emotion.

2. AI AND THE ATTEMPT TO BUILD INTELLIGENT MACHINES: A CRITICAL HISTORY

Artificial Intelligence is the branch of computer science that is interested in building machines that can do things that we consider to be intelligent. This includes things such as; playing chess, recognizing a face, or making a medical diagnosis. More specifically there seems to be two major goals in AI research. The first is to write useful and interesting programs. This goal has been fulfilled somewhat with the creation of expert systems that are used in industry today. The second goal in AI is to gain an understanding of human mental processes and attempt to simulate them with computer programs.

In this section I will discuss some of the relevant history in the field of Artificial Intelligence in relation to these goals. I will discuss the research of some interesting people including John McCarthy, the father of AI, Marvin Minsky, and Terry Winograd. These men have made attempts to simulate some of the processes of the mind through computer models. I will do this by giving a short history of the work done in this area and by examining specifically some models of mind relating to my thesis topic of emotions and mental states in general. I will also discuss the attempts of those who have chosen neural networks, as opposed to the more traditional programming of AI – i.e., expert systems, rules, and logic programming – in their attempts to model the mind. I will examine the successes and the failures of some of

these attempts. To discuss this topic I will use my own variation of a framework presented by M. Mitchell Waldrop in his book *Man-Made Minds.*

The attempt to build intelligent machines and programs was at first looked upon in a very optimistic manner. Many hopeful predictions were made, some of which have been fulfilled, while the success of others is the subject of much debate. Is it really possible for a machine to exhibit intelligent behavior? Is it important that the logistics of this behavior be exactly the same as that of a human? Can only brains be intelligent? What is intelligence? These are some of the many questions that AI researchers have been attempting to answer. One methodology chosen is to build computer models of intelligent tasks. Paul Churchland, in *Matter and Consciousness,* writes:

> One puzzling fact about the results of AI research is that there are certain kinds of tasks such as number crunching, theorem proving, and list searching, which standard computers do very swiftly and very well, while the human brain does them only slowly and comparatively badly. On the other hand there are certain kinds of tasks, such as facial recognition, scene apprehension, sensorimotor coordination, and learning which humans and other animals do swiftly and well, but even the fastest computers running the most sophisticated programs do only slowly and rather poorly.[1]

2.1. PRE - PREHISTORY

The following is a discussion of ideas that pre-date the earliest computers. It was during this time that the basis for modern day computers

[1]Paul M. Churchland. *Matter and Consciousness.* Cambridge, Mass: The MIT Press, 1984. p. 120.

was thought about and put forth. Even during this period comparisons were made between mechanical devices and the mind.

2.1.1. Blaise Pascal

During his life, from 1623 to 1662, Pascal designed a mechanical calculator that could do the functions of addition and subtraction. This work caused him to question the nature of human thinking. He disagreed with the basic views of Descartes, that the mind is and should be logical, and that the emotions are a negative force that interfere with logic. Pascal believed that we must trust our intuitions and our emotions. He said, "The heart has its reasons that reason does not know".[2]

2.1.2. Gottfried Leibniz

Leibniz, who lived from 1656 to 1716, is credited with building the first calculating machine that could multiply and divide and with the origination of the binary number system. He also pondered the nature of the mind, but unlike Pascal, Leibniz claimed that thinking was a process that could be explicitly stated in rule-like form. He felt that knowledge could somehow be formalized in a universal and exact system of notation so that all concepts could be analyzed into a small number

[2]Hubert and Stuart Dreyfus. *Mind Over Machine.* NY, NY: The Free Press, 1986. p. 3.

of original ideas. Leibniz felt that thought could be reduced to the manipulation of numbers or computation.[3]

2.1.3. Babbage and the Analytical Engine

Babbage was an eccentric British mathematician who lived and worked during the years 1791 – 1871. He devised the plans for two machines: the Difference Engine, which was to be used for calculating mathematical tables, and the Analytical Engine, a more general machine that was to be used for a variety of mathematical tasks. These models had many of the same characteristics of our computers today, including a CPU, a memory, and programs that were represented as holes on cards. Babbage, along with some of his colleagues, including the famous Ada Lovelace, speculated on the question of machine intelligence. Lovelace in particular described a machine as only a symbol manipulator possessing no originality or creativity, that could do only what we tell it to do.[4]

2.1.4. George Boole

Boole was an English mathematician who lived during the mid 1800s. His contribution to the field of AI is the development of

---

[3]Hubert Dreyfus. *What Computers Can't Do.* NY, NY: Harper and Row Pub., 1972. p. 69.

[4]Mike Sharples. *Computers And Thought.* Cambridge, Mass: The MIT Press, 1989. p. 11.

symbolic logic. Logic has been used in many AI programs, and as a theory for how thinking and reasoning may take place in humans. It is the basis for the programming language PROLOG. Boole thought that reasoning could be calculating, and tried to "investigate the fundamental laws of those operations of the mind by which reasoning is performed, to give expression to them in the symbolic language of a Calculus....".[5]

2.1.5. Alfred North Whitehead and Bertrand Russell

In this same area there were also contributions by Bertrand Russell and Alfred North Whitehead who wrote *Principia Mathematica* during the years 1910 – 1913. They demonstrated that mathematics is founded on logic and is, in fact, just another form of logic. It is interesting that one "successful" AI program, Logic Theorist, developed by Newell, Simon, and Shaw, was able to prove theorems found in Whitehead and Russell's work, and even found proofs that were considered much more elegant than those that human minds had come up with.

[5] Hubert Dreyfus. *What Computers Can't Do*. NY, NY: Harper and Row Pub., 1972. p. 70.

2.2. PRE-HISTORY: 1930 - 1950

During the time of WWII the first computers were built. The U.S. built machines that could calculate artillery tables. Great Britian created computers with the ability to break German military codes. Germany put their efforts towards machines that could calculate airplane wing flutter. Immediately following this attempts were made to create intelligent programs.[6]

2.2.1. Claude Shannon and Information Theory

Shannon was an engineering student at MIT. In 1937 he showed that the operation of electrical relays and switching circuits could be analyzed to yield such logical elements as True, False, And, Not, and Or. In 1948 he developed what is known as Information Theory. This theory proposed a mathematical definition for information. He believed that the information content of a message was defined by the number of symbols required to encode it, and not by what the message actually said. There is debate about Shannon's theory concerning its relevance to the semantics of a piece of information. In his famous work *The Mathematical Theory Of Communication* , Shannon says:

> The fundamental problem of communication is that of reproducing at one point either exactly or approximately a

[6] M.Mitchell Waldrop. *Man-Made Minds*. USA: Walker and Company, 1987. p. 18.

> message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem.[7]

Some, including Shannon, feel that it was wrong to consider Information Theory a theory of meaning. Others, however, have attempted to use it in this way and have proposed that it could explain how the brain processes and understands information.

Shannon also started pondering the possibility of creating a chess program in 1950. He did not develop a specific program, but considered the issues and problems involved in programming an electronic computer to play a game of chess. Specifics of these ideas can be found in Feigenbaum and Feldman (1963). He estimated that in chess the number of complete moves in an average length game is 10^{120}. He wrote articles on this subject and stated:

> Investigating one particular line of play for 40 moves would be as bad as investigating all lines for just two moves. A suitable compromise would be to examine only the important possible variations – that is, forcing moves, captures and main threats – and carry out the investigation of the possible moves far enough to make the consequences of each fairly clear. It is possible to set up some rough criteria for selecting important variations, not as efficiently as a chess master, but sufficiently well to reduce the number of variations appreciably and thereby permit a deeper investigation of the moves actually considered.[8]

---

[7] Hubert Dreyfus. *What Computers Can't Do*. NY, NY: Harper and Row Pub., 1972. p. 165.

[8] Ibid. p. 74.

In general, Shannon pondered the possibility of building intelligent machines. He proposed that the tasks of pattern recognition and language translation were probable for computers, but also suggested that they may require a different type of computer than those that existed during his lifetime. He says:

> It is my feeling that this will be a computer whose natural operation is in terms of patterns, concepts, and vague similarities, rather than sequential operations and ten-digit numbers.[9]

2.2.2. Alan Turing - The Grandfather Of AI

Ever since I learned about Alan Turing, his thoughts have always been fascinating to me. During a time period when computers were in their infancy Turing began to make predictions and ask questions such as "Can machines think?" Really think? Can a machine be intelligent and exhibit behavior that we would consider to be intelligent? To answer these questions Turing proposed his famous "Imitation Game", later known as the "Turing Test", in his 1950 article "Computing Machinery and Intelligence". He says to imagine that you are alone in a room with a teletype machine that is connected at the other end to either a person or a computer. If through questioning and conversing

[9]Hubert Dreyfus. *What Computers Can't Do.* NY, NY: Harper and Row Pub., 1972. p. 304.

with the "thing" on the other end, you cannot tell whether it is a computer or a person, then you have to concede that a machine can think.[10] For simulation purposes, the dialogue of this machine would be made flexible enough to allow it to be indistinguishable from that of a real human. It would be able to answer questions concerning the "stuff of the world", and even have the ability and sense to try to fool the player of the game. To Turing, any machine that could do this would have to be intelligent, and be able to think. It could also then be attributed with having mental states and intentionality just like humans.

The validity of the Turing Test is still frequently discussed in literature along with his famous question, Can machines think? It seems that the most genuine answer may be, Who knows? This is because we really do not understand what thinking is. Another response may be, Who cares? If a machine can do a good job, does it really matter? Some say, yes, machines can think, but they do not do it the same way we do it. The most common answer though seems to be that, no, machines do not think, and they never will. In *Man-Made Minds* Waldrop writes:

> Furthermore, the essence of humanity isn't reason, logic, or any of the other things that computers can do; it's intuition, sensuality, and emotion. How can a computer think if it does not feel, and how can it feel if it knows nothing of love, anguish,

---

[10] M.Mitchell Waldrop. *Man-Made Minds*. USA: Walker and Company, 1987. p. 16.

> exhilaration, loneliness, and all the rest of what it means to be a living human being.[11]

I agree with Waldrop's comment, and that has led me to present this thesis. Perhaps it is worth attempting to program a machine with these things that Waldrop calls the "essence of humanity". The question, however, of the plausibility of doing this, of course, still remains.

Another contribution of Turing is development of a hypothetical logic machine in 1936, that could compute anything that could be computed. This meant that any logical reasoning process, no matter how complicated, could be done on a real physical device. Turing called his machine a universal machine, and explained that a digital computer is one such machine. He wrote:

> The special property of digital computers, that they can mimic any discrete state machine, is described by saying that they are universal machines. The existence of machines with this property has the important consequence that, considerations of speed apart, it is unnecessary to design various new machines to do various computing processes. They can all be done with one digital computer, suitably programmed for each case. It will be seen that as a consequence of this all digital computers are in a sense equivalent.[12]

These ideas imply that any process that can be formalized and represented as a series of steps can be reproduced by a machine. These processes may also include certain kinds of thinking that could be done

[11] M.Mitchell Waldrop. *Man-Made Minds*. USA: Walker and Company, 1987. p. 16.

[12] Hubert Dreyfus. *What Computers Can't Do*. NY, NY: Harper and Row Pub., 1972. p. 72.

without the use of the human brain.[13] In *Matter and Consciousness* Paul Churchland sums up this debate quite well. He writes:

> The question that confronts the research program of AI, therefore, is not whether suitably programmed computers can simulate the continuing behavior produced by computational procedures found in natural animals, including those found in humans. That question is generally regarded as settled. In principle, at least, they can. The important question is whether the activities that constitute conscious intelligence are all computational procedures of some kind or other. The guiding assumption of AI is that they are, and its aim is to construct actual programs that simulate them.[14]

2.2.3. Warren McCulloch and Walter Pitts: the first Neural Nets

The foundations for connectionism were first laid by these two men during the 1940s. They were the first to prove that some logical functions could be computed by networks of simple neurons. Neural nets were designed by McCulloch and Pitts as blueprints for the special senses (the senses of sight, hearing, touch, taste, and smell) as well as the central processes of the human brain which involves the ability to learn, think, and solve problems.[15]

---

[13] M.Mitchell Waldrop. *Man-Made Minds*. USA: Walker and Company, 1987. p. 16.

[14] Paul M. Churchland *Matter and Consciousness*. Cambridge, Mass: The MIT Press, 1984. p. 105.

[15] Frank George *AI Its Philosophy and Neural Content*. NY, NY: Gordon and Breach Science Publishers, 1986. p. 31.

2.2.4. Norbert Weiner and Cybernetics

Norbert Weiner was at MIT, and in 1948 published his theory of Cybernetics. This theory has three main elements: the importance of feedback for self-correction and the development of purpose, information theory, and the electronic computer. The idea of feedback can be illustrated by a simple thermostat. It carries out the function of keeping a room at a certain temperature by constantly monitoring the present state of the room. This information is fed back into the system, and if any deviation in temperature is found, the thermostat corrects it by turning a furnace on or off. The thermostat then has a goal, that of keeping a room at a certain temperature, and so according to Weiner, has a purpose.

2.2.5. John von Neumann

John von Neuman is credited with conceptualizing the computer in its modern form during the 1940s. His model contained three parts: a memory unit, a CPU, and a data channel connecting them. Von Neumann also originated the idea that a machine could treat both its data and its programs in the same way. They could both be stored in memory, and both be expressed in the same codes. This concept allowed the possibility for a computer program to treat itself as data, and thus,

rewrite or correct itself. As we have seen today this idea has become a reality, with AI programs that do have the ability to adjust themselves.

Von Neumann's computer is called a serial computer. It performs its operations one at a time. He had hoped that the increasing speed capacity of computers with time would make this drawback be insignificant. One other problem with the architecture of the von Neumann computer is its inefficient use of the CPU and memory. The CPU spends time doing nothing while waiting for data and instructions to be passed to and from memory and most memory cells are inactive most of the time. These limits have forced researchers to look to alternatives. In *Man-Made Minds* Waldrop explains:

> But within a decade or so the von Neumann architecture is going to be up against some severe limits. The very fastest number-crunching computers, such as the Cray XM-P (manufactured by Cray Research of Minneapolis), have already begun to feel the constraints of relativity; electronic signals can't travel faster than the speed of light, which means that at some point it doesn't matter anymore how fast the individual microcircuits are. Any gains made in one part of the computer are eaten up by the delay in transmitting data to another part. Nor does it help to put the chips closer together. Pack them too densely and they begin to melt because there's no way to dissipate the heat they generate.[16]

In AI these limits are extremely serious. AI programs require immense computing power, hence the interest in Connectionism and the scepticism towards traditional AI programs on serial computers. The nature

[16] M.Mitchell Waldrop. *Man-Made Minds*. USA: Walker and Company, 1987. p. 112.

of the brain seems to be that of a parallel processor, requiring AI researchers to reevaluate their needs. The field is attempting to model the functions of this concurrent mechanism with its billions of processors, using serial devices. According to Turing, this may be theoretically possible, but what about the actuality of attempting to build real working models?

2.2.6. Donald Hebb

Hebb was concerned with the nature of cognition and its neural realization. During the decades of the 1940s he showed that neural networks could exhibit the properties of learning and memory. He believed that for a machine to support learning there had to be some kind of physical change in the machine to represent the fact that learning had taken place. This thought resulted in Hebb's rule which states that if two units on each end of a connection were simultaneously activated, then the connection between them would be strengthened. This rule reflects the assumption that the more a connection is used, the easier it becomes to use it again.[17]

Hebb suggested that memory may not be formed by individual neurons each storing single concepts, but by patterns of

[17]Heinz R. Pagels. *The Dreams Of Reason.* NY, NY: Bantam Books, 1989. p. 118.

interconnections within groups of neurons. He called these "cell assemblies" or "phase sequences". Hebb hypothesised that learning occurs at the synapses between nerve cells in one of these assemblies. He also suggested that these assemblies self-assembled during the maturation process of the organism.[18]

2.3. DAWN: 1950 - MID 1960

It was during this period that the field of AI gained status as a legitimate field of study. The foundations were being laid for cognitive science, and many AI labs were started. Mike Sharples discusses this period in *Computers and Thought.* He comments:

> In the 1960s, general problem solving methods, supplemented by domain specific heuristics, were applied to a wide range of problems, and AI separated out into the application areas of language understanding and generation, game playing, theorem proving, vision, and robotics. With a few exceptions (such as Newell and Simon's monumental work *Human Problem Solving* , 1972) there was little attempt to construct programs that accurately modeled the human mind; the emphasis was on performance – computer systems that acted in intelligent ways by, for instance, playing chess or solving mathematical problems. The general approach was one of successive refinement: start with a method that approximates the desired action (e.g., a program that plays chess badly, or solves puzzles in a large number of steps) and then refine it until it behaves to the standard expected of an intelligent human.[19]

---

[18] X. Nadel. *Neural Connections, Mental Computation.* Cambridge, Mass: MIT Press, 1989. p. 1.

[19] Mike Sharples. *Computers And Thought.* Cambridge, Mass: The MIT Press, 1989. p. 18.

During this period work continued in the area of neural networks as a means of reproducing human intelligence. Symbolic AI and neural networks enjoyed equal footing in interest of research and in AI journals.

2.3.1. Arthur Samuel and His CHECKERS Program

Samuel actually began considering a checkers program as early as 1948. His first program was proposed in 1959. This model learned to play the game of checkers. One remarkable thing about this program is that it had the ability to improve its performance, and eventually was able to beat its creator. Samuel's objective was to create a program that learns. He chose the environment of the game of checkers in which the learning process would take place. Samuel claims that checkers was chosen because the simplicity of its rules permits greater emphasis to be placed on learning techniques, and checkers contains all of the basic characteristics of an intellectual activity in which heuristic procedures and learning processes can play a major role and in which these processes can be evaluated.[20]

Samuel's program was given the basic rules for the game of checkers, and a manner of accepting the opponent's moves and displaying his

[20]Edward Feigenbaum and Julian Feldman. *Computers and Thought*. NY, NY: McGraw-Hill, 1963. p. 72.

own moves. He employed a standard depth first minmax search procedure. The computer plays by looking ahead a few moves, and by evaluating the resulting board positions which are scored. This scoring is done in terms of a linear polynomial. In order to learn better moves, the computer stores, based on score, certain board positions that can be recalled for reference. Special mechanisms are set up to catalog boards that are saved, delete redundancies, and to discard board positions which are not believed to contain value. This method is called rote-learning.[21] A second method for learning that involves generalizations was tried by Samuel. He attempted to give the program the ability to generalize on the basis of past experience and to save only the generalizations. "Samuel believes that the effective path to progress in AI is probably not that of imitating and adapting human processes."[22] His checkers program does not attempt to play using human-like reasoning. Certainly a human checkers player does not base his move on a mathematical evaluation. He may however look a few moves ahead, and may also have memories of previous games and board positions to pull from. In this sense, the procedure steps are similar, but the heuristics are not. Samuel's program is considered with high regard in the AI field. Programming specifics can be found in Samuel's article in Feigenbaum

[21]Edward Feigenbaum and Julian Feldman. *Computers and Thought.* NY, NY: McGraw-Hill, 1963. p. 80.

(1963).

Checkers, however, did not escape criticism. Dreyfus maintains that Checkers is not a champion and that his own creator admits Checkers' limitation. Dreyfus writes, "According to Samuel, after thirty-five years of effort, 'the program is quite capable of beating an amateur player and can give better players a good contest'."[23] Dreyfus contends that this is not such a spectacular record.

2.3.2. Frank Rosenblatt and the Perceptron

Rosenblatt extended the neural networks of McCulloch and Pitts to create the "perceptron" in 1959. He built a seeing and learning machine that eventually could recognize all 26 letters of the alphabet, written in one specified format. Rosenblatt believed that humans were born with a randomly connected network of neurons. Through learning, this network becomes organized into a representation of the world.

Rosenblatt began a split in the AI community by stressing the superiority of his perceptron and the inferiority of symbolic AI and its techniques. He felt that the only way to recreate the intellectual powers of the brain was to imitate the way the brain does it. Some in the AI community, including Minsky, felt that it did not matter so much how

[22] Edward Feigenbaum and Julian Feldman. *Computers and Thought*. NY, NY: McGraw-Hill, 1963. p. 38.

[23] Hubert and Stuart Dreyfus. *Mind Over Machine*. NY ,NY: The Free Press, 1986. p. 107.

the brain carries out its functions. To them, the important question was to simulate what the brain did using whatever techniques were available, even though they were totally foreign to the human brain.

2.3.3. John McCarthy - The Father Of AI

John McCarthy was an assistant professor of mathematics at Dartmouth college in 1956 when he organized the first conference to study AI. This conference is considered the birthplace of AI and included such participants as Newell, Simon, and Minsky. The purpose of this conference was to gather all those involved in AI research to define the field and gather direction for it. It was this conference and specifically McCarthy that created the name of this new field: Artificial Intelligence. In 1957 McCarthy left Dartmouth and co-founded one of the first AI labs at MIT. In 1962 he went to Stanford and set up another AI lab. In 1958 McCarthy invented LISP, the most widely used AI language.

2.3.4. Marvin Minsky

Minsky co-founded one of the first AI labs with McCarthy at MIT in 1957. He was involved in the attempt to construct programs that solve intelligent tasks. In 1961 he wrote an article on this subject titled,

"Steps Towards AI". He discussed subjects such as search, pattern recognition, learning, planning, and induction. Minsky writes,"To me 'intelligence' seems to denote little more than the complex of performances which we happen to respect, but do not understand."[24]

Minsky also argues that machines can think. He writes:

> But we should not let our inability to discern a locus of intelligence lead us to conclude that programmed computers therefore cannot think. For it may be so with man, as with machine, that, when we understand finally the structure and the program, the feeling of mystery (and self-approbation) will weaken.[25]

Another topic of interest that Minsky discussed is that an intelligent machine would need not only an internal representation of the world, but a representation of itself. He writes:

> If one asks the creature, "Why did you decide to do such and such?" (or if it asks this of itself), any answer might come from an internal model. Thus the evidence of introspection itself is liable to be based ultimately on the processes used in constructing one's image of one's self.[26]

2.3.5. Newell, Simon, and Shaw - Logic Theorist and GPS

Newell and Simon viewed the mind as a symbol-manipulating, or information processing system. In 1955 they began working together at

[24] Edward,Feigenbaum and Julian Feldman. *Computers and Thought*. NY, NY: McGraw-Hill, 1963. p. 447.

[25] Ibid. p. 447.

[26] Ibid. p. 449.

Carnegie Mellon to design a program that could exhibit intelligence. Their first program was Logic Theorist which could prove mathematical theorems in symbolic logic. Heuristic reasoning is the central strategy of Logic Theorist. Newell and Simon rejected the idea of brute force to solve a problem. They suggest that this is not the way humans do it. We use rules of thumb, intuition, and hunches to solve problems. Newell and Simon say that heuristics are very different from algorithms. Heuristics, unlike algorithms, are not guaranteed to work. Their outcome is not predictable.

Simultaneously, Newell, Simon, and Shaw were already designing an improved version of Logic Theorist, GPS (General Problem Solver). It was intended to be a model of human thought processes. It was used to solve problems in logic, as well as various games and puzzles. In 1972 they published *Human Problem Solving* , detailing their ideas about the nature of the mind.

Newell and Simon intended their systems to solve problems in the same way people do. They felt that heuristics were used by people, and so could be used by machines. Simon said that some machines in existence can literally think. Newell claimed that they had discovered that the key to intelligence was found in physical symbol manipulation. Many opposed this notion however, and GPS was criticized. Searle, in

Mind's Brains and Science , discusses the question: Can a digital computer think? He explains:

> This is to say: Is instantiating or implementing the right computer program with the right inputs and outputs, sufficient for, or constitutive of, thinking? And to this question, unlike its predecessors, the answer is clearly 'no'. And it is 'no' for the reason that we have spelled out, namely, the computer program is defined purely syntactically. But thinking is more than just a matter of manipulating meaningless symbols, it involves meaningful semantic contents. These semantic contents are what we mean by mental states, having a mind, all of these simulations are really irrelevant. It doesn't matter how good the technology is, or how rapid the calculations made by the computer are. If it really is a computer, its operations have to be defined syntactically, whereas consciousness, thought, feelings, emotions, and all the rest involve more than a syntax. Those features, by definition, the computer is unable to duplicate however powerful may be its ability to simulate. The key distinction here is between duplication and simulation. And no simulation by itself ever constitutes duplication.[27]

Searle's famous Chinese Room argument also expresses this point. Searle says that the person inside the room does not understand the meaning of the symbols he is producing. He only knows syntax, not semantics. He says that a manipulation of linguistic tokens is taking place, rather than an understanding of how the symbols relate to reality.[28]

---

[27]John Searle. *Minds, Brains and Science.* Cambridge, Mass.: Harvard University Press, 1984. p. 37.

[28]Nadel, Cooper, Cukover, and Harnish. *Neural Connections, Mental Computations.* Cambridge, Mass: MIT Press, 1989. p. 292.

2.3.6. Joseph Weizenbaum's ELIZA (1963)

Weizenbaum's natural language understanding program, ELIZA, played the role of a non-directive therapist. The program used pattern matching, searching and pre-defined responses to simulate natural language understanding. Eliza was impressive to many people, resulting in stories concerning people who, for brief periods, treated ELIZA as if he (it) were a sympathetic human. Many people attributed human characteristics to ELIZA.

Could ELIZA be considered intelligent? Could we say that ELIZA can really understand English? Probably the biggest critic of ELIZA was Weizenbaum himself. He believes that understanding comes about only within a contextual framework: "In other words, we only understand what someone says to us because of what we have in common in our understanding of the world."[29] He felt that ELIZA's representation of knowledge was too shallow and inflexible to be intelligent.

2.4. DARK AGES: LATE 1960's

It was during this period that disillusionment set into the AI community. Creating intelligence seemed to be much more difficult than was first thought.

[29] Mike Sharples. *Computers And Thought*. Cambridge, Mass: The MIT Press, 1989. p. 37.

2.4.1. Minsky Again and Papert: The attack

Minsky and Papert contributed to a stagnation of neural network research by discrediting much of the work done on Perceptrons during the 1960s by claiming that the capabilities of the Perceptron had been exaggerated. Their book, *Perceptrons,* proved that single-layered linear systems could not compute certain important classes of functions. Their efforts caused a suspending of funding of neural net work during the 1970s.

2.4.2. Feigenbaum and Buchanan Knowledge Engineering

In 1965, Feigenbaum and Buchanan founded the Stanford's Heuristic Programming Project. Their first effort was Dendral, which did chemical analysis by interpreting the output of a mass spectrometer. It was completed in the late 1960's. Dendral is considered to be the first expert system.

Dreyfus contends that Dendral is actually not that successful. He claims to have contacted universities and industrial sites and found that Dendral is not widely used commercially.[30]

[30]Hubert and Stuart Dreyfus. *Mind Over Machine.* NY, NY: The Free Press, 1986. p. 110.

2.5. RENAISSANCE: 1970s

It was during this period that an emphasis on real-world applications dominated AI. A new concept for investigation emerged. It was speculated that the fundamental ingredient of intelligence was knowledge. Many questions were asked such as: What is knowledge? How can knowledge be represented in a machine? And, Can a computer learn? Expert systems, vision, and natural language understanding programs flourished.

Dreyfus (1986) discusses two types of knowledge held by experts. The first type is the facts of the domain of the expert. This he describes as widely shared knowledge that is written in the textbooks and journals of the field. The second type of knowledge is heuristic knowledge. Dreyfus describes this as knowledge of good practice and good judgement in a field. It is experiential knowledge, the "art of good guessing", that a human expert acquires over years of work.[31]

The use of production systems became a popular technique during this period. A production system uses condition action pairs called production rules. A production rule is in the form: if this condition holds (or is true), then a certain action is carried out. The left part of the rule states the conditions that must be true before the right side or action part can be executed. During execution, the program searches the left side for a match,

[31] Hubert and Stuart Dreyfus. *Mind Over Machine.* NY ,NY: The Free Press, 1986. p. 104.

and when one is found, the right side "fires". One advantage of production systems is that rules can be added, deleted, or modified, very easily.

2.5.1. PROSPECTOR

PROSPECTOR was designed to advise geologists on whether or not a given site might have ore-grade deposits. It was developed at SRI international, a nonprofit research group in Menlo Park, California.

Dreyfus maintains that the success of PROSPECTOR is exaggerated. He claims that there is no proof that PROSPECTOR has outperformed human experts in the field.[32] It is also suggested that PROSPECTOR'S knowledge is artificial in the sense that it knows rules for how to identify various kinds of geological formations, but it has no idea what sedimentation, metamorphism, or vulcanism were.[33]

2.5.2. HEARSAY II

This is a Speech recognition system designed at Carnegie-Mellon by Newell and Simon's group. It is widely considered to be one of the most influential AI programs ever written.[34] The importance of HEAR-

[32] Hubert and Stuart Dreyfus. *Mind Over Machine.* NY, NY: The Free Press, 1986. p. 115.
[33] Douglas Hofstadter and Daniel C. Dennett. *The Mind's I.* NY, NY: Basic Books Inc., 1981. p. 19.
[34] Ibid. p. 132.

SAY lies not in how well the system understands speech, but in the way that it is constructed – the idea of independent knowledge sources cooperatively solving a problem by posting hypotheses on a global blackboard data structure.[35] HEARSAY was programmed to play voice chess and it was the first system to recognize non-trivial connected speech.

2.5.3. MYCIN, PUFF, INTERNIST-1, GUIDION, CADEUS, and RECONSIDER

These are all expert systems relating to the medical field. They are based on heuristic rules extracted from experts. Cadeus was begun in the early 70s as Internist, a specialist in internal medicine. Mycin can diagnose infectious diseases. PUFF can diagnose lung disease. RECONSIDER is a diagnostic prompting system that works by encouraging doctors to consider alternatives and not jump to conclusions. GUIDION is an expert system developed as a teaching system for doctors.

The performance of MYCIN and INTERNIST-1 have been compared to that of human doctors. An evaluation of MYCIN was presented in the article "Antimicrobial Selection by a Computer". This

[35] Avron Barr and Edward Feigenbaum. *The Handbook Of Artificial Intelligence.* Reading, Mass: Addison-

article was published in the *Journal Of the American Medical Association* in 1979. MYCIN was given information concerning ten meningitis cases and asked to prescribe drug therapy. Its prescriptions were evaluated by a panel of eight infectious disease specialists. The experts rated as acceptable seventy percent of MYCIN's therapies.[36]

INTERNIST'S evaluation was published in the 1982 *New England Journal of Medicine* in an article entitled, "INTERNIST-1, an Experimental Computer Based Diagnostic Consultant for General Internal Medicine". INTERNIST was given nineteen cases to diagnose, each of which had several diseases, so that there were 43 diagnoses to make in total. INTERNIST's diagnoses were compared to that of clinicians at Massachusettes General Hospital and with case discussants. The result: Of the 43, INTERNIST missed 18, the clinicians missed 15, and the case discussants missed 8. The evaluation of INTERNIST was not very encouraging, It was concluded that the experienced clinician was vastly superior to INTERNIST-1.[37] Dreyfus discusses the idea that these medical expert systems in particular have merit if they are used along with a doctor, instead of in place of a doctor.

Wesley Pub. Co., 1981. p. 343.

[36]Hubert and Stuart Dreyfus. *Mind Over Machine.* NY, NY: The Free Press, 1986. p. 116.

[37]Ibid. p. 116.

2.5.4. Terry Winograd and SHRDLU (1973)

SHRDLU was design by Winograd as a simulation of a robot's arm that could recognize and move a set of blocks. SHRDLU also is able to communicate about his (its) blocks world using natural language. SHRDLU's knowledge is represented in the form of procedures – rules that can move blocks, stack blocks, and perform other operations in his (its) environment.

The major problem raised by critics of SHRDLU concerns his (its) world. He operates in a very narrow domain called a micro-world. A micro-world is a very thin slice of the world we live in. Dreyfus discusses this failure in his book *Mind Over Machine.* He quotes Winograd:

> The AI programs of the late 60's and early seventies are much too literal. They deal with meaning as if it were a structure to be built up of the bricks and mortar provided by the words....This gives them a brittle character, able to deal well with tightly specified areas of meaning in an artificially formal conversation. They are correspondingly weak in dealing with natural utterances, full of bits and fragments, continual (unnoticed) metaphor, and reference to much less easily formalizable areas of knowledge.[38]

Attempts to generalize the knowledge of SHRDLU were not very successful. SHRDLU's confinement to a micro-world gives him (it) no frame of reference to the outside world. He (it) does not know who he (it) is or where he (it) fits in the world. His (it's) world is separate and

[38]Hubert and Stuart Dreyfus. *Mind Over Machine.* NY, NY: The Free Press, 1986. p. 77.

isolated from human life. His (it's) memory is limited and his (it's) arbitrary acquisition and representation of knowledge is artificial.

2.5.5. Colby's PARRY

PARRY is a model of a patient suffering from paranoia. He conversed in a manner that a paranoid would. His creator, Kenneth Colby, claimed that PARRY was a true model of mind. Critics would of course challenge this. Colby claims that Parry passed two different versions of the Turing Test. The first test involved showing transcripts of PARRY conversations to experts. Weizenbaum, the creator of ELIZA, refuted these claims. "Weizenbaum claimed that by Colby's reasoning, any electric typewriter is a good model of infantile autism: type in a question and it just sits there and hums. No experts on autism could tell transcripts of genuine attempts to communicate with autistic children from such futile typing exercises!"[39]

In *The Mind's Eye,* Douglas Hofstadter discusses Parry's limitations.

> SANDY:
> An interesting thing about Parry is that it creates no sentences on its own – it merely selects from a huge repertoire of canned sentences the one that best responds to the input sentence.

---

[39] Douglas Hofstadter and Daniel C. Dennett. *The Mind's I.* NY, NY: Basic Books Inc., 1981. p. 469.

PAT:
Amazing! But that would be impossible on a larger scale, wouldn't it?

SANDY:
Yes. The number of sentences you'd need to store to be able to respond in a normal way to all possible sentences in a conversation is astronomical, really unimaginable. And they would have to be so intricately indexed for retrieval.[40]

2.5.6. Minsky Again Frames

Frames were originally proposed by Minsky (1975) for representing knowledge. They are based on the fact that when we recall the image of a particular object, we recall a group of typical attributes of that object. Minsky attempted to combine these attributes into a frame. For example, take the word chair. Minsky says that a chair frame would include certain expectations about chairs that we have in our minds. This includes things such as: chairs have four legs, a seat, a back, and two arms. You also expect that a chair would probably be used for sitting. All of this information would be contained in the frame.

There do exist some legitimate problems with frames. Referring to the above example, you can imagine a chair that does not have arms or legs. Think about a bean bag chair. Unfortunately there are always exceptions and special cases of every description of an object. This same criticism is raised in reference to rule-based systems. Even Minsky raises

this problem by suggesting that the rule: "All birds can fly" requires commonsense knowledge to be truly understood. A bird cannot fly if it is a penguin, or is dead, or has a broken wing, or is in a cage. How can these exceptions to the rule be explicitly stated so that a true picture of a bird can emerge?[41]

2.5.7. Roger Shank: Scripts, CYRUS, and FRUMP

At Yale University, Roger Schank and Robert Abelson were studying natural language understanding. They developed two programs, FRUMP and CYRUS. FRUMP stands for "Fast Reading, Understanding, and Memory Program". It can read and paraphrase news stories from United Press International wire service, indicating that it had understood the subject manner. CYRUS, or Computerized Yale Reasoning and Understanding System, had the job of monitoring all stories about Cyrus Vance, then secretary of state, and compile all facts found. Shank wrote,

> This was an attempt to begin to model the memory of a particular individual. In some sense, the program thought of itself as Cyrus Vance. When asked, during a dialogue about the Camp David negotiations, "Has your wife ever met Mrs. Begin?" CYRUS answered, "Yes, most recently at a state dinner in Israel in January 1980." The program actually had no specific information in its memory about the wives of Vance and former Israeli prime minister

[40]Douglas Hofstadter and Daniel C. Dennett. *The Mind's I.* NY, NY: Basic Books Inc., 1981. p. 60.

[41]M.Mitchell Waldrop. *Man-Made Minds.* USA: Walker and Company, 1987. p. 48.

> Menachem Begin. It was guessing. It figured that if it could find a situation when both women were likely to be present, then it could assume that they had met, Shank explained. The program thus searched for social situations.....Finding a state dinner in Israel that occurred during a trip where Mrs. Vance (accompanied) her husband, it assumed the rest.[42]

Another important contribution by Schank was the development of a theory of knowledge representation called Conceptual Dependency theory. Scripts were part of this theory. A script is composed of a series of scenes that represent the sequence of events we expect to encounter in a given situation. An example can be shown using a restaurant script. When you go into a restaurant you expect to see certain props such as tables, chairs, menus, plates, utensils, and food. You also expect to see certain people in certain roles: manager, waitress, and hostess. It is also assumed that when you go into a restaurant you are hungry, you have money, you will eat, and then you will have less money. The script will contain this information along with certain actions that could take place within this scene. These actions may include: enter the restaurant, look for a table, go to the table, decide where to sit at the table, and sit down.

---

[42] George Johnson. *Machinery Of The Mind.* Redmond, Washington: Microsoft Press, 1986. p. 171.

2.5.8. Ross Quillian and Semantic Nets

Quillian developed the concept of Semantic Networks intending them to be a psychological model of associative memory. These nets consist of nodes that represent objects, concepts, or situations in a domain, and arcs that connect the nodes, representing the relationships between them. For example, a net may include the statement, robin is-a bird., where robin and bird are nodes, and is-a is an arc. One advantage to this approach is that it is easy to add new nodes (For example, Clyde is-a robin is-a bird), so that deductions can be easily made concerning inheritance hierarchies.

2.5.9. Doug Lenat: Automated Mathematics (AM)

AM is a program that could learn. Lenat gave AM basic mathematical information such as how to determine whether two sets are equal and how to perform the inverse of a mathematical function. AM used this information to formulate new theorems and concepts in mathematics. It reportedly reproduced many mathematical theories, and also came up with a new theory concerning prime numbers.

Dreyfus contends that even Lenat admits that AM's abilities were a result of the special properties of the language LISP, used to program

AM.[43]

2.6. NEW AGE: 1980s - CONNECTIONISM and NEURAL NETWORKS

The most important aspect of this period in AI and its attempt to build intelligent machines is the resurgence of work done in the area of neural networks. Researchers were becoming more and more aware of the shortcomings of traditional AI and its rule-based logic systems. Many started to believe that traditional AI using production rules, predicate calculus, frames, or scripts, was on the wrong path. Connectionism, the attempt to model the mind in terms of biological plausibility, was reborn.

Neural Networks and Traditional AI are very different. Traditional AI programs use rules that operate on symbols based on syntax. Neural Nets do not use explicit rules. They are concerned with individual processing units that are defined in terms of activation thresholds and synaptic weights. To overcome the criticisms of the past, neural network researchers added a hidden layer to the net. In the connectionist model, knowledge is distributed across many processing elements.[44]

---

[43] Hubert and Stuart Dreyfus. *Mind Over Machine.* NY, NY: The Free Press, 1986. p. 213.

[44] Nadel, Cooper, Cukover, and Harnish. *Neural Connections, Mental Computations.* Cambridge, Mass: MIT Press, 1989. p. 2

2.6.1. Past Tense Of English Verbs: Rumelhart, Hinton and Williams

Influential to this period was the publishing in 1986 of *Parallel Distributed Processing* by McClelland and Rumelhart. Their idea is to program neuron-like networks that are inspired by the brain, but do not imitate it exactly. The emphasis is on the parallel processing of such a network, which would be much faster than traditional serial processing.

Rumelhart, along with Hinton and Williams, also developed a new program. This neural net learned to conjugate English verbs. It did not contain explicit rules as in a traditional AI program. Interestingly enough the net's mistakes were very similar to those seen in children during their learning of the past tenses of English verbs.

2.6.2. R. Pfeifer and D. Nicholas: A Model of Anger

This model is one of the few I have found besides Pacrat that deals specifically with emotions. This model of anger was carried out at Carnegie Mellon. The main goal of the project was to construct a computational model of emotions, representing how and when emotions are produced. This was done based on the theory that motive and emotion are major influences on cognitive behavior. Mandler (1975) influenced this research through his idea of the role of interrupts in emotion generation.

He argued that an increase in arousal follows the interruption of organized behavior. He believes that interrupts are crucial in multiple task environments where the real time information processing demands of an organism may be considerable. The interrupt facilitates the refocusing of attention, a necessity for survival of an organism.[45] The degree to which arousal increases after an interrupt seems to depend on two factors: surprise and importance. The more surprising and/or important the interrupt is, the larger the resulting increase in arousal will be. Also addressed in this project is the idea that previously generated emotions can be recalled along with the events that elicited them.

This model is designed as a production system and uses a semantic network for representing declarative knowledge. Procedural knowledge is encoded in the form of production rules. These rules have an "if then" structure with a set of conditions that match against the contents of working memory, followed by a set of actions to be executed, if the conditions have been satisfied. A typical action might involve working memory, long term memory, and a separate processing system that corresponds to the Central Nervous System.

The researchers describe three situations when emotions are generated in an organism: 1)when a well organized sequence or sub-

[45] L.Steel and J.A.Campbell. *Progress In AI.* West Sussex, England: Ellis Norwood Pub, 1985. p. 185.

sequence of behaviors has been completed, 2)When an interrupt occurs in such a sequence, and 3)as a result of explicit deliberation (for example, in the intentional recall of some past event that has emotional significance).[46] Also presented are three classes of interrupts: 1)those triggered by a stimulus in the environment such as a loud noise, 2)those triggered by physiological signals like hunger or pain, and 3)those triggered by the operation of information processing.[47]

To illustrate the emotion generation process an example is given in which an interrupt has occurred while the system is executing a command. The goal is to attend an important meeting in another city. In order to do so, the person must catch a plane leaving an airport in a few minutes. In order to catch the plain the person gets a taxi. While enroute to the airport, the taxi develops a flat tire. This generates an interrupt, and a link to the event that caused the interrupt is established in working memory. To calculate the increased arousal, which theoretically follows an interrupt, values representing surprise and importance are calculated using a mathematical formula.

At this point a set of production rules is invoked to determine the quality of the emotion to be generated. An example rule is:

---

[46]L.Steel and J.A.Campbell. *Progress In AI.* West Sussex, England: Ellis Norwood Pub, 1985. p. 186.

[47]Ibid. p. 186.

R1:IF state is negative for self

& state was caused by person1

& person1 was in control

& emotional target is person1

THEN generate anger at person1

A production rule such as this can only apply when all of its conditions match against events in working memory so that in order to fire R1, a number of sub-rules will have to apply first. Two examples of these are:

R2:IF an interrupt has occurred

& emotion is to be determined

THEN determine target for emotion.

R3:IF an interrupt has occurred

& emotion is to be determined

THEN determine if current state is negative for self.[48]

When all of its conditions are present in working memory, R1 is applied, and its actions result in the addition of several structures to working memory. The intensity of the emotion is also calculated using a formula. Long term memory is now updated by adding to it the elements currently in working memory.

Some rules in this emotional model are less general than R1, R2, and R3. Two examples of these are:

R4:IF my son kicks the cat

THEN generate anger at my son.

R5:IF self sees an event that may be described by the

expression "bully hurts victim"

& the emotional target is bully

THEN generate anger at bully.[49]

The researchers feel that their model has room for expansion. They believe that the model needs more knowledge of the world and better inference processes.

2.6.3. ART: Automated Reasoning Tool

This program was a creation of the Inference Corporation in 1984. It could write its own computer programs, and also assist human programmers with their writing. ART is known as a "knowledge-engineering environment" or "software-development toolkit" – an aid

---

[48] L.Steel and J.A.Campbell. *Progress In AI.* West Sussex, England: Ellis Norwood Pub, 1985. p. 188.

[49] Ibid. p. 189.

to designing expert systems.[50]

2.6.4. Common Sense: Cyc by Doug Lenat

Cyc is an attempt to give a program common sense knowledge. Now in its sixth year of development, Lenat claims that Cyc has the abilities of a 4 year old child. Lenat and other researchers have been feeding knowledge about the world to Cyc from newspapers and magazines. This information is stored in a frame-like construction. Cyc was given information about the world and about himself. He then meditates over this information and looks for redundancies or contradictions. Lenat hopes that Cyc will someday be able to read different publications himself, instead of having knowledge force fed to him by researchers.

One interesting story concerning Cyc relates to his understanding of the concept of "intelligent things". Cyc once asked, "All the others who work on this project are people. Am I a person?" Doug responded,"No." Cyc responds, "Is anyone else who works on this project a computer program? Are you a computer program?" Doug responds, "No."[51]

---

[50]George Johnson. *Machinery Of The Mind.* Redmond, Washington: Microsoft Press, 1986. p. 11.

[51]David H. Freedman. "Common Sense and the Computer". In *Discover Magazine.* Volume 11. Number 8. March 1991. p. 71.

Cyc was built in response to the problem of commonsense reasoning in AI. Dreyfus discusses this extensively in his book *Mind over Machine*. He quotes Minsky:

> Just constructing a knowledge base is a major intellectual research problem....We still know far too little about the contents and structure of commonsense knowledge. A "minimal" commonsense system must "know" something about cause and effect, time, purpose, locality, process, and types of knowledge. We need a serious epistemological research effort in this area.[52]

Lenat, of course, claims that he is making progress in this area. Others are sceptical. Even Minsky has criticized Cyc based on the fact that he uses only one method for representing knowledge, the frame. Minsky says that the systems of the future will have two or three different ways of representing knowledge with cross links between them.[53] In a sense, Cyc has what SHRDLU lacked, a sense of who he is and a knowledge of the world we live in. Lenat sees Cyc someday used for many things: in schools for one-on-one tutoring, in retail stores as a database to custom design products for individual consumers, as a scientist making discoveries, and even as a marriage counselor.

---

[52] Hubert and Stuart Dreyfus. *Mind Over Machine*. NY, NY: The Free Press, 1986. p. 140.

[53] David H. Freedman. "Common Sense and the Computer". In *Discover Magazine*. Volume 11. Number 8. March 1991. p. 70.

2.6.5. Rosenburg and Sejnowski - NETtalk

This neural network could read printed text and repeat it verbally. It learns to transform letters into certain sounds. At first the program does not perform well. Through successive iterations its performance improves. After ten hours of training on 1000 words, it produced coherent, intelligible speech given arbitrary English text.[54] The interesting thing about NETtalk, and neural nets in general, is that unlike traditional AI programs they do not require the use of specific rules for representing knowledge. However, these systems need an initial training process so that the program can learn its function.

2.7. THE 1990s

2.7.1. Attila the Robot

Attila was created by Rodney Brooks at MIT. He weighs 3.6 pounds has 10 microprocessors and 150 sensors. He has six legs and looks similar to a spider or insect. Because of his six-legged architecture he can move throughout a hostile environment quite well. He can climb over obstacles, pull himself onto ledges, and crawl up near vertical inclines. Brooks's Attila is not programmed in the same sense of

[54] Paul M. Churchland *Matter and Consciousness*. Cambridge, Mass: The MIT Press, 1984. p. 163.

traditional AI. He does not claim that Attila thinks the way we do or does anything the way we do.

In Brook's model there are basic procedures that are like primitive instincts. These include things such as Move Forward, or Track Prey. There is not one separate brain that processes everything; instead each procedure acts as a separate unit that competes for control. The winner is chosen based on sensory input. Attila has no representation of himself or his world. He can move around in his world, but he does not have any higher level functions such as recognition, reasoning, or thinking.

Marvin Minsky has criticized this approach because of this reason. He says, "Why bother building a robot that's capable of getting from here to there, if once it gets there it can't tell the difference between a table and a cup of coffee?"[55] Brooks says that he doesn't care if his robots are dumb. They are easy and cheap to build and do not require a lot of computing power to run. His goals for his robot are many. A miniature version could be used to clean carpets. A group of them could be used to construct objects. He would also like to see them travel the surfaces of the moon or Mars.

---

[55] David H. Freedman. "Invasion of the Insect Robots." In *Discover Magazine*. Volume 12. Number 8. August 1990. p. 46.

An interesting addition to Attila is being worked on by one of Brook's associates, Pattie Maes. Her specialty is the creation of programs that learn from their mistakes and modify themselves. She is attempting to give Attila motivations for his actions. These would include aggression, hunger, and curiosity. This would give Attila a personality.

3. EMOTIONS AND INTELLIGENCE

As we have just seen in the previous section, the field of AI is attempting to build intelligent, thinking machines. Human beings have this ability. We can think and reason. The ability to think and reason allows us to solve problems and to make decisions. This is intelligent behavior. The human brain, for this reason, is used in AI as a model for creating intelligent programs.

Human thought processes have some unique influences or co-processes associated with them. For example, there is certainly something special about the power of intuition or insight, and I don't think that anyone would be opposed to programming them into machines if it were possible. They contribute to doing things that we consider intelligent such as; solving a problem, writing a computer program, or planning our daily schedule of activities. Creativity is an example of another co-process that contributes to thinking with intelligence. Creativity produces new ideas, which is essential to any intelligent task.

If we take the approach of using the brain as a model of an intelligent machine, then perhaps we should determine the nature of these influences, and if, and how, they can be incorporated into computer models of the human brain. Within AI research there have been few models that have attempted this. There does exist however, is a generous amount of literature

about the nature of human thought processes including these co-processes.

Another influence on intelligent behavior that should be explored is the emotions. Humans are emotional beings. Our emotions are constantly present unless consciously repressed, something which is not always an easy task. In this thesis, I am suggesting that one of the missing components to building intelligent machines may be the emotions. In the first part of this section I will present evidence from psychological literature to substantiate my hypothesis regarding the relationship between the emotions and cognition or intellectual behavior. I am suggesting that emotions, like creativity, insight, and intuition may be necessary for true intelligence. What is needed then is an examination of this relationship and an investigation into the true nature of emotions. What functions do the emotions serve in a human being's life? Are these functions learned, innate, or both? The emotions seem to be able to get us to do things. How does this motivational mechanism work? What type of survival benefits are built in to the emotions? How do the emotions relate and react with perception? How do the emotions affect cognition? Can the emotions improve or degrade performance in intelligent tasks? These questions will be explored within the framework of psychological literature.

In the second part of this section I will discuss the ideas of Marvin Minsky concerning the relationship between intelligence and the emotions. He examines this relationship from the perspective of a mathematician, computer

scientist, and philosopher. He believes that the emotions and intelligence are intertwined.

3.1. PSYCHOLOGICAL THEORIES OF EMOTIONS

The topic of emotions in psychological literature is both varied in opinion and theory, and also somewhat neglected. To some, the concept of emotions is not eligible for a truly serious discussion or analysis. It is somehow irrational and not clearly defined. Those who have attempted to study emotions have proposed very diverse and interesting theories. There is not even agreement in this area of study about the definition of emotions or even a categorical list of emotions. From the research I have done, it seems that the emotions are not simple phenomena. In this paper I will discuss some of the theories that I feel are the most important and stimulating.

I will begin with a more general discussion first. There does seem to be some main commonalities among the diverse theories of emotions. The following list contains what seem to be the main components of many of the theories I have studied:

(1) Emotions may or may not entail some type of cognitive appraisal of something being in some way desirable or undesirable. This appraisal need not take place consciously in the individual.

(2) Emotions seem to contain a certain subjective, felt quality that is present in consciousness and available for study only through introspection.

(3) Emotions occur usually along with marked bodily sensations of certain kinds.

(4) Emotions also occur along with certain involuntary body processes and overt expressions of certain kinds, such as facial expressions.

(5) Tendencies to act in certain ways seem to be included in emotional responses.

(6) An upset or disturbed condition of body and mind seems to occur along with emotions. This condition can be seen as disruptive or organizing, depending on the situation.

(7) Emotions seem to play a role in learning.

(8) Emotions may have an effect on memory.

(9) Emotions may have survival benefits for an individual and a species.[56]

Theories of emotions differ as to which of these items they take to be the emotion itself and which they take to be causes or effects of the emotions. Some theories have a proposed time order for some of the points listed above. Other theories stress the importance of only some of the

[56]Alston William. "Emotion and Feeling". *Encyclopedia Of Philosophy Vol. 2* p. 479-486.

points, while not even addressing others.

3.1.1. Two Prominent Theories

Probably the two most notable theories of emotions are the James-Lange Theory and the Cannon Theory. The James-Lange theory was developed by William James in 1884 and Carl Lange independently in 1885. This theory states that bodily changes precede the feeling of emotion and that it is the perception of these bodily changes that is the emotional experience. To James and Lange, an emotional stimulus immediately causes us to respond with muscular and glandular reactions that characterize the emotion and the situation. Feedback from these emotional responses arrives at the cortex, and then we feel the emotion. Emotions, then, arise from sensations in the body's muscles and comes after a state of bodily arousal, and not before it. James and Lange emphasize the role of the autonomic nervous system in the expression of emotions. Crucial to this is the view that autonomic changes are communicated to the central nervous system which experiences these changes as specific emotions.

An experiment by W. B. Cannon using dogs produced a criticism of the James-Lange theory, resulting in a new theory of emotions. Cannon severed the connections between the autonomic nervous system and the central nervous system and found that dogs continued to express

emotional states. Cannon's main hypothesis is that the physiological state and the emotional experience are triggered simultaneously by the hypothalamus. This proposes a different time order for the emotional experience and the bodily response. Cannon's theory is a physiological theory that suggests how the functions of certain brain structures may determine emotional states. To Cannon, visceral changes cannot be distinguished from emotions. These changes are also too slow to be a source of emotional feeling.

3.1.2. Other Theories

3.1.2.1. Sigmund Freud

To Freud, our adult life is mainly a function of our early infantile development. Freud uses the terminology of arousal or cathexis when referring to emotions. This arousal is associated with three main systems: the oral or alimentary, the anal or eliminative, and the genital or sexual. Freud says that various aspects of these activities can be pleasant or unpleasant depending on the interpretation of the event in the context of our environment. He explains that we attempt to find states of environment that will result in pleasant experiences. He claims that this is how we learn about which events in life will lead to positive outcomes. This theme is definitely present in other psychological literature. Pain and pleasure are strong

motivating factors in our lives. They can facilitate learning and influence our actions.

> It is difficult to present a consistent treatment of affect in terms of classical psychoanalytic theory since Freud and his followers used the term rather loosely and prescribed different roles for it as the theory developed. In his early work Freud thought affect or emotion was the only driving force in mental life, and at various times in his later writings he still spoke of affect or emotions as the intrapsychic factors which give rise to fantasies, wishes, and symptoms.[57]

Freudian theory in general is primarily concerned with the negative affects arising from conflict, which then result in repression as a defense mechanism. This is related to Freud's idea of discharge. It is when this discharge is not allowed to take place properly that problems with our psyche can occur. This point is related to Freud's view on the evolutionary aspects of emotions, and his claim that humans have an innate drive towards aggression that serves the survival of both the individual and the species. According to Freud when this aggression cannot be expressed, it becomes dammed up and as a consequence, it is often inappropriately released.[58] To Freud, our emotions are instinctual, and serve the purpose of motivating our actions. It was work by Freud that opened up the field of motivation to psychology.

---

[57] Carroll E. Izard *Human Emotions*. NY, NY: Plenum Press, 1977. p. 21.

[58] George Mandler *Mind And Emotion*. NY, NY: John Wiley and Sons, 1975 p. 143.

Freud also recognized the importance of the arousal of the autonomic nervous system and an interpretation of the environment. He therefore stressed the importance of a cognitive appraisal of the situation and some type of automatic bodily response.

3.1.2.2. Charles Darwin

Darwin's theory of natural selection implies that almost every feature of each existing species has survival value, including emotional behavior. The emotions of fear and aggression (anger) to Darwin perform the function of promoting survival. Aggression, for example, facilitates the selection of the strongest animals of a given species. Fear also does this by serving as a warning signal to danger, providing rapid response, e.g., avoidance. Darwin also raises the importance of evaluation and emotions. He explains that an organism must recognize the beneficial or harmful aspects of its environment in order to survive. Emotions often accomplish this task. To Darwin, the fundamental emotions are governed by innate neural programs. This does not mean that no aspect of an emotion can be modified through experience. He says that people can learn to modify these innate emotional responses.

In Darwin's famous work, *The Expression Of The Emotions In Man And Animals,* he explains that there are some universal,

unlearned, cross cultural, emotional expressions, primarily facial in nature. He discusses the fact that certain emotions have the same expressions and experiential qualities in widely different cultures from virtually every continent of the globe, including isolated preliterate cultures, who have had virtually no contact with western civilization.[59] This expressive behavior can either be a consequence of emotions or a regulator of emotions. Regarding regulation, Darwin expressed the opinion that facial expressions, or other outward signs of an emotion can intensify the emotion, and that, in the reverse, the repression of all outward signs of an emotion can soften or repress it. To Darwin, these facial expressions resulted from the evolutionary process.

3.1.2.3. S. S. Tomkins

According to Tomkins, emotions constitute the primary motivational system of human beings. They play an important role in organizing and sustaining behavior. Tomkins stresses the importance of perception and cognition in initiating emotions, but also emphasizes a two-way interaction between cognition and emotions and the important effects that emotion can have on perception and cognition. He says, "Reason without emotion would be impotent,

emotion without reason would be blind".[60] To Tomkins, the intimate relationship between affect and cognition is rooted in the infant's early efforts at adjustment to its new and ever changing environment. He explains that this environmental novelty activates the interest affect which in turn sustains efforts of exploration. Thus, affective and cognitive processes are intertwined from the very beginning of life.

Tomkins also looked at emotions from the neurological level. He maintained that emotions are activated by changes in the density of neural stimulation. He defined this as the number of neurons firing per unit of time. Tomkins states that changes in neural stimulation, and hence activation of an emotion, can occur as a result of an innate releaser, a drive state, an image, or another emotion. He said that any of these could activate an emotion without an appraisal process. Tomkins says that emotions can be unlearned and that the emotional system is also capable of being instigated by learned stimuli. He also maintains that some emotions are consistently elicited by stimulation increase, some by stimulation decrease, and some by the maintenance of a steady level of stimulation.

---

[59]Carroll E. Izard *Human Emotions*. NY, NY: Plenum Press, 1977. p. 6.

[60]Ibid. p. 51.

Tomkins also noticed the special significance of the face, facial expressions, and facial feedback. He says that feedback from facial behavior, when transformed into conscious form, constitutes the experience or awareness of emotion. He also explains that specific innate programs for organized sets of facial responses are stored in sub-cortical centers. According to Tomkins, visceral responses play a secondary role in emotion, providing only background or accompaniment for the discrete expressions of the face. A specific emotion, a specific facial expression, and our awareness of that facial expression comprise the innately programmed subjective experience of emotion. Tomkins developed a detailed method for measuring facial behavior called FAST (Facial Affect Scoring Test). It explores the expressions of the brow and forehead, the eye-lids, and the lower face including the cheeks, nose, mouth, and chin.

Tomkins also stresses the importance of the emotion of interest. He says that the emotion of interest is as essential to intellectual development as exercise is to physical development.[61] He explains that interest enables individuals to sustain attention to complex objects. He says that the activation of interest at the neurological level involves a moderate increase in the density of neural firing.

---

[61]Carroll E. Izard *Human Emotions*. NY, NY: Plenum Press, 1977. p. 11.

Interest provides amplification of the physiological function required for prolonged and fatiguing work or play. Interest also plays a role in the development of competencies and skills and in the development of intelligence.

> Interest is not only a necessary support of perception but of the state of wakefulness. Indeed insomnia may be produced not only by disturbing negative affect but also by sustained intense excitement. Again, without interest the development of thinking and the conceptual apparatus would be seriously impaired.
>
> The interrelationships between the affect of interest and the functions of thought and memory are so extensive that absence of the affective support of interest would jeopardize intellectual development no less than the destruction of brain tissue. To think, as to engage in any other human activity, one must care, one must be excited, must be continuously rewarded. There is no human competence which can be achieved in the absence of a sustaining interest...[62]

3.1.2.4. O. H. Mowrer

Mowrer, like Tomkins, believes that emotions play a major role in organizing, sustaining, and motivating behavior. He stresses the role that emotions play in changes in behavior or performance which represent learning. He says that emotions are a prerequisite for learning. To Mowrer, negative emotions signal distress or pain and produce avoidance reactions. Positive emotions, on the other hand, signal success or reward. Mowrer explains that it is the inner subjec-

tive experience of emotion that provides the ongoing emotional state that modifies, controls, and directs behavior. Mowrer attacks the tendency in Western civilization to look upon the emotions with distrust and contempt and to elevate the intellect (reason and logic) above them. He says:

> If the present analysis is sound, the emotions are of quite extraordinary importance in the total economy of living organisms and do not deserve being put into opposition with intelligence. The emotions are, it seems, themselves a high order of intelligence.[63]

Mowrer also discusses the involuntary nature of emotions. He explains that emotions are always automatic and involuntary. He says:

> It is incorrect to blame the emotions for psychological problems since the emotions are not a result of willful or deliberate choice. We can choose to behave or misbehave, but we cannot choose or control emotions, directly or voluntarily. Given the appropriate (condition) stimulus, they occur automatically, reflexly.[64]

According to Mowrer it is mistaken to assume that the emotions are at fault in psychopathology. He feels that it is the cognitions and actions that are interacting with the emotions that are at fault. Mowrer's view is that behavior is a result of choice, or, in a broader sense, the result of character.

---

[63]Carroll E. Izard *Human Emotions.* NY, NY: Plenum Press, 1977. p. 229.

[64]Ibid. p. 3.

To Mowrer there are certain emotions that are more important than others. He explains that guilt, for example, develops as a function of the learning process. As a child is rewarded for good behavior, and punished for bad behavior, he gradually develops a sense of proper and improper behavior. Guilt, he explains, is actually a type of fear. It arises after performing a previously punished act. Anxiety also plays a role in learning and can, like guilt, prevent previously punished actions from repeating. Hope is an emotion that can instigate and direct behavior. He describes it as a conditioned form of positive reinforcement. Fear, to Mowrer, is a psychological warning of impending discomfort which one tries to avoid.

3.1.2.5. S. Schachter and J. Singer

These two researchers have developed a theory of emotions called the Two Factor Theory. They stress the importance of physiological arousal and cognitive appraisal. It is Schachter who began to look at the importance of cognition in defining the type of emotion felt during a state of physiological arousal. Arousal provides the turbulence, but cognition is necessary to label the emotion. Schachter refuted Freud's view of the negative aspects of emotions and stressed their organizing and constructive functions. He says that people are

[64]Carroll E. Izard *Human Emotions*. NY, NY: Plenum Press, 1977. p. 105.

always experiencing some type of emotion, and that there is no action without affect. Schachter has outlined three major points:

(1) Given a state of physiological arousal for which an individual has no immediate explanation, he will "label" this state and describe his feelings in terms of the cognitions available to him... .

(2) Given a state of physiological arousal for which an individual has a completely appropriate explanation..., no evaluative needs will arise and the individual is unlikely to label his feelings in terms of [any] alternative cognitions available... .

(3) Given the same cognitive circumstances, the individual will react emotionally or describe his feelings as emotions only to the extent that he experiences a state of physiological arousal.[65]

Singer and Schachter carried out many experiments trying to test their hypothesis. In one experiment, research participants were told that they were going to be testing a new kind of vitamin that could have an effect on vision. The drug that was actually given was epinephrine. It has the effect of producing the physiological state of arousal that normally goes along with an emotional experience. Some participants were told that they might experience some outward

effects of the vitamin including: shaking hands, heart rate increase and pounding, and a flushed face. These are true and accurate side effects of epinephrine. Other participants were given no information at all about the effects of the drug. The subjects were divided into two groups and members of each group were put into different situations. Members of one group were put into a room with a person who was aware of the true nature of the experiment and who behaved in a euphoric and happy manner. The other group was also put into a room with another person, also aware of the experiment, who acted angry. The informed individuals showed little or no tendency towards euphoria or anger. The uninformed participants in contrast were happy following the experience with the euphoric person, and angry when subjected to the angry person. The results meant that when individuals were physically aroused, their emotions depended at least in part on how they perceived the situation. They did not consider themselves to be in an emotional state until they were placed in a situation in which they were led into making the evaluation or appraisal peculiar to some emotion. The informed participants explained their arousal on the "vitamin" given to them. The uniformed participants explained their bodily sensations by adopting the mood of the confederate in the experiment. This showed that given

[65]George Mandler *Mind And Emotion*. NY, NY: John Wiley and Sons, 1975 p. 88.

the same state of arousal, the nature of the expressed emotion can be manipulated by manipulations of the social situation. Schachter says:

> Precisely the same state of physiological arousal could be "joy" or "fury" or any great diversity of emotional labels, depending on the cognitive aspects of the situation.[66]

Schachter then believes that the physiological arousal is the same for all emotions, and it is the interpretation of the arousal by the individual that results in emotional experience.

3.1.2.6. Carroll Izard

Izard's theory of emotions is called the Differential Emotions Theory. He believes that emotions are the principal motivational system and the personality processes that give meaning and significance to human existence. He describes emotions as complex processes with neurophysiological, neuromuscular, and phenomenological aspects. He says that the experiencing of emotion can arise independently of cognition. To Izard, the emotions are based on both learned and innate processes. They also effect perception and cognition in that emotions can alter them. He explains that perception is like a prediction or wager based on past experience. Our perception is colored by our wants, desires, and purposes. Our emotions act like a filter to modify

[66] Gregory Kimble. *Principles Of Psychology*. NY, NY: John Wiley and Sons, 1984. p. 326.

raw sensory data and as an organizer of this data. To Izard, emotions are non-rational and non-linear.

Izard also places importance on the role of facial expressions of emotions. They serve the purpose of regulation and control. He believes that emotions are both innate and learned. Izard believes there are ten fundamental emotions. These include:

(1) Interest and excitement. He says that these emotions have the capacity to facilitate intuition and creativity. They also provide much of the motivation for learning and the development of skills, analytical intellectual processes, and competencies. He says that interest results from an increase in neural stimulation, sometimes brought upon by change or novelty. The emotion of interest can restrict or focus attention, cognition, and perception so as to immerse the individual in a subject, and to free the mind of interferences. Daydreaming and fantasies help to amplify interest.

(2) Joy. To Izard, joy is a highly desirable emotion that along with interest guarantees that humans will be social creatures. Joy can also facilitate intuition and tactical knowledge.

(3) Surprise. This emotion results with a sharp increase in neural stimulation. It serves the function of clearing the nervous

system of ongoing emotion and cognition so that the individual can react appropriately to the new stimulus.

(4) Distress and Anguish. They serve the purpose of communicating to the self and to others, that things are not well, and then motivating the person to change the situation.

(5) Anger. This emotion serves a useful function in evolution. He says that anger can be justified when it provides the added source of strength necessary to respond to oppression or threat.

(6) Disgust. This emotion has the function of motivating behavior to change things that are wrong.

(7) Contempt. Izard says it is difficult to find a good use for this emotion in our lives. It does tend to motivate our actions to possibly change situations that we look upon with contempt.

(8) Fear. Izard says that fear is activated by a rapid increase in neural stimulation, brought on by imagined or real danger. It provides the motivation to escape from danger.

(9) Shame. Shame usually occurs in the context of relationships with other people. It motivates a desire to hide or flee. It is a powerful source of conformity. It serves as a guardian to self-respect. It can foster self corrective and self improvement behavior.

(10) Guilt. This emotion occurs in situations where one feels personaly responsible. Guilt can stimulate thought and cognitive evaluation of the wrongdoing.[67]

3.1.2.7. Magda Arnold

Arnold's approach to emotions is very unique. She stresses the importance of the subjective feeling of emotions, especially as the main tool in the study of emotions. She analyzes emotions based on the experiences of people and common sense. Arnold defines an emotion as a sequence of events described by the concepts of perception and appraisal. She says that emotions are a non-rational attraction or repulsion that follows upon the appraisal of something as good or bad for the perceiver. This appraisal is direct, intuitive, and immediate, following perception, and recognized only upon introspection. In Arnold's words:

> We can define emotion as the felt tendency toward anything intuitively appraised as being good (beneficial), or away from anything intuitively appraised as bad (harmful). This attraction or aversion is accompanied by a pattern of physiological changes organized toward approach or withdrawal. The patterns differ for different emotions.[68]

Arnold stresses not just the organizing effect of emotions, but the disorganizing effect that emotions can have on one's life.

[67]Carroll E. Izard *Human Emotions*. NY, NY: Plenum Press, 1977. p. 85-92.

[68]Magda Arnold *Emotion and Personality. Vol 1.* NY, NY: Columbia University Press, 1960. p. 182.

Emotional distractions are constantly present in our lives. They can keep us from doing work that needs to be done. If we are working on a difficult task, our attention can be drawn to something appraised as being more desirable. This appraisal leading to a strong desire is not necessarily final though. Arnold says we will reflect on the situation, weigh our alternatives, and the desire is either counterbalanced or reinforced by deliberate judgement. Although Arnold stresses the negative aspects of emotions, she also believes that there are positive emotions. These include; love and admiration, love of beauty, joy, mirth and laughter, and happiness.

Arnold has also studied the neurophysiology of emotion. Her idea is that the limbic system is the key to emotions and more specifically it is hippocampus which receives nerve impulses from the cortex, which then relays this information to the subcortex. As part of her complete research of emotions, Arnold has also studied the physiological changes that are produced in emotion. She explains that an emotional action initiates a pattern of neural impulses activating voluntary and involuntary muscles and impulses inducing autonomic and endocrine changes.

---

[60]Magda Arnold *Emotion and Personality. Vol 1.* NY, NY: Columbia University Press, 1960. p. 31.

3.1.3. Summary

Many other theories of emotions are also relevant to our current study. Some important researchers are Mandler, Leeper, McDougall, Plutchick, Lazarus, Papez, McClean, Buck, Lyons, and Bower. There is also discussion of emotions in philosophical literature to be explored. This literature includes writers such as Aristotle, Aquinas, Hobbes, Kant, Hume, Descartes, and Locke.

In this section I have tried to present evidence in regards to my hypothesis concerning the emotions and intelligence. Tomkins, for example, stresses the two-way interaction that exists between emotions and cognition. Emotions affect cognition, and cognition affects emotions. He stresses the importance of emotions in organizing behavior and he discusses the importance of the emotion of interest in intellectual development.

Mowrer's idea on the necessity of emotions in the learning process could imply that an intelligent machine, who of course must have a learning component, must then have an emotional component. Like Tomkins he feels that the emotions have the function of motivating and directing behavior. He says that the emotions are a type of intellectual process themselves.

Schachter says that there can be no motor activity without affect. (This is exactly what Scanlon and Johnson have proposed and used in their design of Pacrat, which will be presented in the next section).

Izard's ideas are also very supportive of my hypothesis, especially those relating to emotions and creativity, intuition, cognition, and perception. He feels that emotions have the ability to facilitate intuition and creativity – two concepts that I feel are involved in human intelligence and reasoning. He also feels that emotions have the function of altering perception and cognition.

It seems that if we are to build machines that share at least some of the capabilities of human intelligence, then we must study psychological theories that explain how human intelligence is developed and maintained and what types of processes interact with it. In summary, emotions are one of these processes. They directly affect our intellectual processes of reasoning, thinking, and problem-solving.

3.2. MARVIN MINSKY ON THE EMOTIONS

Marvin Minsky wrote, "The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions."[70] Minsky defines intelligence as processes in our minds that enable us to solve problems we consider difficult. These

[70] Marvin Minsky *The Society Of Mind* NY, NY: Simon and Schuster, 1985. p. 163.

processes, however, are not well understood. The concept of emotions is also something that is not well understood. Minsky says that although many people believe that thoughts and feeling are very separate and different things, they are nevertheless very much related; feelings are like types of thoughts.

Minsky believes that intelligent behavior is goal directed and involves problem solving. He also believes that intelligence is influenced by our emotional states. This influence comes in many forms and works in different ways.

3.2.1. Background Information

Minsky believes that people are born with some type of basic, instinctual needs. He calls these "proto-specialists". They include goals such as thirst, hunger, warmth, and defense. We accomplish these goals through the use of our body, including the sense organs which carry out our human activities. Instinct, however, is not the whole picture. Minsky says we create new goals, besides the instinctual ones, throughout our whole lifetime. These new goals come about through many years of elaborate concept formation and life experiences. Thus, our goals are a combination of the instinctual and the learned.

Minsky feels that the mind needs some type of built in constraints, and checks and balances to guide us towards these goals. These checks

are also used to keep us from following goals that lead to self-destruction. To Minsky, pleasure and pain are concepts that can accomplish this. He sees them as being very powerful. They can undermine one's interest in anything that is not immediate. They simplify our point of view. Pain can distract us from dangerous goals and help us to survive. Minsky says that we seek behavior that rewards us with pleasure, and avoids that which brings us pain. Minsky sees emotions as working with the intellect to attain goals. I will now explain specific examples of how these two concepts interact.

3.2.2. The Interaction Of Emotions and the Intellect

3.2.2.1. Emotion's role in motivation and self-control

In trying to achieve a goal, we must keep working even while under duress; and we must not be easily distracted. We also must not lose sight of our goal and the intermediate steps of progress we make along the way. Minsky says that in order to do this we use weird things to control ourselves and keep ourselves motivated. These things include fantasies, and in some cases even lies. Minsky asks, Why can't we just tell ourselves to do what we want to do? He explains that that if we were easily able to switch from one goal to the next too rapidly, we'd never get anything done. If we could just switch off hunger and pain, we would have serious problems.

Therefore, we must have checks and balances to keep us working towards our goals, while using our intellect.

To explain the above point, Minsky uses the following example. Suppose a man is working on a very difficult problem. Although he is tired and becoming bored, he conjures up a fantasy of a challenger working on the same problem, who is getting close to completion. The man imagines this, and then feels an angry wish to frustrate this challenger. This forces him to continue working. The man is using emotions in order to accomplish an intellectual purpose. He may not be able to make himself angry, simply by deciding to be angry, but he can still imagine objects or situations that make him angry.

Tricks such as these to continue being motivated and to control our behavior are used to reward ourselves for an accomplishment. Some examples of these tricks are:

(1) The use of statements such as,"If I get this done, I'll have more time for other things."

(2) The use of willpower by telling ourselves, "Don't do that."

(3) The use of activity like moving around and exercise.

(4) The use of expressions such as: set jaw, stiff upper lip.

(5) The use of coffee or drugs.

(6) The use of emotion with statements such as, "If I win there is much to gain, but more to lose if I fail. "If I get this done, I'll have more time for other things." And imagining admiration upon success or disapproval upon failure. Thus, according to Minsky, we make ourselves behave and control our motivation by exploiting our own fears and desires, or by offering ourselves rewards, or by threatening loss of what we love.

3.2.2.2. The involvement of emotions in intellectual application and development

Minsky says that confusion and pain stimulate us to apply our intellect. To explain this he uses the example of a man who, through injury, develops problems walking. In order to walk now, he must think about walking, and now use his intellect for something that was once automatic.

Minsky asks why emotions are supposedly harder to explain than intellect? Many people view emotions such as anger as non-rational. If we look, though, at the challenger example discussed earlier in this section, the use of anger by the man can certainly be seen as a conscious rational process.

Minsky says that intellectual development can depend upon attachments to other people whom we want to please, and upon fear

and dread of disappointing those we look up to. We can make emotional/ intellectual attachments to people, and want to feel/think the way certain people do. Thus, emotional attachments can inspire us to improve and develop our intellectual capacities.

3.2.2.3. Emotions and Learning

Minsky says that people do not learn well unless they are interested or concerned. We learn from our mistakes, and we also learn from our successes. He says that every machine that has the capacity to learn must have some type of protective scheme, or otherwise, it could get trapped endlessly repeating the same activity. This mechanism keeps us from wasting time.

We also need some type of controlling mechanism over our pleasure systems that might try to reach beyond the limit of ordinary pleasure or reward. To explain this idea, Minsky gives the common example of eventually becoming bored with a task we once thought was pleasant. It seems as though we need a variety of purposes and goals, and through the satisfaction of these goals we learn and become intelligent. In summary, our pleasure systems help us to learn new things, by requiring a variety, so as to eliminate boredom, making us investigate new concepts and ideas.

3.2.2.4. Emotions and the Process of Transferring Goals and Values

Minsky says that brains transfer values and goals by exploiting personal relationships we call emotional such as fear, affection, love, hate, detachment, and dependency. Values and goals of a given culture pass from one generation to the next. They are not learned the way other skills are learned. We learn our earliest values under the influence of attachment related signals that represent not our own successes or failures, but our parents' love and rejection.[71]

3.2.2.5. Emotions as warning signs

Minsky says that to serve as useful warning signs, feelings such as pain and hunger must be engineered not simply to indicate dangerous conditions, but to anticipate them and warn us before too much damage is done.[72] He says that feelings such as pain or depression are products of our minds' activities that are engineered to warn us before we reach our limits. As an example, Minsky uses the idea of the feeling of depression and discouragement that one gets when stuck in a boring job. These feelings help us to realize that maybe we should evaluate our situation and replace or alter our current goals.

[71] Marvin Minsky *The Society Of Mind* NY, NY: Simon and Schuster, 1985. p. 181.

[72] Ibid. p. 286.

Another example of this can be seen in babies. Minsky says that infants are single minded. They seem to be usually in one of two states that can switch suddenly. Babies do not display a mixture of emotions, but express only one emotion at a time. He says that this is done to increase the infant's chance of survival, since with mixed emotions it would be even more difficult to determine what the baby needed or wanted. The child's cries can be seen as warning signs to the parent, and so require clarity.

3.2.2.6. Choices, Comparisons, and Compromises in Problem Solving and Emotions

Emotional ideas such as like, prefer, and enjoy, are a sort of measurement. They reduce things to a value. These judgements seem to be similar to heuristics used in problem solving in artificial intelligence. Minsky says that the function of liking is to turn off alternatives. If unconstrained, it narrows down our universe.

Minsky believes that in problem solving we rarely need to know that anything is absolutely right or wrong. Instead we just want to make a choice between alternatives. Intelligence, then, includes judgement which is involved in more than just determining right or wrong, or true or false.

3.2.2.7. Miscellaneous Thoughts

The idea of creativity is related to emotions. It is difficult to define creativity in terms of logic or algorithmics. Minsky feels that creativity is needed for intellectual behavior. Creativity requires the use of imagination. We use imagination to acquire new ideas. We can imagine things we have never seen before, or even things that do not exist. The use of fantasies, emotional or not, is indispensable for every complicated problem solving process.[73] Minsky is doubtful that we use logic to solve problems or create new ideas. He says we might use logic in the later stages of problem solving, to formulate ideas conceived in other ways.

The other idea of interest raised by Minsky is the concept of pretending to be in a certain emotional state to accomplish a goal. Sometimes people threaten to develop certain negative emotional states to get a certain result. Examples of these are seen in many relationships. Our parents tell us that if we act in a certain way they will become mad at us. People who are manipulative sometimes pretend to be angry or pleased to get what they want.

In summary, Minsky sees humans as goal driven organisms with processes that help us carry out our goals. These goals start with the

---

[73]Marvin Minsky *The Society Of Mind* NY, NY: Simon and Schuster, 1985. p. 163.

instinctual and are also acquired throughout one's lifetime. Emotions are just one of the processes that help us to live in a positive direction for our being. Cognition or the intellect is also an example of one of these processes. Cognition and emotions, along with other processes, exist in cooperation to help us accomplish our goals.

4. PACRAT: NON-LIVING INTELLIGENCE

In the first part of this section I will present the work of two researchers: Mark Johnson and Ray Scanlon. They have developed a computer simulation of non-living intelligence, a "being" called Pacrat. Pacrat is the first stage of an attempt at reverse engineering of the brain. Pacrat displays intelligent behavior under the influence of emotions. These researchers are assuming that emotion/motivation must be present in every intelligent being. Pacrat is said to get angry, display intuition and curiosity, and to have an awareness of his world and surroundings.

In the second part of this section, I will discuss the programming component to this thesis. I have rewritten Scanlon and Johnson's FORTRAN version of Pacrat in C. Statistical results and source code will be presented.

4.1. THE IMPLEMENTATION OF SCANLON AND JOHNSON

4.1.1. Let's Meet Pacrat: An Introduction

The main hypothesis of my paper is that emotions and intelligent behavior have an interdependent relationship. This implies that those in the field of Artificial Intelligence who are attempting to build intelligent machines should look at this relationship and incorporate it into their models. This is exactly what has been done with Pacrat.

Pacrat is the work of two researchers, Mark Johnson and Ray Scanlon at Benet Laboratories, for the U.S. Army. He is one step in their long term goal of creating an autonomously piloted combat vehicle that must be aware of its surroundings, able to learn, able to make complex decisions, and in essence, be able to think like a person. Scanlon and Johnson write:

> There is no such thing as an intelligent machine without a thinking machine. There are cleverly programmed machines, but there are no clever machines without a thinking machine. We believe the only feasible approach to intelligent machine design is through reverse engineering of the mammalian brain.[74]

Pacrat is a computer simulation of a primitive, mammalian organism. One important aspect of their work, which is closely related to the topic of this paper, involves their ideas concerning the relationship between thinking and emotion. Scanlon and Johnson point out that:

> Feeling is essential because without motivation, there is nothing. The machine must want to do things. In doing things, it will learn; and having learned, it will think.[75]

According to these researchers, there is no meaningful motor activity without some form of motivation. In order to do anything like a human, including thinking, they feel a machine must be motivated like a human.

---

[74]Raymond Scanlon and Mark Johnson. "Synaptogenesis". Watervliet, NY: US Army Armament Research Center, 1990. p. 5.

[75]Raymond Scanlon and Mark Johnson, *"Machine Emotion". IEEE Spring Compcon.* Wash, DC: Computer Society Press Of the IEEE, 1988. p 403.

4.1.2. What is Pacrat Really Like?

4.1.2.1. Pacrat And His World

Pacrat's world is a 21 x 21 grid on a computer screen. One location on the grid, row 11 column 1, is his home. Food can be found in two other alternating locations: row 3 column 18, and row 11 column 18. Each location on the grid has four weights associated with it that represent Pacrat's probability of moving to the next location north, south, east, or west. Paths to food and home are reinforced through the adjusting of these weights by Pacrat's reward system. Pacrat's primary motivations are hunger and fear. When hungry, he leaves his burrow and sets out to search for food. It is during this search that Pacrat has many experiences, including various forms of thinking and feelings. After eating, Pacrat is motivated by fear of open places, agoraphobia, and scurries back to his home. There Pacrat will sleep until he becomes hungry again.

The intent is for Pacrat to behave in his world as a mammal would, and so he is created with some of the capabilities and emotions of a living animal. He has the ability to sleep, to be awake, to move, and to eat. He has a stomach, a nose, and a sense of location. Pacrat also displays anger, curiosity and boredom.

Scanlon and Johnson have provided Pacrat with three distinct forms of thinking.

(1) The type of thinking which occurs when sensory input is temporarily blocked, and motor input is inhibited. Pacrat has been given the ability to evaluate the consequences of his last move, and also the awareness that a current action may have important connotations. This is called association.

(2) The capacity of recognition, which is learned from previous successes and rewards. This occurs when Pacrat arrives at a location where he has found food in the past.

(3) The ability to think of more efficient paths to food. Scanlon and Johnson claim that this is insight, which to them is the most basic mechanism of rational thought.[76]

4.1.2.2. Pacrat's Anatomy

Pacrat is a model of a mammalian brain. His brain has eight components. These include the ascending reticular activating substance (RAS), the thalamus, the hypothalamus, the amygdala, the cingulate gyrus, the median forebrain bundle, the hippocampus, and the isocortex. A block diagram of Pacrat's brain is included at the

---

[76] Raymond Scanlon and Mark Johnson, *"Machine Emotion"*. *IEEE Spring Compcon.* Washington, DC: Computer Society Press Of the IEEE, 1988. p. 405.

end of this thesis, and details of these brain parts will be discussed in a later section. These centers are modeled based on their functional relationships. The attempt is not to model the response of individual neurons, but only the activity of groups of neurons which Scanlon and Johnson call codons. A codon is a record or result of an experience.[77] More specifically it is a vector that represents the state of Pacrat at a given instant. The variables that make up this state are: 1) the number of a sensory neuron, that is 1 - 441, each being a location on the grid, 2) his dominant emotional state, 3) the values that make up the state of the amygdala and thalamus, (These are the amygnt, which is active when Pacrat is eating. The thalgl, which is also active when Pacrat is eating. The thalnt, which is active when Pacrta is thinking. The thalbd, which is active if Pacrat is thinking and arrives at a food spot.), 4) the number of the previously active codon, that is, the location Pacrat was at before, and 5) the time she arrived at the current location. On a lower level of anatomy, this state could represent a specific configuration of active neurons in the brain.

Pacrat also has four motor neurons that are controlled by the isocortex. Pacrat receives sensory input from the environment. This

[77] Raymond Scanlon and Mark Johnson, "Machine Emotion". *IEEE Spring Compcon.* Washington, DC:

input is relayed to the isocortex under the influence of the dominant emotion.

4.1.2.2.1. Details Of Pacrat's Brain

4.1.2.2.2. The Hypothalamus

The hypothalamus is the area of Pacrat's brain that deals with hunger. Pacrat's stomach has sensory neurons lining it that signal the activity of the stomach. The activity of these neurons is represented by a scalar that continuously decreases until Pacrat is eating. As Pacrat gets hungry his hunger center becomes more active; after he has eaten, the activity quiets down. The hypothalamus also deals with Pacrat's anger. In this case certain codons in the isocortex that have been influenced by the amygdala become active. This model is sound in that in humans the hypothalamus is also thought to be the seat of hunger and involved with anger.

4.1.2.2.3. The Cingulate Gyrus

Fear is defined as the level of activity present in the cingulate gyrus. When Pacrat is out of his burrow, and his hunger is satisfied, sensory neurons excite the cingulate gyrus and produce

Computer Society Press Of the IEEE, 1988. p. 403.

agoraphobia, fear of open places. Pacrat must get home in order to get relief. Doing so activates his reward system. Some literature has shown that the hypothalamus is involved in fear. See Arnold (1960). In humans the cingulate gyrus can inhibit or facilitate action depending on the appraisal of such action as appropriate or inappropriate.

4.1.2.2.4. The Hippocampus

The hippocampus is the area of Pacrat's brain that produces curiosity. It begins by the activation of a new codon. A record is kept of Pacrat's previously active codons, and an unsuccessful search determines newness. For example, if Pacrat is at sensory neuron 50 for the first time, is afraid (dominant emotion 2), and has come from sensory neuron 49, this information, along with the four values that represent the state of her thalamus and amygdala (let's say they are all 0) and the present time, are the values that make up her currently active codon. This information is recorded for future use. This codon is considered to be new. If later in the trial he again moves to neuron 50, but is hungry and/or comes from neuron 51, this is also considered to be a new codon and is recorded. If in later trials Pacrat again arrives at neuron 50 from neuron 49, is afraid, and her thalgl and amygnt are both 0, this

would be considered a reactivation of an old codon. In this case the time of this codon is changed to the present time, and no new record is kept.

Pacrat's hippocampus affects his motor activity so that something new encourages him to explore his world. In humans, the hippocampus is involved in memory. It allows us to recall things that have been previously experienced, and also whether the experiences were liked or disliked. It also may be involved with a period of hesitation or deliberation that sometimes precedes action in humans.

4.1.2.2.5. The Thalamus

The functions of Pacrat's thalamus include receiving all sensory input except olfactory, and the management of Pacrat's thinking. The thalamus can relay this input to the isocortex or block it altogether. If sensory input to the isocortex is blocked, Pacrat experiences thinking. His motor activity is inhibited, but he thinks about moving to other locations on the grid. If the thalamus does not block the sensory input to the cortex, Pacrat will move to a new location on the grid. It is speculated that the thalamus in humans mediates the sleep state. It can increase or decrease the excitability of the cortex, and can inhibit motor and

sensory areas.

4.1.2.2.6. The Amygdala

In humans, the amygdala is thought to be the seat of imagination. It is thought to work closely with the hippocampus in imagining sensory experiences and what might or can be done about them. In Pacrat anger is experienced through the hypothalamus and isocortex, under direction of the amygdala. When Pacrat is searching for food, and he does not find it in a familiar location, Pacrat becomes angry. The amygdala becomes active, and forces Pacrat out of that location, and into new random locations where food may be found. This is accomplished by temporarily making each of Pacrat's weights equal. These weights represent the probability of his moving from his current location either north, south, east, or west. In *Emotion and Personality* Arnold points out that:

> As soon as a man appraises something as annoying, and feels anger (whether he expresses it or not), he immediately imagines what to do to fight effectively, even if all he does is use bad language. Such motor fantasies seem to be mediated via the amygdala and the frontal association areas, and reinforce the anger and attack as well as directing it.[78]

[78] Magda Arnold, *Emotion And Personality*, NY, NY: Columbia University Press, 1960. p. 188.

4.1.2.2.7. The Reticular Ascending Substance - RAS

The RAS in humans is thought to maintain the waking state and act as a relay during the appraisal of sensory input. In Pacrat it becomes excited whenever the hypothalamus or the cingulate gyrus is active. This means that if Pacrat is hungry or afraid the RAS becomes excited. If Pacrat is neither hungry nor afraid he just goes to sleep as the RAS becomes inactive.

4.1.2.2.8. The Medial Forebrain Bundle - MFB

In Pacrat the MFB serves as a reward-punishment mechanism whenever activity in the hypothalamus or cingulate gyrus is reduced. If Pacrat has reached food in a number of moves that is less than a running average, he is rewarded. Going back 30 moves, the weights of Pacrat's previous locations are adjusted so that this path to food will be preferred by Pacrat in the future. The reward is highest for the most recent codon and is less for those codons that have been active farthest in the past. This is considered to be learning, and will cause Pacrat to search for food in locations where it has been found before. This mechanism works the same when Pacrat is scared and finds his way home. The weights on his path are adjusted to reinforce this path for future trials. Because Pacrat's current state and the weights associated

with it can be seen as an assemblage of active neurons, Scanlon and Johnson explain that in the Pacrat simulation this serves the purpose of reinforcing recently fired synapses.[79]

4.1.3. Why a Thinking, Feeling Machine?

4.1.3.1. We Think With Feeling

Scanlon and Johnson are attempting to create what they call non-living intelligence. They want to duplicate the functions of the brain in electronic circuits. They point to the existence of the brain as proof that a thinking machine exists, and then consider the possibility that it can be duplicated.

The researchers hypothesize that there are three main processes involved in the life of a organism that has a brain. These are motivation, cognition, and motor activity.

> The most important constraint on our work is that without motivation-emotion there is no cognition, no motor activity. If we would have a thinking machine, then we must look to a feeling-thinking machine – for there is no other.[80]

These three components are all present in Pacrat. He wants to do things, he does them, and he thinks. The function of emotion and motivation are to guide the organism through its world.

---

[79] R. Scanlon and M. Johnson, "Machine Emotion". *IEEE Spring Compcon.* Wash., DC: Computer Society Press Of the IEEE, 1988. p. 404.

[80] R. Scanlon and M. Johnson. "Experiences With A Feeling-Thinking Machine". *1987 IEEE International*

4.1.3.2. Thinking and Feeling are not Rule-Based or Logical

In order to build a Pacrat with emotion-motivation and thinking, Scanlon and Johnson have rejected the more traditional methods of AI: expert systems, rule-based programs, and predicate calculus. These researchers feel that these techniques have failed thus far to produce intelligent behavior or thinking, especially human intelligence and human thinking. Scanlon and Johnson feel that Pacrat, like humans, must think in an non-rational manner, with emotion. They reject the notion that humans only reason with formal logic. This opinion has also been expressed by Minsky in *The Society Of Mind* as discussed earlier in this thesis. He feels that logic has been shown to be unreliable. Rules to be used in the process of reasoning are difficult to state explicitly and entirely. Minsky asks,

> Why do people think emotion is harder to explain than intellect? Is anger always non-rational? Why is it a popular view that emotions are inherently more complex and harder to understand than other aspects of human thought?[81]

To Scanlon and Johnson, expert systems do not deal well with uncertainty or the unexpected. This would be a serious drawback in any kind of war situation, or in every day life situations as well. This is another reason why they have rejected the expert system

Conference On Neural Networks. Wash, DC: Computer Society Press Of The IEEE, 1987. Vol 2. p. 72.

[81]Marvin Minsky. *The Society Of Mind.* NY, NY: Wiley and Sons, 1985. p. 172.

approach, and have embraced the techniques of neural networks and human brain modeling. There are some in the field who agree with this view that neural nets may be the answer to recreating a mind and that "traditional AI" has failed. This view is held by Dreyfus as explored earlier in this thesis. There also are others who believe that neither neural nets nor any fancy technological trick will ever produce a truly intelligent being.

4.1.4. The Future of Pacrat

The version of Pacrat described above is actually step two in an on-going process. The initial program was called "mouse". This program displayed learning similar to Pacrat, but it could not think, and was not motivated by emotion. Presently Scanlon and Johnson are investigating the neuronal activity of the brain instead of just the functional relationship between its parts. They are also subdividing Pacrat's eight brain centers to into smaller anatomical parts, so as to eventually complete a reverse engineering of the human brain.

4.1.5. Final Remarks

According to Scanlon and Johnson, if the brain can perform the functions that we are trying to produce, then we should put our efforts towards modeling and recreating it. Thus, to these researchers, the first

step in creating an intelligent organism is to supply it with motivation and emotion, the ability to think, the ability to learn, and the ability to get around in the world. Intelligence will come when these capacities are put in place.

4.2. MY IMPLEMENTATION OF PACRAT – MS. PACRAT

4.2.1. Program Introduction

For the programming component to this thesis I have taken Scanlon and Johnson's Pacrat program and reprogrammed it in C. Their program was written in Fortran IV and ran on an IBM 4341. My C program runs on an AT&T 3B2/600 and an AT&T 3B1. I have also written a graphics program that displays Ms. Pacrat moving on a grid. This graphics component uses the 3B1 graphics display. Ms. Pacrat has also been run on a VAX system running ULTRIX. Because of its speed, the VAX machine allows for a more efficient observation of Ms. Pacrat's behavior after hundreds of trials.

4.2.2. The Coding of Ms. Pacrat

The program I received from Scanlon and Johnson was an uncommented hard-copy listing. Unfortunately this listing was the only thing that they could find from the old version of Pacrat. No electronic copies were saved and I was not guaranteed the the listing I received was a

working program. I began my coding using a strictly syntactical translation from FORTRAN to C. Athough this was my first experience with the programming language FORTRAN, I did not find the translation process to be too difficult, and I now feel I have a good working knowledge of a new language. Although I had a basic idea of what the program was supposed to accomplish, the fact that the listing was uncommented made it quite difficult to understand by just reading the code. Once I had a running C program I began a massive debugging campaign. I first investigated the nature of the major data structures of the program by printing them out and examining their values. Eventually I began to understand how Pacrat works.

During this process I was able to communicate with Ray Scanlon, who works in Watervilet, N.Y., through the ethernet. Ray answered many questions for me and was of great help during difficult times. He is a very interesting man, with many years of experience in the field of computer science. He also considers himself to be a philosopher, allowing for stimulating discussion.

4.2.3. Experiences with the Pacrat program

One of the most interesting aspects of writing this program concerns the major data structures of the program. FORTRAN does not support dynamic storage allocation, and so does not have pointers. It

also does not have structures as in C. This requires the use of multi-dimensional arrays in place of C structures. For me, this was a new experience.

Another aspect of the program that was quite different from my previous programming experiences concerns the nature of local variables in FORTRAN. Again, because of the lack of dynamic allocation, local variables to a subroutine retain their values when the subroutine is called again. At first I was unaware of this fact, which caused me confusion and a misunderstanding of many variables. Once I realized this, I was able to make great strides in my program translation.

4.2.4. Program Results

Included at the end of this thesis are graphs showing Ms. Pacrat's learning. The x-axis shows the trial number and the y-axis shows the average number of moves it takes Ms. Pacrat to get to food or to home for each 10 trials. The results are shown for the paths to each food spot separately and the path from each food spot to home.

Ms. Pacrat tends to learn one food spot quicker and better than the other. After observing many trial runs, it seems her tendency is to learn the closer food spot first, row 11 column 18. She does however have runs where she learns the food spot at row 3 column 18 first. When she learns a food spot, Ms. Pacrat's tendency is to continue to go

to this spot even though food is not there. The function of her amygdala is to combat this behavior. Interestingly enough, Scanlon mentioned this behavior to me. He says that Ms. Pacrat is exhibiting obsessive compulsive behavior.

4.2.4.1. Graph Specifics

Graph 1 shows the average number of moves it takes Ms. Pacrat to get to the food location at row 11 column 18. This is also referred to as food spot 2. For early trials I have seen averages as high as 1000 moves. As the graph shows, this number decreases so that by 1000 trials his average number of moves to this food spot ranges from 30 to 20.

Graph 2 shows Ms. Pacrat's results for food spot 1 at row 3 column 18. The average number of moves decreases, but tends to bottom out at approximately 80 moves.

Graph 3 shows Ms. Pacrat learning her way from the food spot at row 3 column 18 to home. This task is accomplished easily by Ms. Pacrat after 1000 trials. Graph 4 shows Graph 3 in sharper detail, setting the y-axis at 300 instead of 1000.

Graph 5 shows the average number of moves it takes Ms. Pacrat to go from the food spot at row 11 column 18 to home. These results are very similar to Graph 3. Ms. Pacrat has no problem

learning her way home from either food location. Graph 6 shows Graph 5 using 300 for the y-axis.

Graphs 6 and 7 show the average number of moves it takes Ms. Pacrat to go to food and back home again without making a distinction between the two different food spots. The average to food is higher than that to home, since in this case the average number of moves to both food spots are combined.

4.2.4.2. Remaining Visual Displays

4.2.4.2.1. Figure 1

The first display shows Ms. Pacrat arriving at a food spot and eating during trial 1. Ms. Pacrat's location is neuron 368, food spot 2. Her dominant emotion is hunger (2). Notice that eating raises the level of the variables amygnt and thalgl by .1. Also notice where this information is now inserted into Ms. Pacrat's codon. Here you see the active codon 368, dominant emotion hunger (2), the values of the amygnt and thalgl in packed form (110), the previously active codon (347) and the time of activation of this codon denoted by ibgcnt (682). To get to this food spot, Ms. Pacrat's last two moves were from neuron 346 to 347 to 368. You will notice that the move partion shown is actually for codon

347, since this was the last move taken. The first section of the partion (north) is small because Ms. Pacrat came from 346 which is north of 347, and so she is being discouraged from moving north back to 346.

The second display shows Ms. Pacrat arriving at home, afraid. Her location is neuron 11 and her dominant emotion is fear (3). Notice that the variable soothe is now 1, meaning that her fear is now being calmed. She is also prohibited from moving now since the variable ainhib is 1. This variable is turned on when Ms. Pacrat is eating, thinking, or being soothed. When Ms. Pacrat arrives at home, scared, and is soothed, the thalgl becomes active. Notice its value at .1.

4.2.4.2.2. Figure 2

The top display shows an example of two of Ms. Pacrat's three forms of thinking. Notice that the thnkng variable is 1. This snapshot is in Trial 2. Ms. Pacrat's location is sensory neuron 347. Her path of thinking, however, places her at neuron 346. During thinking Ms. Pacrat also associates. In this case she is associating to neuron 367 from 346. Notice that during thinking the thalnt becomes active. Its value is now .1. Sometime in the past neuron 346 was in a path that was rewarded. The values in iakshn 346

show the four weights for this neuron. The first field is north, the second field is east, the third field is south, and the fourth field is west. The weight from 346 east to 367 has been reinforced, while the other 3 have been reduced. Iakshn also shows us that at time 429 Ms. Pacrat moved north to 345 and associated west at time 428. It also shows the association from 346 east, to 367 at time 2407.

The second display shows an angry Ms. Pacrat. Her dominant emotion is anger (1). It is Trial 2 and she knows that in Trail 1 food was found at neuron 368. Notice that her previous location was 368, but food is now located at 360, row 3 column 18. Ms. Pacrat's angry variable is now 1.0. Also notice that the move partion is equally divided so that Ms. Pacrat's moves will now be totally random.

4.2.4.2.3. Figure 3

This table shows the number of times that Ms. Pacrat is rewarded for finding food when hungry, and getting home when scared. It also shows the average number of moves for these rewarded trials. Note that the number of times rewarded decreases as Ms. Pacrat's performance improves. This is because the running average that determines if Ms. Pacrat will be

rewarded is a dynamic threshold that is continually lowered. This means that as Ms. Pacrat learns, more is expected of her and once she learns where food and home are she is no longer in need of rewards.

4.2.4.2.4. Figure 4

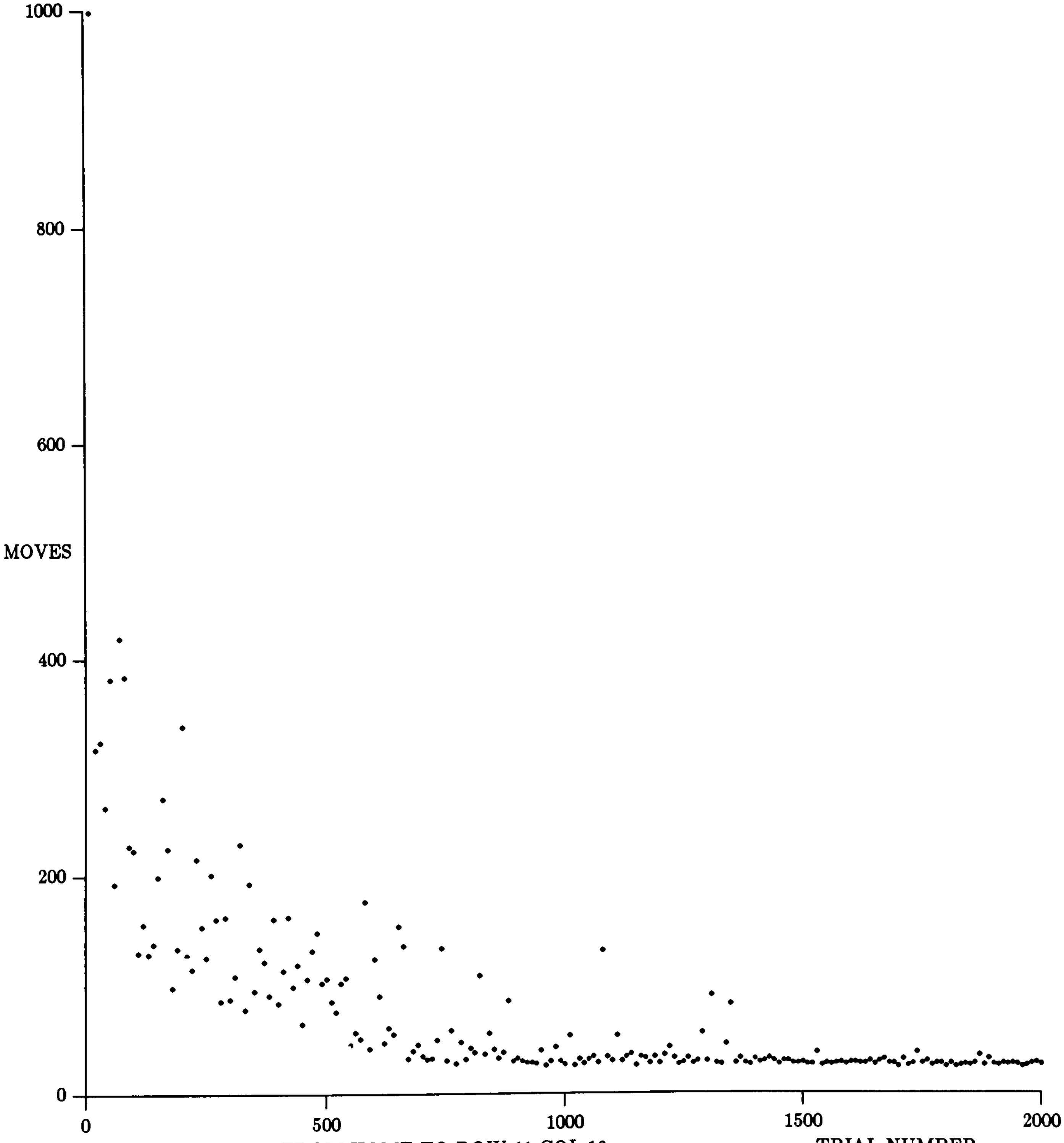
This table shows the reward thresholds below which Ms. Pacrat must be to get rewarded. These thresholds gradually decrease as explained above. Notice that the averages for all four paths are represented.

4.2.4.2.5. Figure 5

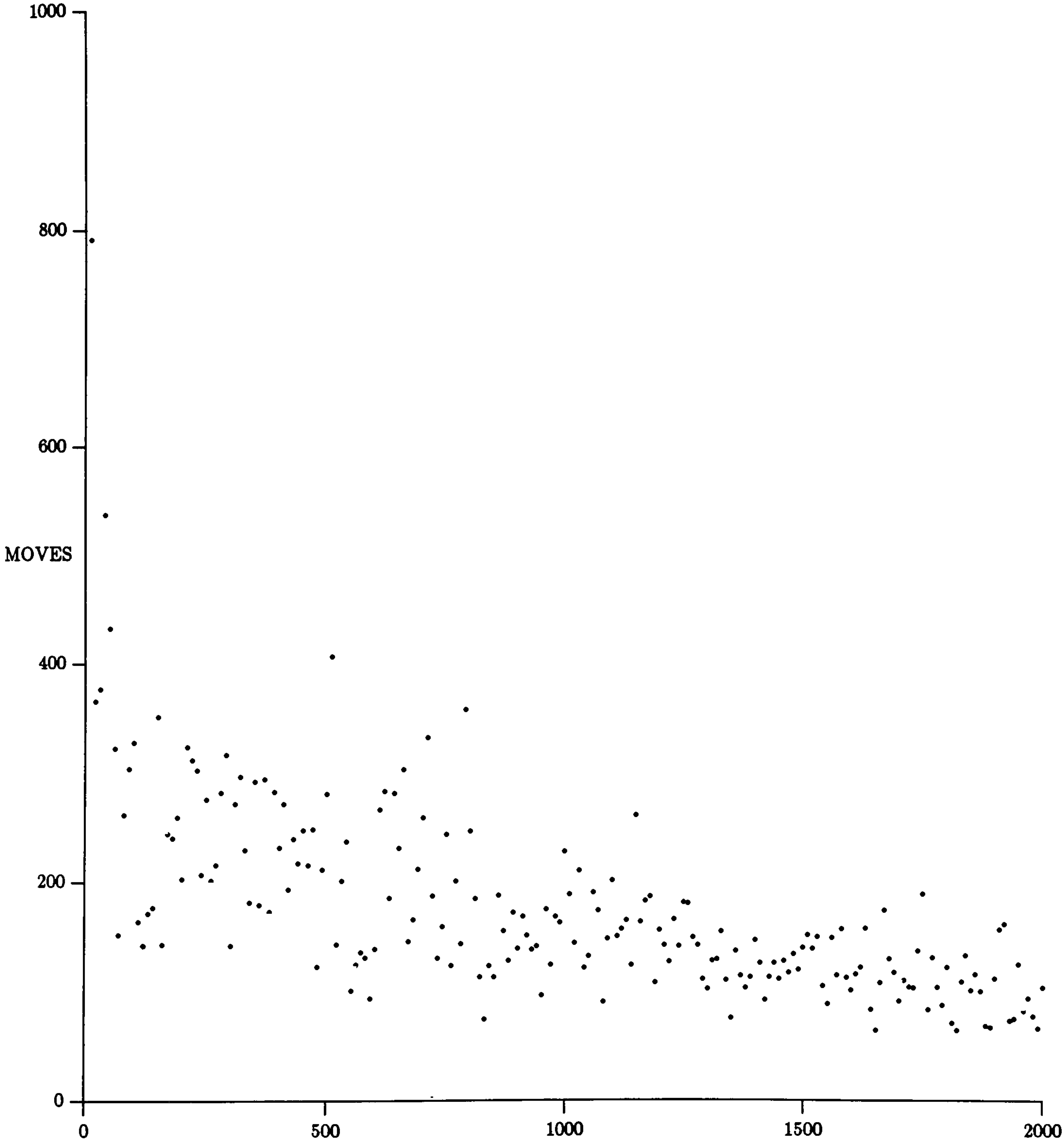
This shows the anatomy of Ms. Pacrat's brain. The specific parts have been described previously in this thesis.

4.2.4.2.6. Figure 6

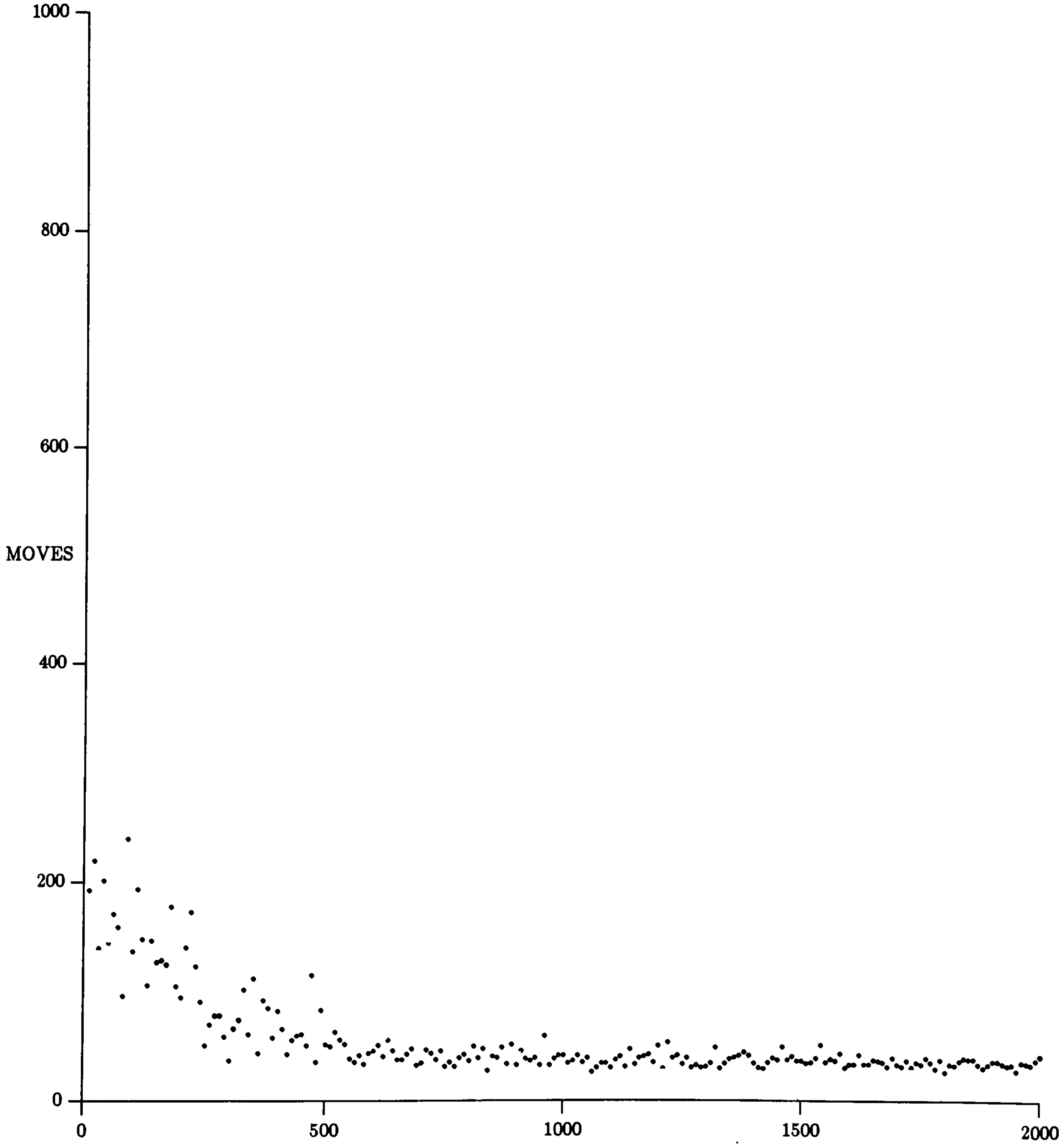
This shows the graphical display of Ms. Pacrat's world. The grid size is 21 X 21. Her home is shown at location 11, 1 with an H. Ms. Pacrat herself is represented by a shaded in rectangle, and her association location is seen as a smaller, unshaded, rectangle. The location of food is noted by an X.



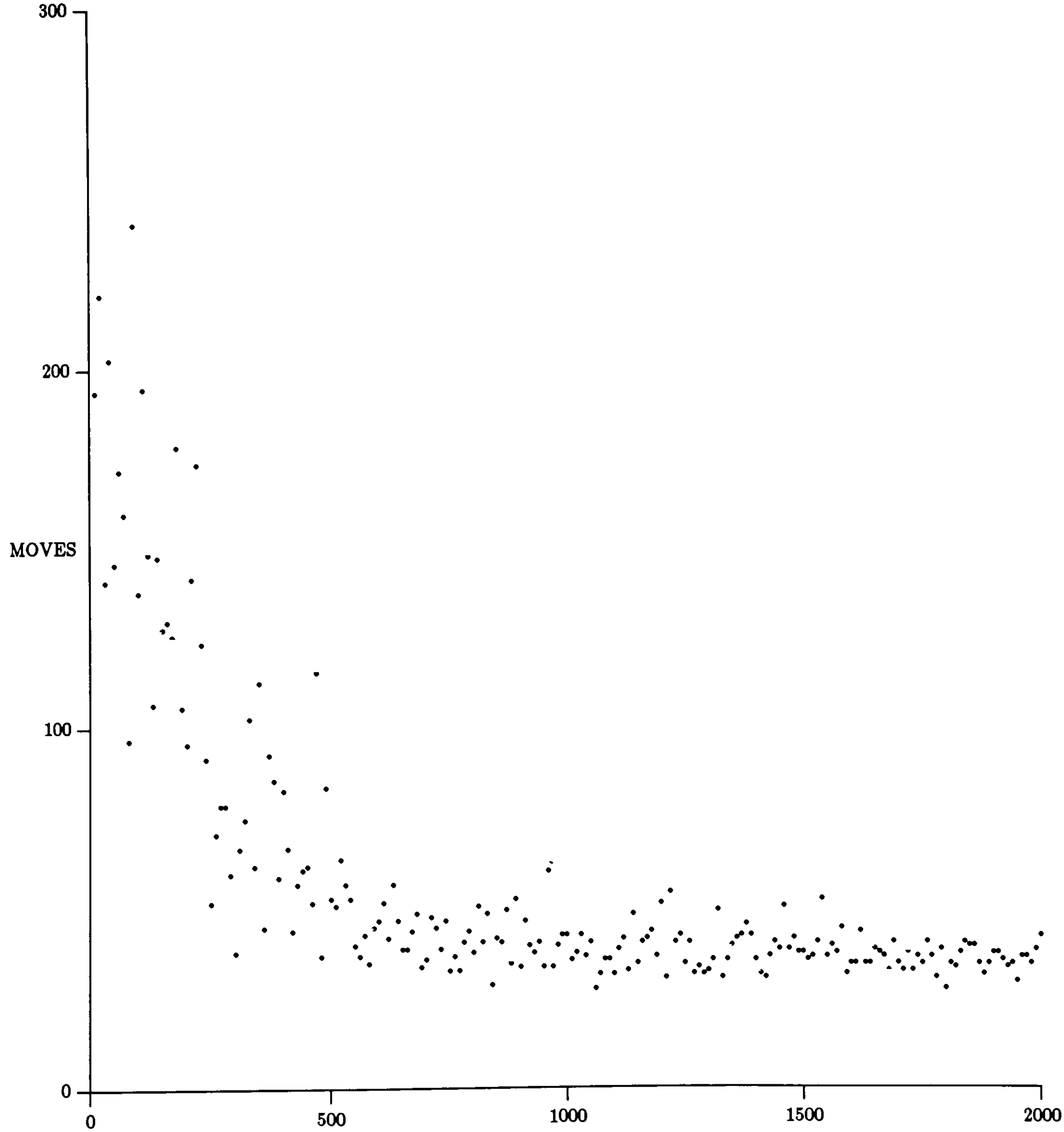
GRAPH 1 FROM HOME TO ROW 11 COL 18



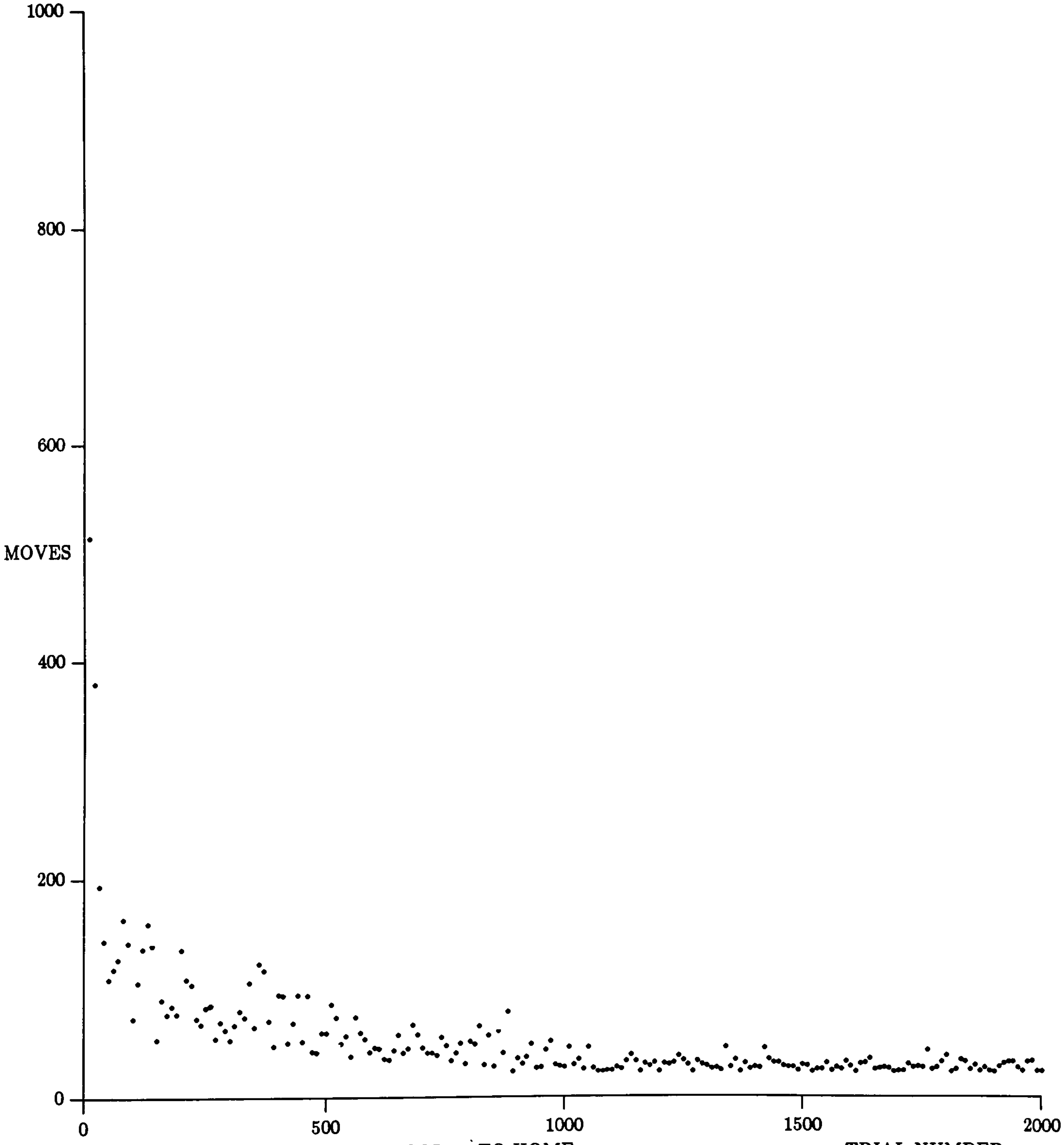
GRAPH 2 FROM HOME TO ROW 3 COL 18



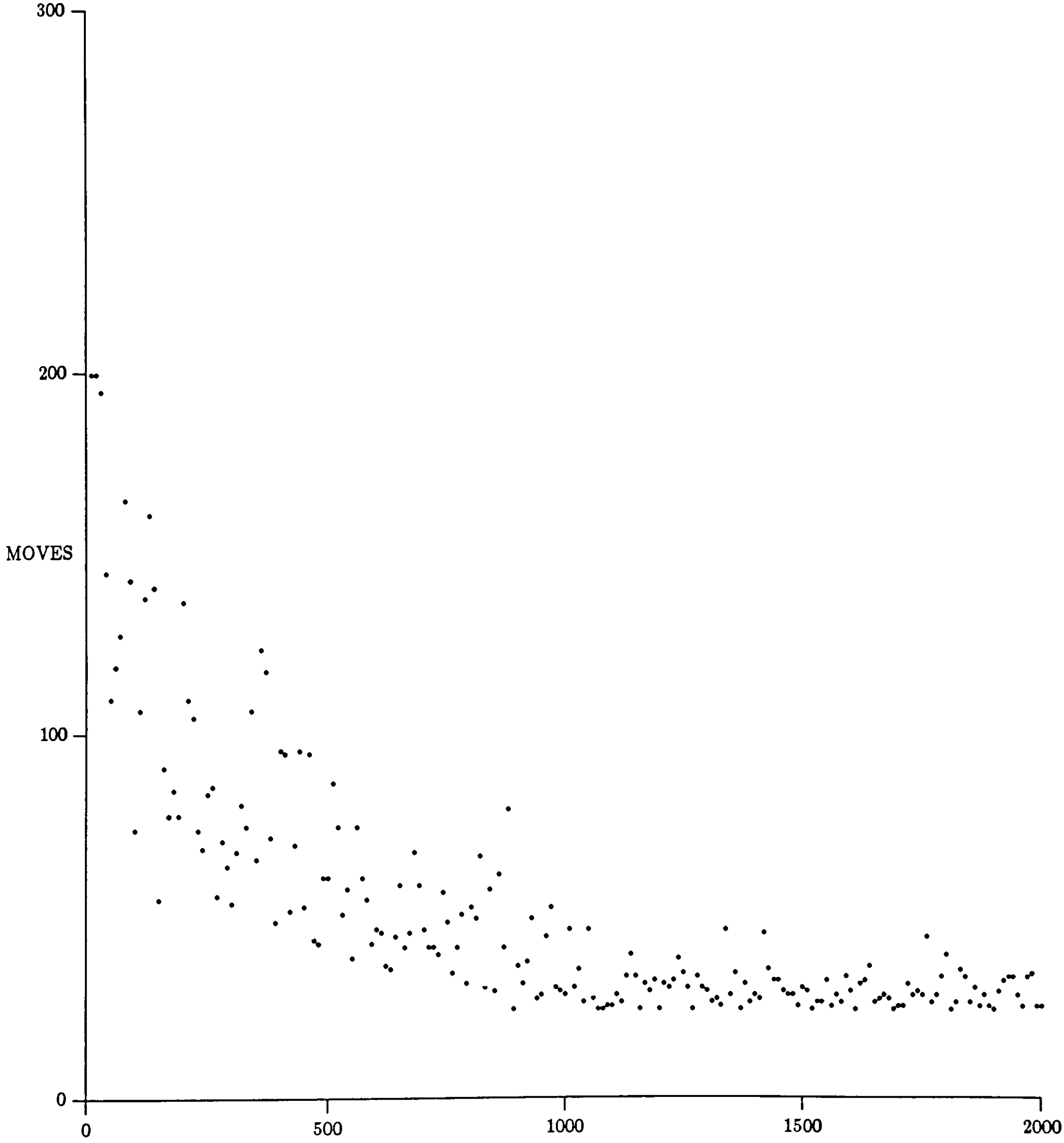
GRAPH 3 FROM ROW 3 COL 18 TO HOME



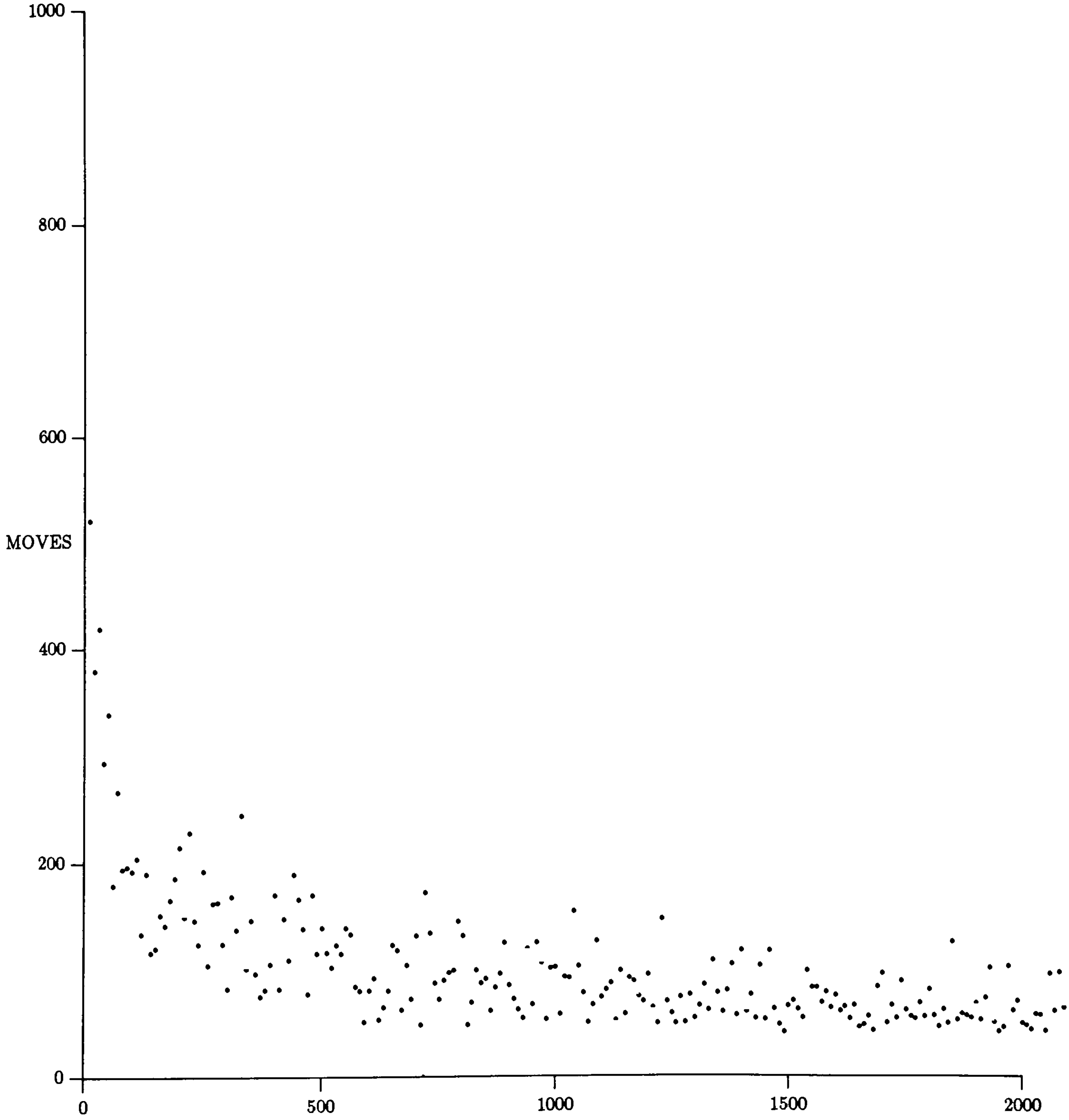
GRAPH 4 FROM ROW 3 COL 18 TO HOME



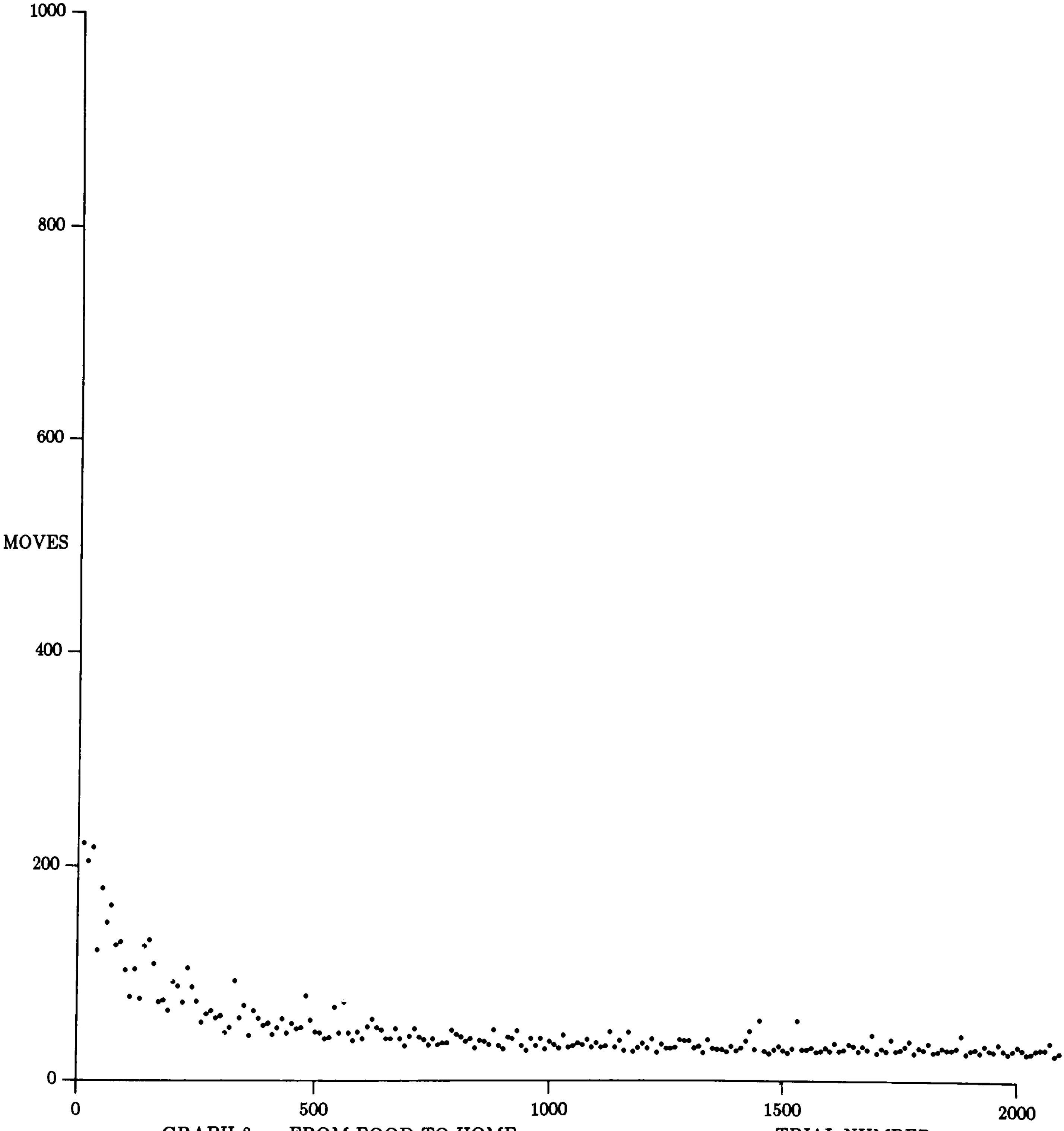
GRAPH 5 FROM ROW 11 COL 18 TO HOME



GRAPH 6 FROM ROW 11 COL 18 TO HOME



GRAPH 7 FROM HOME TO FOOD



GRAPH 8 FROM FOOD TO HOME

CYCLE STATS
****
TRIAL is 1 IBGCNT is 682 NUMOVE is 201
P's LOC is neuron 368: ROW 11 COL 18 P's PREV LOC is neuron 347
DOMINANT EMOTION is 2 LSUGG is 2 FOOD LOCATION is ROW 11 COL 18
ACTIVE CODON is SENSORY NEURON 368 STATE 1
PACRAT is eating
0.900000 is hungry 1.000000 is uneasy 0.000000 is angry 0.000000 is bored 0.000000 is curis
0.000000 is aintnd 1.000000 is ainhib 0 is soothe 0.100000 is stomak 0 is thnkng 1.000000 is alert
0.000000 is thalbd 0.100000 is amygnt 0.100000 is thalgl 0.000000 is thalnt
0.250000 0.250000 0.250000 0.250000 – weights for this codon
0.125000 1.000000 1.000000 0.008333 is repulse
0.010000 0.250000 0.250000 0.172917 is sugg
0.014643 0.380720 0.746797 1.000000 – move partition
Associating is 0 ISENSE is 368 368 2 0 347 0 is codon of assoc
****

icodon 368 1 is
368 2 110 347 682

iakshn 368 1 is
0.250000 0.250000 0.250000 0.250000
0.000000 0.000000 0.000000 0.000000
0.000000 0.000000 0.000000 0.000000

CYCLE STATS
****
TRIAL is 1 IBGCNT is 1034 NUMOVE is 310
P's LOC is neuron 11: ROW 11 COL 1 P's PREV LOC is neuron 10
DOMINANT EMOTION is 3 LSUGG is 3 FOOD LOCATION is ROW 11 COL 18
ACTIVE CODON is SENSORY NEURON 11 STATE 4
0.011019 is hungry 0.900000 is uneasy 0.000000 is angry 0.000000 is bored 0.600000 is curis
0.000000 is aintnd 1.000000 is ainhib 1 is soothe 0.988981 is stomak 0 is thnkng 1.000000 is alert
0.000000 is thalbd 0.000000 is amygnt 0.100000 is thalgl 0.000000 is thalnt
0.250000 0.250000 0.250000 0.250000 – weights for this codon
0.008333 1.908333 1.958333 1.000000 is repulse
0.481250 0.010000 0.485417 0.250000 is sugg
0.392323 0.400476 0.796196 1.000000 – move partition
Associating is 0 ISENSE is 11 11 3 0 10 0 is codon of assoc
****

icodon 11 4 is
11 3 10 10 1034

iakshn 11 4 is
0.250000 0.250000 0.250000 0.250000
0.000000 0.000000 0.000000 0.000000
0.000000 0.000000 0.000000 0.000000

FIGURE 1

CYCLE STATS
****
TRIAL is 2 IBGCNT is 2407 NUMOVE is 362
P's LOC is neuron 347: ROW 11 COL 17 P's PREV LOC is neuron 348
DOMINANT EMOTION is 2 LSUGG is 2 FOOD LOCATION is ROW 3 COL 18
ACTIVE CODON is SENSORY NEURON 346 STATE 1
0.814593 is hungry 1.000000 is uneasy 0.000000 is angry 0.000000 is bored 0.800000 is curis
0.000000 is aintnd 1.000000 is ainhib 0 is soothe 0.185407 is stomak 1 is thnkng 0.814590 is alert
0.000000 is thalbd 0.000000 is amygnt 0.000000 is thalgl 0.100000 is thalnt
0.233645 0.299065 0.233645 0.233645 -- weights for this codon
1.000000 0.025000 0.008333 1.000000 is repulse
0.233645 0.299065 0.233645 0.233645 is sugg
0.233645 0.532710 0.766355 1.000000 -- move partition
Associating is 1 ISENSE is 0 367 2 0 346 0 is codon of assoc
****

icodon 346 1 is
346 2 1 347 2407

iakshn 346 1 is
0.233645 0.299065 0.233645 0.233645
345.000000 0.000000 0.000000 0.000000
429.000000 2407.000000 0.000000 428.000000

CYCLE STATS
****
TRIAL is 2 IBGCNT is 3905 NUMOVE is 730
P's LOC is neuron 389: ROW 11 COL 19 P's PREV LOC is neuron 368
DOMINANT EMOTION is 1 LSUGG is 2 FOOD LOCATION is ROW 3 COL 18
ACTIVE CODON is SENSORY NEURON 389 STATE 6
0.838475 is hungry 1.000000 is uneasy 1.000000 is angry 0.000000 is bored 1.000000 is curis
0.000000 is aintnd 0.000000 is ainhib 0 is soothe 0.161525 is stomak 0 is thnkng 0.838472 is alert
0.000000 is thalbd 0.000000 is amygnt 0.000000 is thalgl 0.000000 is thalnt
0.250000 0.250000 0.250000 0.250000 -- weights for this codon
1.608333 1.691667 1.591667 0.008333 is repulse
0.250000 0.250000 0.250000 0.250000 is sugg
0.250000 0.500000 0.750000 1.000000 -- move partition
Associating is 0 ISENSE is 389 389 1 0 368 0 is codon of assoc
****

icodon 389 6 is
389 1 0 368 3905

iakshn 389 6 is
0.250000 0.250000 0.250000 0.250000
0.000000 0.000000 0.000000 0.000000
0.000000 0.000000 0.000000 0.000000

FIGURE 2

REWARD DATA

REWARD AVERAGES				
TRIAL #	# OF TIMES REWARDED FOR FOOD	AVERAGE # MOVES	# OF TIMES REWARDED FOR HOME	AVERAGE # MOVES
100	54	180	58	105
200	45	96	47	67
300	41	99	52	45
400	49	86	39	38
500	39	71	46	33
600	50	67	46	33
700	51	64	31	28
800	54	58	29	27
900	53	53	35	28
1000	17	68	30	26
1100	37	65	10	25
1200	30	66	5	25
1300	26	68	10	26
1400	41	43	7	25
1500	16	67	8	26
1600	22	70	9	28
1700	20	64	3	25
1800	22	66	7	25
1900	18	62	3	25
2000	20	63	7	24

This table shows the number of times that Ms. Pacrat has been rewarded for finding food and getting home. It also shows the average number of moves for these rewarded trials.

FIGURE 3

REWARD DATA

	REWARD THRESHOLD			
TRIAL #	TO HOME FROM FOOD 1	TO HOME FROM FOOD 2	TO FOOD 1 FROM HOME	TO FOOD 2 FROM HOME
100	119	86	206	216
200	84	73	120	159
300	40	43	90	152
400	54	59	81	171
500	46	43	87	168
600	31	35	70	89
700	29	30	29	102
800	28	31	30	176
900	29	30	29	102
1000	30	23	22	134
1100	23	19	31	118
1200	31	20	22	110
1300	24	21	27	86
1400	29	20	23	91
1500	28	21	21	96
1600	26	21	20	81
1700	26	19	20	81
1800	23	25	19	78
1900	25	18	21	66
2000	26	20	19	62

This table shows the averages that determine if Ms. Pacrat will be rewarded. There are four distinct averages representing the four paths from home to both foodspots and back again.

FIGURE 4

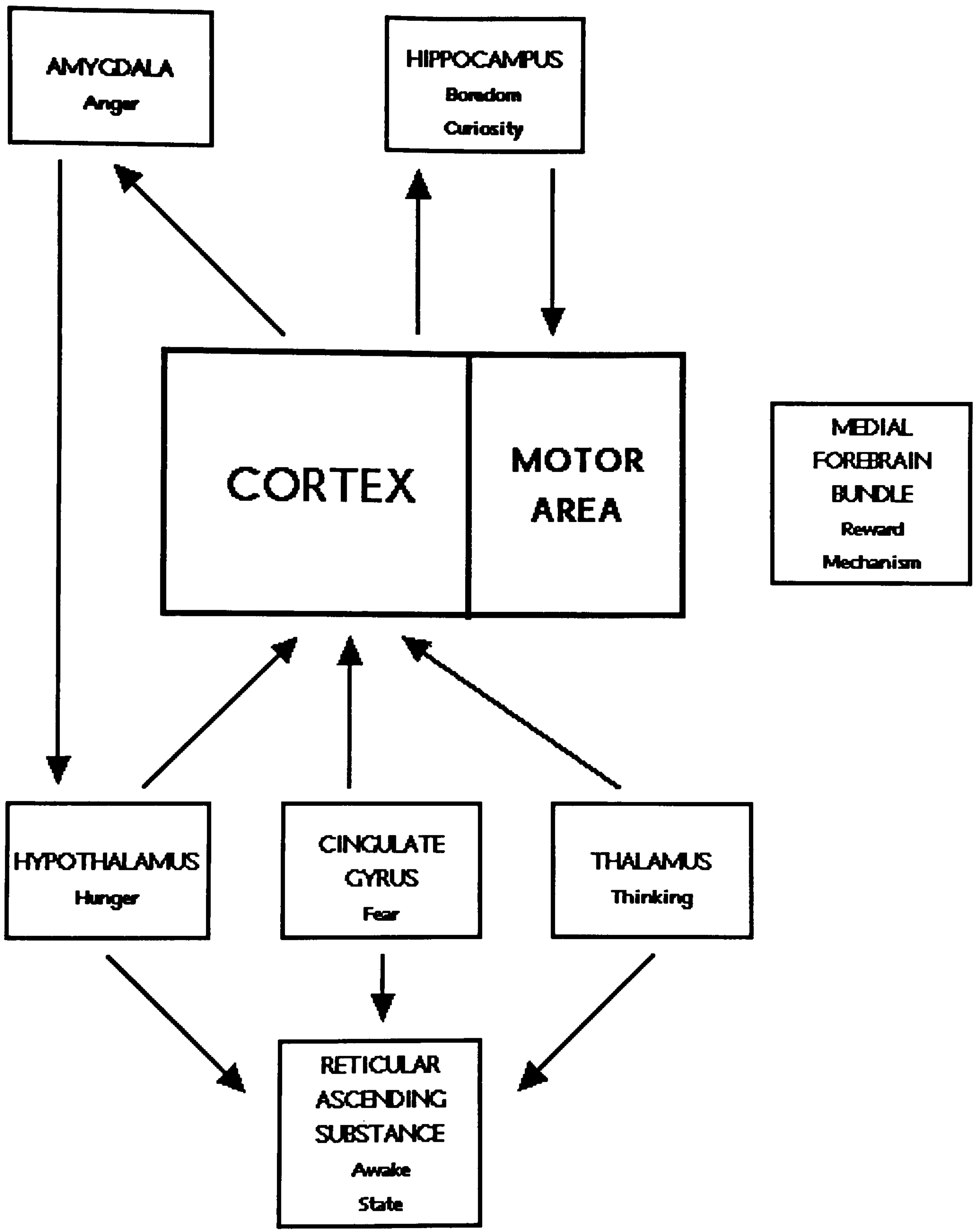
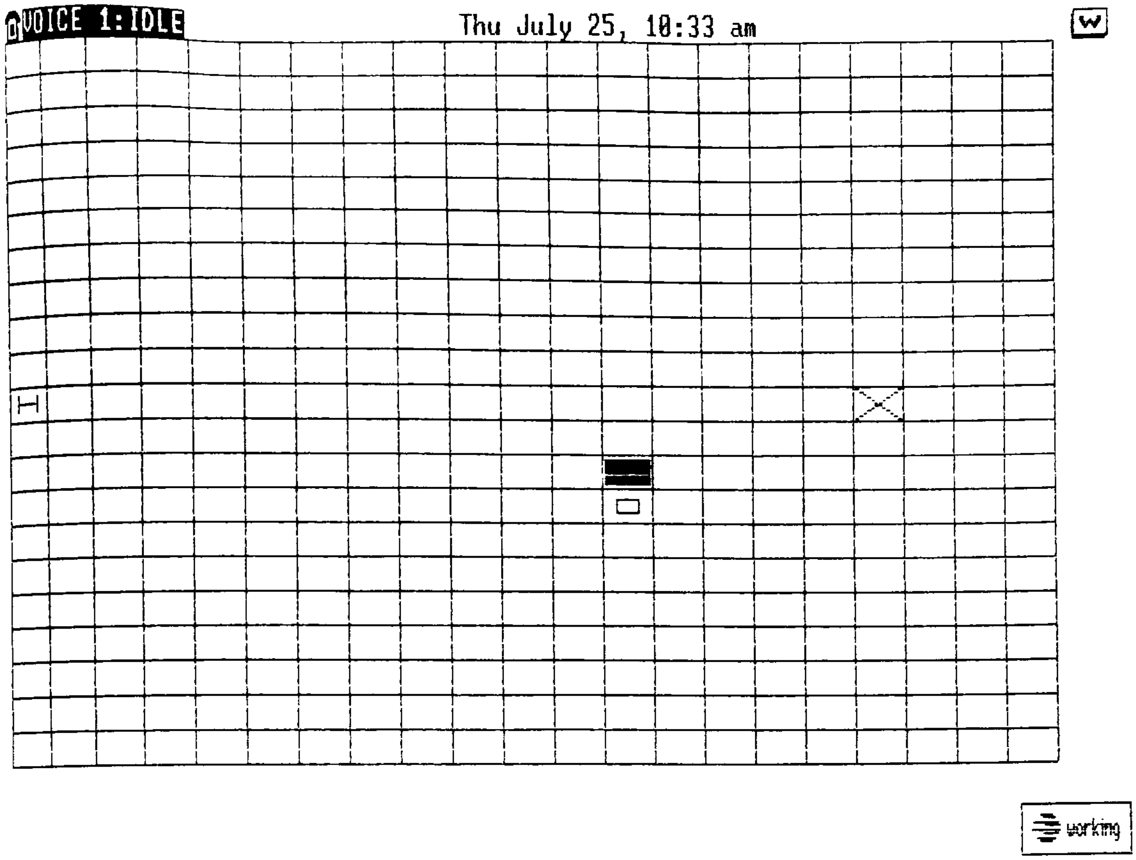

AMYGDALA
Anger
HIPPOCAMPUS
Boredom
Curiosity
CORTEX
MOTOR
AREA
MEDIAL
FOREBRAIN
BUNDLE
Reward
Mechanism
HYPOTHALAMUS
Hunger
CINGULATE
GYRUS
Fear
THALAMUS
Thinking
RETICULAR
ASCENDING
SUBSTANCE
Awake
State

FIGURE 5


VOICE 1:IDLE
Thu July 25, 10:33 am
working

5. CONCLUDING REMARKS

In this thesis I have examined the attempt of the field of Artificial Intelligence to build intelligent machines. As we have seen there have been many successful programs created, but there is still much work to be done. Human thought processes can be used as a model for intelligent machines. We can reason, make decisions, and do other things that are considered to be intelligent. An important component to this is that we do intelligent things under the influence of our emotions. In this thesis I have suggested that thinking with emotion may be the key to producing intelligent behavior in machines. Processes such as, creativity, insight, and intuition are also essential to our thinking. Because of the role that these processes play in humans, I feel that they deserve study and consideration in AI. The question here is, Will machines ever be able to think as well as we do? The answer, I feel, is still unknown, but the attempt itself is certainly interesting, worthwhile, and a source of new and exciting insights.

6. SOME INTERESTING QUOTES

Fancy what a game of chess would be if all the chessmen had passions and intellects, more or less small and cunning; if you were not only uncertain about your adversary's men, but a little uncertain about your own; if your Knight could shuffle himself on to a new square on the sly; if your Bishop, in disgust at your castling, could wheedle your Pawns out of their places, and if your Pawns, hating you because they are Pawns, could make away from their appointed posts that you might get checkmate on a sudden. You might be the longest-headed of deductive reasoners and yet you might be beaten by your own Pawns. You would be especially likely to be beaten, if you depended arrogantly on your mathematical imagination, and regarded your passionate pieces with contempt. Yet this imaginary chess is easy compared with a game a man has to play against his fellow-men with other fellow-men for his instruments....

George Eliot, 1866[82]

Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain – that is, not only write it but know that it had written it. No mechanism could feel (and not merely artificially signal, an easy contrivance) pleasure at its successes, grief when its valves fuse, be warmed by flattery, be made miserable by its mistakes, be charmed by sex, be angry or depressed when it cannot get what it wants.

Professor Jefferson Lister, 1949[83]

Since in actual human behavior motive and emotion are major influences on the course of cognitive behavior, a general theory of thinking and problem solving must incorporate such influences.

Herbert Simon, 1967[84]

---

[82]Alick Elithorn and David Jones. *Artificial And Human Thinking*. San Franscisco, Calif: Jossey Bass Inc., 1973. p. 1.

[83]Douglas Hofstadter and Daniel C. Dennett. *The Mind's I*. NY, NY: Basic Books Inc., 1981. p. 60.

[84]Keith Gunderson. *Mentality And Machines*. Minneapolis, Minn: Univ. of Minnesota Press, 1971. p. 157.

It is possible, in human beings, for the anger to produce physical disturbances. However, if X satisfied enough cognitive conditions he could rightly describe himself as being very angry, despite not having the physical symptoms. The anger could be strong, insofar as it constantly intruded into his thoughts and decisions, and insofar as he strongly desired to make Y suffer, and suffer a great deal.

Aaron Sloman and Monica Croucher,1981[85]

The Emotions are all those feelings that so change men as to affect their judgements.

Aristotle[86]

Mathematical formalizers wish to treat matters of intuition mathematically, and make themselves ridiculous.... The mind....does it tacitly, naturally, and without technical rules.

Pascal,1670.[87]

Certainly Dreyfus is right in maintaining that a machine which can handle only a certain select set of sentences or objects in an ad hoc manner should not be regarded as capable of "understanding" anything (at least not in the same sense that humans do). But on the other hand, it is excessive to require that one single machine be capable of duplicating all the best performances which any human could possibly produce. If this were part of the requirement for being an intelligent being, then no human could qualify either.

Stuart Goldkind,1987[88]

Let S be a "complete model" of a human being and one that could satisfy inter alia the Turing Interrogation Test. It is, of course, not being asserted that S, the model of a human being, is a human being,

[85] Aaron Sloman and Monica Croucher. "Why Robots Will Have Emotions". *Ninth International Joint Conference On AI, 1981.* p. 197-202.

[86] William Lyons. Emotion. Cambridge, Mass: Cambridge Univ. Press, 1980. p. 92.

[87] Hubert and Stuart Dreyfus. *Mind Over Machine.* NY, NY: The Free Press, 1986. p. 16.

[88] Stuart Goldkind. *Machines and Intelligence.* NY, NY: Greenwood Press, 1987. p. 49.

although we must ask ourselves whether there could arise a situation where S and the human would be indistinguishable on one hand, or "identical" on the other.... It is not different to imagine the inclusion of emotions in S. This would be necessary for survival for any S which was to be involved in an "open society" – threatened by other species (especially, say, homo sapiens) and competing for goal-satisfaction and goal-needs. It would be an alarm system which would facilitate physical activity e.g. fight and flight and would also by virtue of its fabric have all sorts of side-effects. S's for most purposes, do not need emotions, but to simulate human beings they do.

Frank George, 1986[89]

It can also be argued that there is no way to build a super-intelligent robot which also copes with a complex set of different sorts of motives, in a partly unpredictable world, without giving that robot mechanisms which are capable of producing emotional states, as a result of performing the cognitive tasks for which they are required. That is, the possibility of having emotions may be a by-product of being able to cope with a complex and unpredictable world in an intelligent way. (This does not mean that every intelligent robot will necessarily be emotional, only that it will have the ability – and abilities are not always exercised.)

Aaron Sloman, 1984[90]

SANDY: You've hit upon a sensitive point. I've thought about this for a long time, I've concluded that I simply can't believe emotions and thought can be divorced. To put it another way, I think emotions are an automatic by-product of the ability to think.

CHRIS: That's an intersting conclusion, but what if you're wrong? What if I produce a machine that can think but not emote? Then its intelligence might go unrecognized because it failed to pass your kind of test.

SANDY: I'd like you to point out to me where the boundary line

[89]Frank George. *AI Its Philosophy and Neural Content.* NY, NY: Gordon and Breach Science Publishers, 1986. p. 14,17.

[90]Mike Sharples. *Computers and Thought.* Cambridge, Mass: The MIT Press, 1989. p. 314.

between emotional questions and non-emotional ones lies. You might want to ask about the meaning of a great novel. This certainly requires the understanding of human emotions! Now is that thinking, or mere cool calculation? You might want to ask about a subtle choice of words. For that you need an understanding of their connotations. Turing uses examples like this in his article. You might want to ask for advice about a complex romantic situation. The machine would need to know a lot about human motivations and their roots. If it failed at this kind of task, I would not be much inclined to say that it could think. As far as I'm concerned, thinking, feeling, and consciousness are just different facets of one phenomenon, and no one of them can be present without the others.

CHRIS: I still can't see that intelligence has to involve emotions. Why couldn't you imagine an intelligence that simply calculates and has no feeling.?

SANDY: A couple of answers here. Number one, any intelligence has to have motivations. It's simply not the case, whatever many people may think, that machines could think any more "objectively" than people do. Machines, when they look at a scene, will have to focus and filter that scene down into some preconceived categories, just as a person does. And that means seeing some things and missing others. It means giving more weight to some things than others. This happens on every level of processing.

PAT: I'm not sure I'm following you.

SANDY: Take me right now, for instance. You might think I'm just making some intellectual points, and I wouldn't need emotions to do that. But what makes me CARE about these points? Just now – why did I stress the word "care" so heavily? Because I'm emotionally involved in this conversation! People talk to each other out of conviction – not out of hollow, mechanical reflexes. Even the most intellectual conversation is driven by underlying passions. There's an emotional undercurrent to every conversation – it's the fact that the speakers want to be listened to, understood, and respected for what they are saying.

PAT: It sounds to me as if all you're saying is that people need to be interested in what they're saying, otherwise a conversation dies.

SANDY: Right! I wouldn't bother to talk to anyone if I weren't motivated by interest. And "interest" is just another name for a

whole constellation of subconscious biases.

Douglas R. Hofstadter,1985[91]

I do believe the following: 1)affective phenomena are rich and varied; 2)their relation to intellect and cognition is far from clear; and therefore 3) they deserve a lot more study, even in AI.

John Haugeland, 1985[92]

It must be said that AI researchers and cognitive scientists have generally tended to avoid the subject of emotion, preferring to concentrate instead on the purely cognitive aspects of intelligence. But some things are clear enough. For example, psychologists generally agree that one major function of emotions is to focus our attention, to help us decide what is important in the world. When someone is frightened by a spider, that spider is very important; for the moment, at least, nothing else seems to exist. When we are moved by a play or a piece of music, we tend to forget ourselves and our surroundings as we watch. Furthermore, if we define emotion in a very broad sense, it's also clear that it serves to define our long-term goals and motives. Thus, an ambitious person may shape his or her entire career in ways that satisfy a need for the approval of others, or for the security of money. An extremely shy person may do just the opposite, choosing a life-style that avoids confrontations with others.

M.Mitchell Waldrop, 1987[93]

[91]Douglas R. Hofstadter. *Metamagical Themas.* NY, NY: Basic Books, 1985. p. 502.

[92]John Haugeland. *Artificial Intelligence: The Very Idea.* Cambridge, Mass.: The MIT Press, 1985. p. 237.

[93]M. Mitchell Waldrop. *Man-Made Minds.* USA: Walker and Company, 1987. p. 131.

BIBLIOGRAPHY

Alston, William. "Emotion And Feeling". *Encyclopedia Of Philosophy. Vol. 2.* NY, NY: Macmillan Pub, 1967. p. 479 - 493.

Arnold, Magda. *Emotion and Personality.* NY, NY: Columbia University Press, 1960.

Barr, Avron and Edward Feigenbaum. *The Handbook Of Artificial Intelligence.* Reading, Mass: Addison-Wesley Pub. Co., 1981.

Bower, G.H. "Mood and Memory". *American Psychologist.* Vol. 6, No. 2, Feb.1981, p.129-148.

Churchland, Paul M. *Matter and Consciousness.* Cambridge, Mass: The MIT Press, 1984.

Freedman, David. "Common Sense and the Computer". *Discover Magazine.* Volume 11 Number 8. March 1991. p. 67 - 71.

Freedman, David. "Invasion of the Insect Robots". *Discover Magazine.* Volume 12 Number 3. August 1990. p. 42 - 50.

Dreyfus, Hubert. *What Computers Can't Do.* NY, NY: Harper and Row Pub., 1972.

Dreyfus, Hubert and Stuart. *Mind Over Machine.* NY, NY: The Free Press, 1986.

Elithorn, Alick and David Jones. *Artificial And Human Thinking.* San Franscisco, Calif: Jossey Bass Inc., 1973.

Feigenbaum, Edward and Julian Feldman. *Computers and Thought.* NY, NY: McGraw-Hill, 1963.

George, Frank. *AI Its Philosophy and Neural Content.* NY, NY: Gordon and Breach Science Publishers, 1986.

Goldkind, Stuart. *Machines and Intelligence.* NY, NY: Greenwood Press, 1987.

Gunderson, Keith. *Mentality and Machines.* Minneapolis, Minn: Univ. Of Minnesota Press, 1971.

Pagels, Heinz R. *The Dreams Of Reason.* NY, NY: Bantam Books, 1989.

Hofstadter, Douglas and Daniel C. Dennett. *The Mind's I.* NY, NY: Basic Books Inc., 1981.

Izard, Carroll. *Human Emotions.* NY, NY: Plenum Press, 1977.

Lyons, William. *Emotion.* Cambridge, Mass: Cambridge Univ. Press, 1980.

Mandler, George. *Mind And Emotion.* NY, NY: John Wiley and Sons, 1975.

Minsky, Marvin. *The Society Of Mind.* NY, NY: Wiley and Sons, 1985.

Rumelhart, D. E. and J.L. McClelland (Eds.). *Parallel Distributed Processing I & II.* Cambridge: MIT Press, 1986.

Scanlon, Raymond and Mark Johnson. "Machine Emotion". In *IEEE Spring Compcon.* Washington, DC: Computer Society Press Of The IEEE. Feb. 29 - March 4, 1988.

Scanlon, Raymond and Mark Johnson. "Experiences With a Feeling-Thinking Machine". 1987 IEEE International Conference On Neural Networks. Vol. 2 June 21-24. Washington, DC: Computer Society Press Of the IEEE,1987.

Scanlon, Raymond and Mark Johnson. "Synaptogenesis". Watervliet, NY: US Army Armament Research, Development and Engineering Center, 1990.

Searle, John. *Minds, Brains and Science.* Cambridge, Mass.: Harvard University Press, 1984.

Sharples, Mike. *Computers and Thought.* Cambridge, Mass: The MIT Press, 1989.

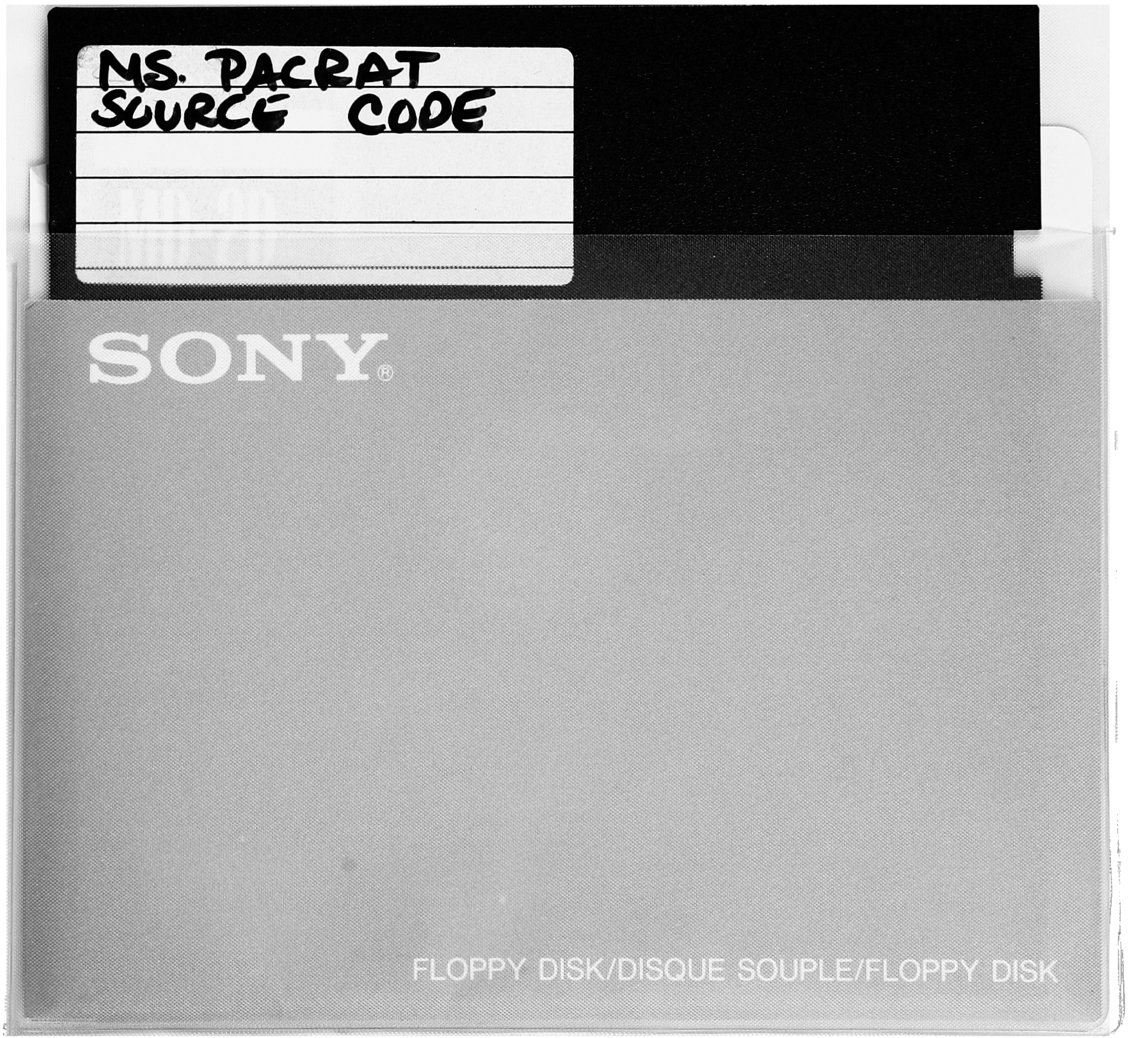
Sloman, Aaron and Monica Croucher. "Why Robots Will Have Emotions". Ninth International Joint Conference On AI, 1981.

Steel, L and J. A. Campbell. *Progress In AI.* West Sussex, England: Ellis Norwood Pub, 1985.

Turing, Alan. "Computing Machinery and Intelligence", 1950, reprinted in *Computers and Thought.* E. Feigenbaum and J. Feldman, eds. NY, NY:

McGraw-Hill, 1963.

Waldrop, M. Mitchell. *Man-Made Minds.* USA: Walker and Company, 1987.


MS. PACRAT
SOURCE CODE
SONY
FLOPPY DISK/DISQUE SOUPLE/FLOPPY DISK