

CARROLL COLLEGE

The Thinking Machine:
A Philosophical Study of
Thought, Sentience and Intelligence
in Computers

A Dissertation Submitted to
The Faculty of the Department of Philosophy
in Candidacy for the Degree of
Bachelor of Arts: Philosophy
Department of Philosophy

By
Christopher M. Perez

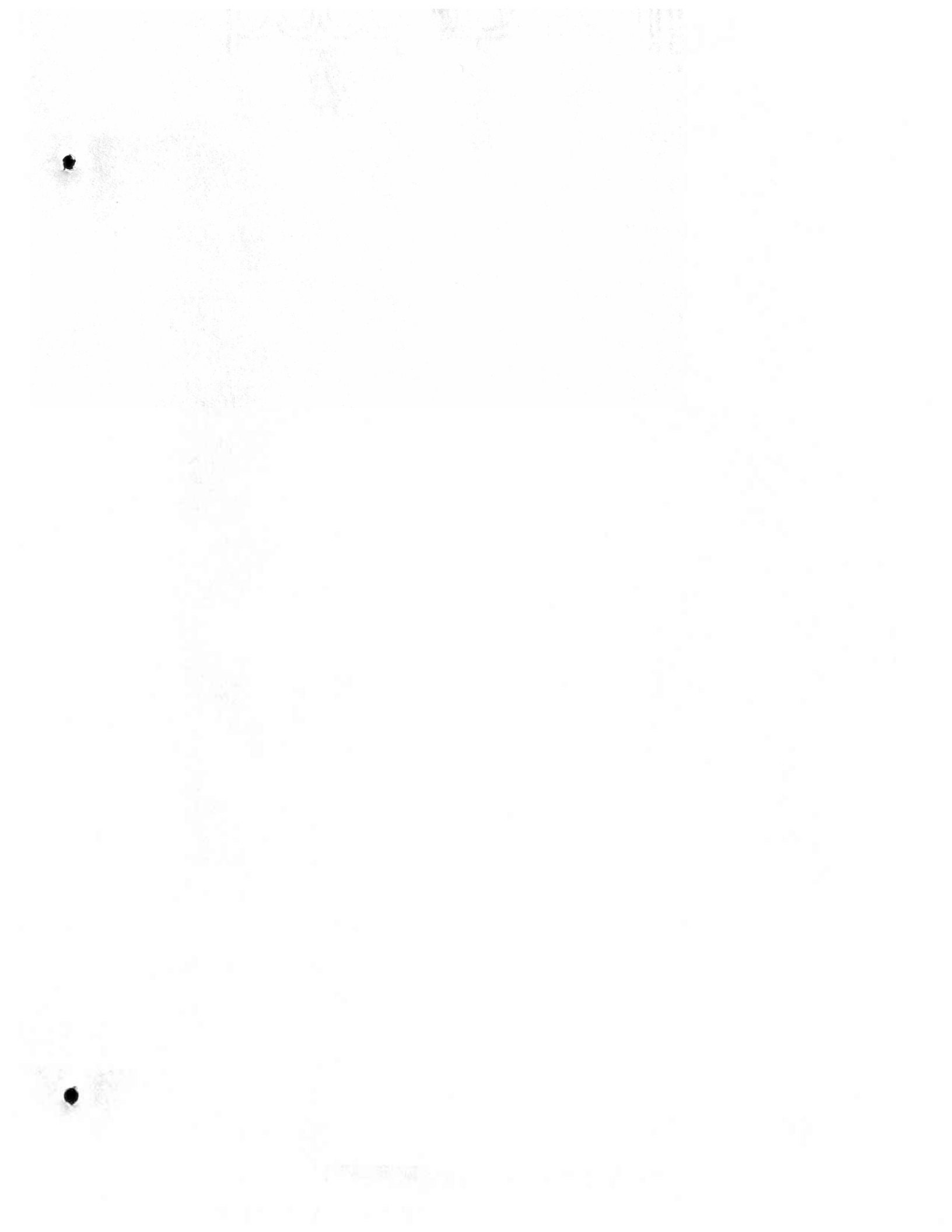
Helena, Montana

April 1989



3 5962 00083 839

CORETTE LIBRARY
CARROLL COLLEGE



SIGNATURE PAGE

This thesis for honors recognition has been approved for
the Department of Philosophy.

<i>Barry Feist, Ph.D.</i> Director	<i>April 10, 1989</i> date
<i>Philip B. Kne</i> Reader	<i>10 April, 1989</i> date
<i>Richard T. Lambert</i> Reader	<i>April 10, 1989</i> date

ACKNOWLEDGMENTS

My special thanks, and all my love, to Anke, my wife, for the understanding and patience during this effort. I also wish to express my gratitude to the Montana Army National Guard for its logistical support.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

Chapter Page

1. INTRODUCTION 1

 Opening Remarks

2. DEFINITIONS FOR "INTELLIGENCE" AND OTHER TERMS 8

 Intelligence

 Sentience

 Thought and Thinking

 Additional Terms for Consideration

 Memory

 Imagination

 Appetitive Abilities

 Emotion

 Artificial

 Mimicry

 Creation

 Artificial Intelligence

3. THE TURING TEST AND ELIZA 21

 A description and explanation of the Turing Test

 ELIZA: a history and explanation

4. THE OBJECTIONS AND THEIR REBUTTALS 28

 The Theological Objection

 The "Missing Ingredient" Objection

 The Consciousness Objection

 The Moral Objection

5. CAN A COMPUTER THINK? 42

 The broad meaning of intelligence

 The limited meaning of intelligence

 "Machine Intelligence" now and in the future

 The meaning of "artificial"

 Mimicry and duplication

 Possibilities and consequences

6. THE HUMAN QUESTION 54

 Synergistic Humanity

 Mechanistic Humanity

Chapter	Page
7. CONCLUSION	64
Summary of main points	
Appendix	
A. ELIZA PROGRAM LISTING	69
B. ELIZA RUN-TIME PRINTOUT	72
SELECT BIBLIOGRAPHY	74

CHAPTER 1
INTRODUCTION

Artificial intelligence, or AI as it is known in the trade journals, is as much a buzzword and a marketing term as it is the title and description of an area of research in computer science or epistemology. To sell a software product, or to sell a programming journal, put the phrase "Artificial Intelligence" on the cover or the wrapper. Editorials abound promising the "common man" the cornucopia of benefits or the Pandora's Box of troubles that AI will bring to their lives. Science fiction, on paper and on film, has depicted intelligent computers as both benevolent and malevolent.

Looking beyond the marketing circus, and disregarding the emotional appeals of the editorials, the question comes to mind, "Does anyone really know if a computer can think?" Researchers in computer science, in the attempt to answer this question, have discovered that they must first answer the more basic question, "What do we mean when we use the word 'think'?" In asking this question, the computer scientist has entered a new realm, a place where the work of the computer scientist intersects the biologist's dissection table, the psychologist's laboratory, and the musings of the

philosopher. In addition, any person who speaks English feels that they know what they mean when they say that they "think." However, if you then ask someone to explain what he means when he says "thinking," problems start to occur.

The concern with artificial intelligence is not a new one. Most people living today would assume that the idea of a thinking machine is necessarily tied to the modern digital computer, and that questions of artificial intelligence therefore arose because of its development. Some would remember the fear of machines replacing men due to factory automation, and would place the concept of AI in the time of Henry Ford and the birth of the large assembly lines in the factories of the world. Fewer still would remember the chess automata of the nineteenth century, and the awe-filled reactions and religious and philosophical arguments that their existence raised on the topic of thinking machines. Even earlier than that, Rene Descartes, the French rationalist philosopher, was asking and answering, in his own mind, the question "Can a machine imitate the thinking of a man?"

In his introduction to his scientific and philosophical treatises, Discourse on the Method of Properly Conducting One's Reason and of Seeking the Truth in the Sciences (1637), Descartes proposes "...if there were machines which had a likeness to our bodies and imitated our

actions....,"¹ while writing in the seventeenth century. The questions of AI could be traced back further yet, with the legends of Roger Bacon and the brass head, the story of the Golem, even back to the Greeks in the legend of Pygmalion.² The idea of the inanimate shaped into the likeness of men and imitating the functions of men, including thinking and all its associated operations, has always held for men a fascination that our modern day endeavors into AI are attempting to satisfy.

Yet, in order for AI researchers to attempt to imitate thinking, they must first understand what thinking is.

When someone attempts to explain "thinking," the terms often mentioned first will be "intelligence" or "thought." "Consciousness" or "self-awareness" will also be associated with "thinking," and should be explained if an understanding of the term is attempted.

Some of the other words that are used in reference to intelligence, and to AI, are "memory," "imagination," "emotion" and "appetitive abilities". These words, and the terms already listed, are also words used in everyday language with everyday meanings for people as they talk about "thinking." The philosopher, studying epistemology,

¹Rene Descartes, Discourse on the Method and the Meditations, trans. F.E. Sutcliffe, (New York: Viking Penguin Inc., 1968), 73.

²Thomas Bulfinch, Bulfinch's Mythology: The Age of Fable (Garden City, NJ: Doubleday & Co., 1968), 65-67.

in trying to answer the question, "How do I know what I know?" uses these words. The computer scientist, studying and researching in AI, also uses and has meanings for these words, especially when speaking of "thinking machines." These meanings are different from the philosophical and everyday meanings. This diversity of meaning prevents any sort of common understanding among the various groups who use these words. It is one of the great problems that anyone who approaches AI in study must face.

A danger in any sort of writing or discussion is that words may take on several meanings or usages, depending on the viewpoint of the individuals communicating. There always exists the common, dictionary usage for a word in a language. However, a person can make changes to the meaning of a word, specifying a broader or more limited meaning, to show how that word will be used by that individual. As long as this shift in meaning is made clear and kept in the open, confusion can be avoided. "Glass" has a common meaning of a hard, clear substance used to make windows and drinking containers. "Glass" has the meaning for an inorganic chemist of a semifluid semiconductor. As long as that chemist tells us that a semifluid substance is meant, and not the hard solid implied in the common meaning, we are not confused as to how the word "glass" is being used.

When considering the words mentioned earlier that are important to understanding the issues and problems of AI,

any definition for these terms must be open, and avoid some of the pitfalls that can occur in writing or discussion. The most common error is a sliding meaning. An individual will openly claim that a certain word will be used in a certain way. When he uses that word, what happens is that he will actually, usually unconsciously, expand his meaning for the word, or slide in the way he uses the word so that it includes some meaning that he has not openly declared. This leads to confusion and possibly invalid conclusions about intelligence, based on usage of a sliding meaning for the word "intelligence."

With this concept of openly declaring definitions to avoid sliding usage, we can attempt now to examine definitions of terms and answer the original question. Can a computer think? The thesis will approach an answer in the following manner.

Professor Alan M. Turing, a pioneer in the development of the modern digital computer, considered this question in a paper entitled "Can A Machine Think?" In this article Professor Turing proposes a test that would determine if a machine could imitate a man's thinking, in such a way the machine would meet all possible objections to calling it intelligent. Professor Turing felt that if a machine passed his test, it could be properly called a thinking machine.

An examination of a simple program that makes an attempt to pass this test, and in some cases does pass this test, would be appropriate.

Even with Professor Turing's test, several objections and questions to the entire concept of AI exist. These will be examined, along with some rebuttals.

After examining these points and the Turing Test in the light of the open definition derived from the meaning of pertinent terms, what conclusions could be drawn? Can the questions be answered? Can a computer possess intelligence? If a computer does indeed possess intelligence, is that intelligence necessarily an imitation of a human intelligence? Or is it something new?

Before proceeding further, certain terms that pertain to computer operations should be defined as they are used in this paper. When the term "computer" is used, a digital computer based on a single processor performing sequential operations will be meant unless otherwise specified. By "digital," what is meant is that this computing machine uses a binary-based logic system that is directly built into the circuits of the machine. By "single processor performing sequential operations" is meant that this computing machine will have only one set of circuits that does the actual processing of information and that it processes this information by treating it as a single long line of items, and takes each item one at a time to act upon. The term

"program" shall refer to the series of instructions kept in a computer's storage area that guides the steps of processing.

CHAPTER 2

DEFINITIONS FOR "INTELLIGENCE" AND OTHER TERMS

Experience has shown that for rational discussion to be possible, all participants in the discussion must agree on the way that the key terms in the discussion are going to be used and understood. Often, some term is found which cannot be easily limited or defined. This is particularly true of the terms that are the point of this discussion. Often, the true nature of the problem or situation becomes clear in the attempt to reach mutual agreed meanings for terms.

The word "intelligence" is a difficult word to define, and what constitutes an acceptable meaning for this word is relative to the point of view that individuals bring to the discussion. One source gives a common usage definition of the word "intelligence" to be "...the faculty of understanding; the capacity to know or apprehend, the intellect as a gift or an endowment..."¹ It goes on further to define "intelligence" as

¹Webster's New International Dictionary of the English Language, rev. ed. (1950), s.v. "Intelligence."

The capacity for knowledge and understanding, especially as applied to the handling of novel situations; the power of meeting a novel situation successfully by adjusting one's behavior to the total situation.²

This same source defines the word "intellect" as

The power or faculty of knowing, as distinguished from the power to feel and to will; sometimes, the capacity for higher forms of knowledge, as distinguished from the power to perceive and imagine; the power to perceive relationships, to judge and to comprehend; also, ability to think; understanding.³

These two dictionary definitions give some feel for the way the word "intelligence" is commonly used in everyday conversation. The meaning is perhaps a little broad, and perhaps a little vague, but it serves for everyday use.

Researchers in artificial intelligence, defining the word "intelligence," have apparently either openly or by error placed themselves into one of two positions.

Researchers position themselves "openly" by defining the term in their speeches and published works, and stick to their definition. They position themselves "by error" in that they fall into the error of sliding meaning discussed earlier. The researcher defines the word, then in the way he uses the word he changes its meaning, bringing in something from outside of his original definition, and therefore reaching conclusions of doubtful validity. Two positions on meaning for the word "intelligence" occur

²Ibid.

³Ibid., s.v. "Intellect."

because the computer scientists have moved into the philosopher's realm in attempting to define the word. Most computer scientists who believe that computers can think fall into the broad meaning camp. Most computer scientists who do not believe that computers can think fall into the limited meaning camp.

Within the broad meaning camp, there also exist several possible positions. One of the larger groups defines the word "intelligence" broadly by claiming that behavior that appears intelligent over a broad range of circumstances to a human being is indicative of intelligence in the object or creature that exhibits the behavior. Alan Turing, as set forth in the article mentioned in the introduction and which will be examined in some detail later, holds a behaviorist position. The whole point of the "imitation game," the Turing Test, is to determine if a machine is intelligent by observing its behavior over a wide range of subjects and situations, with the limitation of only communicating by remote means, e.g. a keyboard and display device.⁴ John Haugeland, a philosopher at the University of Pittsburgh, agrees completely with this definition of intelligence. He says, "How shall we define

⁴Alan M. Turing, "Can a Machine Think?" Mind 59, No. 236 (1950) 2099-2100.

intelligence? ... For practical purposes, a criterion proposed by Alan Turing (1950) satisfies nearly everyone."⁵

Another position within the broad meaning camp is the group of researcher/philosophers who define "intelligence" as the sum of various faculties or abilities.

Forget the validity of the IQ tests themselves for a moment, and ask whether the ability to perceive relationships between objects and patterns, to solve problems of varying complexity and to see fresh ways out of mental puzzles, can reasonably be called intellectual activities. I think all will agree that these are at least some of the pertinent factors, though there may well be others.⁶

Christopher Evans, the researcher in computer science who holds the view expressed here, feels that abstract abilities, derived from IQ tests, can compare to human daily activities that are labeled "intelligent," and then can compare to the daily activities of animals. He asks his readers to accept a demarcation to the meaning of "intelligence" that would include animals as "intelligent." Then, looking at computer operations, he moves his demarcation outward to include computers as members of the group labeled "intelligent." The justification for the

⁵John Haugeland, Artificial Intelligence: The Very Idea (Boston: MIT Press, 1985), 12.

⁶Christopher R. Evans, The Micro Millennium (New York: The Viking Press, 1979), 156.

lines of demarcation is that all members of the group that possess "intelligence" exhibit and use certain faculties or abilities in their everyday operations.⁷

A key component for the entire broad-meaning camp is that whatever precise meaning is given to intelligence, this meaning must include the essential property of being quantitative, that is, that whatever the characteristics that make up "intelligence," these characteristics must be measurable, or in some way reducible to some numeric value. Whether one speaks of the probability of certain behavior when compared to random chance or compared to certain versatility factors, numerical analysis is an essential part of the definition of "intelligence."

For the limited-meaning camp, the essential property of "intelligence" is a qualitative one. The position given here is that machines and animals, while capable of imitative behavior, are lacking in some quality that distinguishes imitative behavior from "true intelligence." Implicit in this position is the argument that no amount of quantitative increase in ability to imitate human behavior, even if this behavior becomes indistinguishable from "true" human behavior, will result in a qualitative leap to "intelligence."

⁷Ibid., 157.

A major opponent of AI, both on epistemological and ethical grounds, is Joseph Weizenbaum, a professor of computer science at MIT. He argues,

Man faces problems no machine could possibly be made to face. Man is not a machine. I shall argue that, although man most certainly processes information, he does not necessarily process it in the way computers⁸ do. Computers and men are not species of the same genus.

Morton Hunt expresses the same viewpoint in a different manner.

Until artificial intelligence can duplicate human mental development from birth onwards; until it can absorb the intricacies and subtleties [sic] of cultural values; until it can acquire consciousness of self; until it becomes capable of playfulness and curiosity; until it can create new goals for itself, unplanned and uninstigated by any human programmer; until it is motivated not by goals alone but by some restless compulsion to be doing and exploring; until it can care about, and be pleased or annoyed by, its own thoughts; until it can make wise moral judgements -- until all these conditions exist, the computer, it seems to me, will not match or even palely imitate the most valuable aspects of human thinking.

A closer look at the common usage for the term "intelligence" shows that it reflects both of these two positions. The first portion of the definition of "intelligence" and the definition of the word "intellect" display a facet of the limited-meaning position in that it

⁸Joseph Weizenbaum, Computer Power and Human Reason: From Judgment to Calculation (New York: W. H. Freeman and Co., 1976), 3.

⁹Morton Hunt, "What the Human Mind Can Do That the Computer Can't", in Philosophy and Contemporary Issues, ed. Milton Goldinger and John R. Burr (New York: Macmillan Publishing Co., 1984), 375.

talks of "higher forms of knowledge" and "the intellect as a gift." These are value judgements that display a qualitative orientation. The second half of the definition for "intelligence" displays a behaviorist approach, speaking of "handling novel situations." Common usage offers no criterion on which to base an answer to "Can computers think?".

"Consciousness," or "sentience," is another key term for this discussion. Here, common usage gives a meaning for "sentience" to be "Sentient being or state; consciousness; mental life." The Dictionary of Philosophy and Psychology gives a meaning for "consciousness" as

The distinctive characteristic of whatever may be called mental life... Wherever there is not total unconsciousness, in the sense that we attribute unconsciousness to a table or a log of wood, the existence of some form of mind we denote by the word consciousness.

Again we find workers in the field of AI split into the same two camps. The members of the broad-meaning position claim "consciousness" to be only detectable by observation of behavior. Being observable makes it to some degree measurable and therefore quantitative. The limited-meaning group claims, as shown by Morton Hunt above, that they consider "consciousness" to be a quality of "intelligence" that is not measurable in the way that the broad-meaning people claim. Descartes' famous "I think, therefore I am," is at stake here. Adherents of the broad

meaning will claim that "consciousness" evolves merely as a by-product of increasing the quantity of connections in the computer system. For the members of the limited-meaning group, "consciousness" is an essential property of human beings, or perhaps even animals, and therefore restricted to biological forms. The key question here is "Is consciousness even an essential property for intelligence?"

"Thought" and "thinking" are two more words deeply entwined with the concepts of AI. Common usage has defined "to think" as

To form or formulate in the mind; to have the thought, notion, or image of; to conceive; imagine... To have in one's mind as an intention or desire; to purpose; to hope; also, Now Rare, to intend doing...¹⁰

The word "thought" is defined in common usage as "Act or process of thinking; mental concentration on ideas as distinguished from sense perception or emotions; cogitation."¹¹ These common usage definitions are somewhat circular in that "thinking" is to have "thoughts," while "thought" is the act of "thinking."

A view of philosophical usage could perhaps avoid this logic problem. The Dictionary of Philosophy and Psychology defines "intelligence" or "intellect" as "The faculty or capacity for knowing; intellection... There is a tendency to apply the term intellect more to the capacity for

¹⁰ Webster's, s.v. "Think."

¹¹ Ibid., s.v. "Thought."

conceptual thinking." The term "concept" is then defined in the same work as "Cognition of a universal as distinguished from the particulars which it unifies...", while this is the meaning for the term "cognition":

The being aware of an Object... it may well be questioned, however, whether in current usage cognition does not imply judgement, at least in a rudimentary form, as well as presentation of an object.

One philosophical viewpoint defines "thinking" as something like "the process of utilizing one's capacity to be aware of and judge upon some universal object which unifies or groups together a number of particular objects." However, this particular philosophical view seems to link the concepts "intelligence" and "thinking" together very tightly. To simplify matters in this thesis, and to prevent problems in definitions and meanings, we will declare our meaning for "thinking." This is simply that "thinking" is the process of using "intelligence," whatever we define intelligence to be. This dodges the question, however, of the possibility of thinking in a "stupid" or unintelligent manner.

There are other words that would seem to require some attention during our declaration of usage for words. "Memory," "emotion," "imagination" and the appetitive abilities like "desire," "wishing" and "willing," are all terms that the limited-meaning camp would insist, as Morton Hunt did, are qualities that are found as part of "intelligence." The broad-meaning camp would declare these

terms to be labels for various abilities or faculties that also exist in the mind, and that some, like "memory," are in fact part of "intelligence" but that some, such as "emotion," are not. The appetitive abilities are mentioned in some of the common usage definitions, but here too the split is apparent, for in the definitions for "intelligence" the "power to will" is set apart from "intelligence" while in defining "thinking," "intention or desire" is considered integral to the action. In the same way, "imagination" is also split between the two terms. This two part division again signifies a philosophical choice between a quantitative, analytic, and to some degree mechanistic, approach and a qualitative, holistic approach.

The only philosopher I have found who seems to make an attempt to straddle the two positions is John R. Searle, whose argument states: "'Could a machine think?' The answer is, obviously, yes. We are precisely such machines."¹² However, the thrust of the article just quoted is that a machine sufficiently complex enough could be labeled "intelligent," but that true intelligence requires intention or will. Searle's point, based on the philosophical theory of intentionality, is that computers as defined in the introduction to this paper cannot think,

¹²John R. Searle, Minds, Brains and Programs, in Mind Design, ed. John Haugeland (Montgomery, VT: Bradford Books, Publishers, 1981), 300.

because they cannot "intend," and without intent they cannot understand meaning. Meaning for Searle then ties knowledge with intent. While accepting a great many of the broad-meaning group's tenets, he ultimately falls into the limited-meaning camp due to his acceptance of the appetative abilities as essential qualities for "intelligence."

Having examined some of the meanings attached to the word "intelligence," the word "artificial" requires consideration for an understanding of the term "artificial intelligence." "Artificial" seems to have two meanings in common usage. The first meaning is mimicry, imitation not capable of functioning in the same way as the object imitated, or "fake." The second meaning in common usage for "artificial" is creation, something man-made, manufactured versus natural, but possessing true functionality in its own right, something that is the product of artifice.

Opponents of the broad-meaning group apply the first common usage meaning of "artificial" to the term "artificial intelligence," to arrive at a definition for AI that states that AI is the attempt to imitate human intelligence that will ultimately fail because certain qualities cannot be imitated by machines of any kind. Morton Hunt's position clearly illustrates this viewpoint in the case he made against the broad-meaning definition for "artificial intelligence."

Supporters for the broad-meaning group claim the second usage of "artificial" and define AI to be the attempt to create a functioning intelligence in a man-made machine. For this second group, imitation of human behavior labeled intelligent only serves as a criterion for functionality. They declare this despite the fact that the Turing Test relies on successful imitation as the criterion for intelligence.

Richard Dorf supports this side of the argument, and defines the term as "ARTIFICIAL INTELLIGENCE: The characteristic of a computer system capable of thinking, reasoning and learning (functions normally associated with human intelligence)."¹³ He then goes on to say:

But it must be admitted that there exists some field of behavior in which men and machines can coexist and in which they can be compared. It has been noted that one often regards an action as "intelligent" until he understands it. In explaining the action, it often becomes routine and mechanistic rather than intelligent.¹⁴

Christopher Evans states his position supporting the broad-meaning camp thus:

For some time computers have been capable of perceiving relationships between different categories of objects in the external world (though admittedly in a rather laborious manner) and are also able to solve certain types of problems. So again, unless one was prepared to take a stand and define intelligence as something that could only be present in biological systems, and not

¹³Richard C. Dorf, Computers and Man (San Francisco: Boyd & Fraser Publishing Co., 1974), 385.

¹⁴Ibid., 386.

things made out of electronic components, one is forced to accept the notion of machine intelligence.¹⁵

This analysis of meanings for the terms used to discuss AI has revealed the opposed philosophies of at least two groups that are at the front of the AI movement today. The limited-meaning group advocates a qualitative approach that rejects any possibility of any type of machine intelligence. The broad-meaning group supports a quantitative approach that permits the possibility of some kind of machine intelligence. These distinctions and definitions will guide us as we attempt to answer the basic question of AI put in its simplest terms, "Can a computer think?" .

¹⁵Evans, 157.

CHAPTER 3

THE TURING TEST AND ELIZA

The "Turing Test" has been mentioned in previous sections as a major weapon used by those who hold to the broad meaning for "intelligence." Its importance to AI and to the questions that are raised for and against the idea of intelligent machines cannot be understated. To understand the approach, and to comprehend the genesis of this position, as well as its strengths and weaknesses, a detailed study of Alan Turing and his "Turing Test" is necessary.

Alan M. Turing was a pivotal figure in the history of computer science and in the development of AI as well. A British scientist and mathematician who worked on the Enigma code-breaking project during World War II, Turing took part in several discussions on the possibility of building a machine capable of thought, and of the means for testing such a machine. These discussions served as the basis for the article that Turing later published in the journal Mind in 1950. In it he proposed a test that could be applied to a machine in such a way that the intelligence or the non-intelligence of the machine would be clearly determined by any number of judges, impartial or otherwise.

This examination or experiment is commonly referred to in the literature and the jargon of the AI labs as the "Turing Test." Turing himself called it the "imitation game."¹ The rules of this game are simple. Turing proposes as an example two people in two rooms. One is a man and the other is a woman. A judge or interrogator sits in the third room. The judge communicates with the two people through a keyboard and some type of output device, such as a printer or a monitor display screen. The judge may ask any questions which do not require any direct sensory perception by him of either of the two people. By means of these questions the judge must determine which of the two is male, and which is female.²

Turing then makes the next obvious step, and proposes that a machine be substituted for one of the people. The test now becomes whether the judge can determine which of the two entities communicating over the apparatus is human, and which is not. The key point that Turing stresses is that if a machine can continually and consistently pass for human for any number of judges, biased or not, then the machine must be judged to be intelligent, or to be human.

Turing is part of that group of philosophers and scientists who hold for a very broad meaning of the term

¹Turing, 2099.

²Ibid., 2099-2100.

"intelligence." He believes that "at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."³ Turing also defends his opinion by answering objections that he felt would be commonly raised. Some of these objections include what Turing labeled "the theological objection," "the mathematical objection," "the argument from consciousness" and "the argument from informality of behavior." In his defense, Turing assumes a behaviorist stance, arguing that if a human interrogator cannot mentally or physically experience the perceptions and state of being of the machine, or experience and feel what it is "feeling" or "thinking," in the same way that the interrogator cannot experience in any way what another human being is "feeling" or "thinking," the only way to determine the presence of intelligence realistically and objectively is by comparison to a system considered to possess intelligence beyond any doubt. In this case, the system is a human said to possess an intelligence, as versus an animal labeled intelligent.

Turing places no restrictions on the imitation game beyond those already stated. He does offer guidelines to explain his position and to describe what he feels constitutes a reasonable test. He offers as a guideline

³Ibid., 2103.

that a digital computer, as defined previously, is perhaps the best machine to test. He grants that the computers of his day are not entirely up to the task, but he projects that near the end of the twentieth century digital computers would begin to approach a level of complexity that would allow them to fool the average interrogator 30% of the time after five minutes of questioning.⁴

This understanding of what constitutes the Turing Test, and what it claims to do, permits a clearer understanding of the following example of a program that has had a profound influence both positively and negatively on the development and justification of AI.

Joseph Weizenbaum, a researcher at MIT, wrote a program in the computer language LISP that he called ELIZA. ELIZA was designed to imitate a psychologist working with a patient in the style of psychoanalysis developed and advocated by Carl Rogers. This project was originally designed to be an experiment in using natural language to communicate with a computer, rather than using a designed computer programming language to communicate. The program ended up as a limited but powerful example of the Turing Test at work. Several of the people who interacted with the program (both the actual research subjects and some researchers who had access to the program) were found to

⁴Ibid.

have developed a relationship to the machine/program combination that was considered by them to be a two-way communication between equals.⁵

Weizenbaum was amazed at these attitudes. "My own shock was not administered by any important political figure espousing his philosophy of science, but by some people who insisted on misinterpreting a piece of work I had done."⁶

First, this situation did not satisfy the initial conditions set forth by Turing. It was not a controlled test. There was no human involved for comparison. The humans involved were intellectually aware that they were interacting with a computer.

Second, the subject matter of the interactions was limited to the areas of psychological problems and psychoanalysis. The questions and answers, and the interactions in general, were not as free ranging nor as probing as Turing required. "A number of practicing psychiatrists seriously believed the DOCTOR computer program could grow into a nearly completely automatic form of psychotherapy."⁷ (DOCTOR is the name Weizenbaum gave to ELIZA in its role as a psychotherapist.)

⁵Weizenbaum, 6-7.

⁶Ibid., 2.

⁷Ibid., 5.

Third, as Weizenbaum pointedly observed, the human subjects were not disinterested observers, but rather were individuals who formed an emotional attachment to the program.

I was startled to see how quickly and how very deeply people conversing with DOCTOR became emotionally involved with the computer and how unequivocally they anthropomorphized it. Once my secretary, who had watched me work on the program for many months and therefore surely knew it to be merely a computer program, started conversing with it. After only a few interchanges with it, she asked me to leave the room. Another time, I suggested I might rig the system so that I could examine all conversations anyone had with it, say, overnight. I was promptly bombarded with accusations that what I proposed amounted to spying on people's most intimate thoughts; clear evidence that people were conversing with the computer as if it were a person who could be appropriately and usefully addressed in intimate terms.

Despite the points listed above, proponents of the Turing Test hold forth this example and state that even within the limited setting that ELIZA provided, some human being interacted with the computer as if it were intelligent.

ELIZA was an important formative event in AI. Those who held with Turing and the broad-meaning camp heralded it as proof and vindication of their position. Weizenbaum and others in the limited meaning camp hold that its obviously limited context and clearly non-intelligent behavior serves as justification for debunking the whole concept of AI, not just the Turing Test.

⁸Ibid., 6-7.

ELIZA as programmed by Weizenbaum was an extremely complex and lengthy program. Several simpler versions exist that allow at least some interaction and indication of what the original ELIZA did. One version was published in the book More Basic Computer Games. The program listing, written in the Microsoft BASIC programming language by Steve North, is included in Appendix A, as well as a sample of an interactive session with this version of ELIZA, in Appendix B.⁹ The human's inputs are marked indented from ELIZA's comments.

⁹Steve North, ELIZA, in More Basic Computer Games, ed. David H. Ahl (New York: Workman Publishing, 1979), 56-58.

CHAPTER 4

THE OBJECTIONS AND THEIR REBUTTALS

The Turing Test and ELIZA, as well as all the descendants of ELIZA that fill the literature of AI today, serve as the bastion and bedrock for the broad-meaning camp in AI. The Turing Test and Alan Turing himself are referred to in the literature with an appeal to almost divine authority. For example, Apple Computer's Jean-Louis Gassee stated in an article that "There is one infallible way to find out whether a machine is intelligent or not. It is the test of Turing, the English mathematician..."¹ Fred Reed, in his column "Defense" in Federal Computer Week said "If we accept the definition that if it looks as if it thinks then it does think, then some military systems ... are tiptoeing toward being intelligent."² Notice here, however, that Reed does not refer to these systems in terms like "he" or "she"; it remains an "it" despite its apparent intelligence. The author Arthur C. Clarke, in his article "Are You Thinking Machines?," said in reference to the

¹Jean-Louis Gassee, "This Idiot Savant: The Proper Study of Machine," ComputerLand Magazine 3, no. 2 (January/February 1988): 34.

²Fred Reed, "Overlooking the Marvelous," Federal Computer Week, 3 October 1988, 12.

1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

definitions of terms. A group of researchers and supporters of AI advocate a broad meaning of "intelligence" and a belief that machines can be made to be able to think. An opposing group advocates a limited meaning of "intelligence" and generally believe that machines cannot and will not be intelligent.

CHAPTER 5

CAN A COMPUTER THINK?

Is the goal implied by the name, "Artificial Intelligence," theoretically or practically attainable? Could a computer possess any kind of intelligence? Can a computer think? At present, the answers to these questions depend on whom we ask.

The Broad Meaning of "Intelligence" Camp Answers "Yes"

Those who respond in the affirmative to "Are computers intelligent?" and "Can a computer think?" generally support a quantitative, broad meaning for intelligence. This broad-meaning camp defines and uses "intelligence" in a far-ranging manner, thus applying it not only to human beings, but to some animals, and to some possible machines. This camp maintains that intelligence is some ability or process that is measurable, and it can be possessed in varying amounts or degrees. These degrees form a scale with humans at the high end of this intelligence scale, animals in the middle, and inanimate objects at bottom. Through evolution and development by human agency, computers are seen by the broad meaning camp as moving up the intelligence scale, attaining steadily higher degrees of

intelligence, and soon to equal humans in the amount of intelligence they possess.

The methods of measurement used by this group are based on observation and comparison of behavior of all beings that might possess "intelligence." In this measurement the level of human intelligence is taken as a standard for a scale, somewhat analogous to the choice of the boiling point of water as the standard for the temperature scales. Internal processes that are not directly observable are only important for the behavior they cause. The closer the computer behavior comes to the standard of human behavior, the more the internal processes of the computer become important, but only as a means, not as a goal in themselves.

The first strength of the broad-meaning approach to AI is that they have a standard by which to measure progress toward the goal of creating intelligence. Second, their method is based upon the accepted practice for science, observation of the external phenomena associated with the object of study. Third, they have openly declared their definition of "intelligence." Their definition has such a wide domain that almost all objections can be answered. They can be answered either with the claim that the point of the objection is included as integral to the broad-meaning position, or with the claim that while the objection has some validity, it applies only to a limited set of

circumstances, and does not invalidate the broad-meaning position.

The broad-meaning position also has several weaknesses. The first weakness is that while the broad-meaning definition of "intelligence" ranges widely enough to answer most objections, it also ranges so widely that it could be asserted that by claiming everything, they prove nothing. For example, if apes are intelligent, shouldn't we include dogs? If we include dogs, shouldn't we include cats? If we include cats, shouldn't we include rats? What shall we decide about the rats' fleas? Where do we lay a line of demarcation between intelligence and nonintelligence? The second weakness is that while they claim human intelligence for a standard, they have been unable to come close to that standard, even by their own measurements, except in very limited and controlled contexts such as ELIZA or computer chess play. Third, while their method uses "objective" observation, the comparison to the standard is judged subjectively. Even in the Turing Test, no effort is made to establish a "double blind," that is, moving the judges one step or layer back from the tests to reduce subjective interaction and influence on the observation. In other words, a judge who is also the eyewitness cannot be said to be objective, either in a court of law or a scientific laboratory.

The Narrow Meaning Group Says "No"
to Machine Intelligence

Those who advocate a negative answer to the question, "Can a computer think?", for the most part, propose a qualitative, experiential definition for "intelligence." They limit the domain of application for "intelligence." The limited-meaning group defines "intelligence" as a characteristic unique to human beings. They hold that while other beings such as animals and computers may possess elements of intelligence, and exhibit some behavior similar to human behavior, no other creature possesses or exhibits intelligence in its entirety.

Because this group stresses the qualitative and experiential aspects of intelligence, objectivity and measurement are not useful in determining the intelligence or lack thereof in computers. The key distinguishing point for this group is that intelligence is not only exhibited, it is experienced. The subjective is important, and so the internal processes of intelligent behavior become the most important area of study. To judge mimicked behavior as a sign of equal intelligence is as invalid for the limited-meaning group as to judge the behavior of a wind-up toy mouse as indicative of life.

The strengths of the limited-meaning position appear in the objections it presents to the broad-meaning approach. First, the theological objection takes the area of

intelligence out of the domain of scientists by adding a supernatural, unobservable element. Scientists, basing their methods on empiricism and experimental manipulation, have only limited means by which even to approach this area. The second objection, the missing ingredient objection, is a direct assault on the logical weakness of the wide domain of the broad-meaning definition of "intelligence." The third objection, while in the style of the missing ingredient objection, stresses the experiential approach of the limited-meaning supporters. An intelligent being is conscious, not only of its environment, but of itself and its own intelligence. This is a subjective state that humans have experienced and described. The limited-meaning camp denies that a machine will ever "experience" anything remotely resembling consciousness of self. The moral objection questions the practicality and usefulness of AI, by questioning the ethics and motives for attempting to create intelligence.

The limited-meaning approach to AI is itself open to attack. The theological objection is outside the domain of science as it is practiced today, and so its validity as an objection to a scientific endeavor is questionable. The missing ingredient objection has strength, but has often been and is still used in an ad-hoc manner, in the form of "Okay, the computer may be able to do or to possess 'x' now, but it still has to do or come to possess 'y' to be

'intelligent.'" The consciousness objection's strength is its weakness. The limited-meaning camp claims the existence of a subjective state of self-awareness in humans. A philosophical or scientific proof for the existence of such a state would require some objective element, which would permit the broad-meaning camp to incorporate it into their standard for intelligence, and would bring the validity of the claim for subjectivity into question.

While the majority of researchers in AI fall into one of these two camps as we have described them, not all fit in so neatly. There are researchers who define intelligence in the broad manner, but who deny the feasibility of AI and of machine intelligence. There are also scientists and philosophers who support the limited meaning for intelligence, yet are extremely optimistic about attaining an intelligent computer.

If we restrict the scope of our search in time and potential by specifying that we mean a computer with today's technology and programming, that is built or could be built today, the answer must be no. A computer built with today's technology, programmed with today's algorithms, cannot be considered intelligent. This is the consensus of the vast majority of all who work in or study AI.

The belief in the eventual accomplishment of AI, in the accomplishment of machine intelligence, is based on the

projection that technology will eventually enable a sufficiently faster, larger computer to do what today's computers cannot. Current research in computer science in the field of complexity indicates that this will not be the case. The basic idea is that for certain types of computational problems, even a ten-fold increase in storage capacity and processing speed would only allow a fractional increase in information processing. Most problems in AI are of this type. Improving computer hardware performance while retaining the basic structural design is similar to the development of automobiles. In the early years, an improvement in performance brought dramatic difference in how cars were used on the road. Today, improvement in automobile technology has not had the effect on the actual practice of driving to the degree found earlier. This is the state of computer development that most computer designers agree exists today.

However, complexity theory, the field of computer science that determines how complex a problem is and how best to solve it if possible, does not close the door on AI. If hardware performance cannot do the job, the algorithms, or problem solving steps, used in the programs possibly can. The best computer chess programs today do not perform brute force searches for the winning combinations. They are programmed to apply rules to reduce the number of choices without actually checking them, and this has greatly

increased their level of performance in play against humans, as well as speeding up their play. This is one area that gives hope for AI.

Another way around the hardware performance problem is to redesign the computer. To continue our analogy, to go farther and faster, instead of improving the car you build an airplane, or even a rocket. In the realm of computer architecture, the new designs that are showing great promise are parallel processing and neural networks.

Parallel processing is based on a simple idea. If one processor is not fast enough, then add another. Expressed more generally, if n processors are not enough to do the job, add more processors. The difference between this approach and the attempt to improve current architectures is the difference between approaching a problem sequentially, and approaching a problem as a whole, each processor taking its small part and contributing to the solution. The Connection Machine, built by a Boston computer company with ties to MIT, has sixty-four separate processors built in. There are plans by several companies to build machines with thousands of processors. The difficulty that is holding up progress with this field is that while parallel processing can and does give tremendous increases in the speed and complexity of processing, not all problems, and particularly not all AI problems, can be formulated in a way to permit solution on a parallel design machine.

Neural networking is a computer architecture that attempts to imitate the actual structure of the brain's neural pathways. Processors are treated as corresponding to neurons, and the processors are connected to each other in a way similar to the neural connections in the brain, a more complex structure than simple parallel architecture. Again a huge potential exists to increase processing speed and complexity; however the degree of difficulty in programming is also increased. The possibility of the existence of problems that may not lend themselves to solution on neural computers may also be a problem here.

New algorithms, neural and parallel computer designs, and the increase in the performance and efficiency of current digital computers are the potentials that AI supporters base their optimism on. These are the tools of the future that will be the ancestors of the thinking machine.

In our work up to this point, "intelligence" has been the watchword and the main premise for the arguments and positions. Artificial intelligence has two components. The second component, "artificial," if considered and incorporated into the discussion, could and will modify our basic question, and the possible answers to it.

"Artificial" is used in two ways in AI. The first usage describes the mimicking of the behavior and function

of human intelligence, "faking" it, as it were. Those opposed to AI say that this is all that AI has been able to do, and this is all that AI will ever be able to do. Based on this usage, opponents of AI claim that machine intelligence is not "real" intelligence, anymore than silk flowers are "real" flowers. Parrots mimic human speech, chimpanzees mimic human behavior, and yet neither is declared to possess human intelligence. Computers may mimic human speech and interaction, but that is not a basis for claims of intelligence equal to human intelligence.

The second usage of "artificial" in AI describes creation by human agency of something that possesses intelligence. AI supporters claim that machine intelligence is "real" intelligence, in the same way that human created x-rays are x-rays, that human created water is water, that human created chemical elements are chemical elements. Artificial in this usage implies creation, not mimicry.

Radiation from human sources is not functionally different from naturally occurring radiation. A physicist would not differentiate between human generated radio waves and quasar generated radio waves. Light from a light bulb illuminates a book as well as light from the sun. Yet all these human created things are different from their natural counterparts. Radio waves from Earth have a different frequency than those from a quasar, as light bulb light has different color and energy levels from those of sunlight.

The end results, the effects of man-made objects as compared to naturally occurring objects, are very close if not identical. Yet the composition and character of the man-made objects do distinguish them from naturally occurring phenomena.

To define "intelligence" in such a way as to make it the exclusive province of human beings denies the common usage and experience of intelligent behavior in other types of beings. We speak of smart dogs, and intelligent apes. A chimpanzee's behavior may at times mimic human behavior, but some of its behavior is self-motivated and directed, displaying qualities that we label intelligent. Is this intelligent behavior at a human level? No, it is at a chimpanzee level. It is chimpanzee intelligence, animal intelligence, and that is how we speak of it in an everyday manner.

In contrast to limiting the scope of "intelligence", to define "intelligence" in so broad a manner that any and all behavior will be labelled intelligent is to cheapen the word. To go to the extreme of labelling the actions of a thermostat on the wall as "intelligent" and "intentional" strain belief. Moreover, to equate all types of intelligence raises its own problems. Is a human being who is mentally retarded through organic damage to his brain as intelligent as most humans? Is his intelligence a human one? Some animals respond more "intelligently" than some of

these poor people. It is just as illogical to declare that because white rats display the ability to learn a maze, that their intelligence is equal to a human intelligence. There is obviously a difference present in these comparisons.

Neither the broad nor the limited definitions of intelligence give a complete picture. Light can be described as having a wave structure, and as having a particle structure. Intelligence does possess characteristics that are measurable and quantifiable. Intelligence also possesses qualities that are immeasurable, and can only be experienced through their effects or subjectively and internally. A correct definition of intelligence must attempt, as common usage does, to incorporate both sides, from quantitative and qualitative approaches.

Do we have, then, a criterion for intelligence? Can we declare without error that object "x" is intelligent? If not, an approach from a different direction is called for.

CHAPTER 6
THE HUMAN QUESTION

Up to this point, we have determined some of the usages for the terminology of AI, and we have examined the two major camps representing opposing views in the field. But is this enough to answer our questions? Perhaps, for the computer scientist, it may be enough to say that AI is a possibility, that AI possesses the potential to be a fruitful area for study and experimentation. The philosopher requires more than this to be satisfied, and must take one step beyond what we have covered so far.

Synergistic Humanity

The philosophical position we must now consider could best be labelled as the synergistic view of humanity. Synergy occurs when the whole object is greater in function or scope than the sum of the functions or scopes of the individual parts or components that came together to compose it. Organic beings, as a class, are synergistic wholes. Life is more than amino acids enclosed in a semi-permeable

membrane. Animals are more than a joining of muscle, bone, blood and connective tissues. Our humanity is a product of a synergistic whole.

This view of humanity does not perceive humans as being capable of simple analysis into basic components. Each facet of humanity is related to all the other facets, and to examine one facet in isolation is to lose some of the essential characteristics that make it part of humanity. According to this position, the determination of whether or not some "x" is intelligent cannot be made apart from considering the level of creativity, imagination, emotion, communication, humor, the object possesses. The correct study of intelligence therefore requires that we study what it is to be human.

The field of artificial intelligence, approached from this position, must understand that intelligent behavior and intelligent communication encompass more than words projected on a monitor, or a computer-chip voice coming from a speaker. When we communicate as human beings with other human beings, all the senses come into play. Not only the content of the words, but the tone of voice, the facial expression, the position of the body and the limbs, and hundreds of other subtle signs and symbols are sent and received in the simple greeting, "Good Morning!". The look of intelligent behavior, the way intelligent communication "feels" to us as we participate in it, involves a

synergistic understanding of what a human is, and what a human does.

From this viewpoint, the most telling criticism of AI, and of many other communications media, is that it is "flat" or "two-dimensional." The Turing Test suffers from this objection. Professor Turing, in attempting to reduce intelligence to its most basic form, reduced it beyond the point of true communication to a level of simple information exchange. While this information exchange could eventually reach a level of sophistication that would cause it to be confused with communication, it would still be lacking the essential quality of action and interaction.

Though humans exchange information with words over display screens or telegraph or fax machine, primary communication comes from face to face meetings. Turing ruled out such actual physical contact. This stacked the deck in his favor. If we only examine the endgrain of a piece of lumber and then declare that the lumber is structurally sound throughout its entire length, we would be very wrong in making that claim. To claim that the Turing Test is a valid test of information exchange in a natural language would be correct. To claim that the Turing Test is a valid test of intelligent communication is incorrect, because it does not test all the other aspects of what humans do when they communicate.

In support of the synergistic view of humanity, there is the tendency of human beings to anthropomorphism. People ascribe human characteristics to many objects with which they interact with at different levels. Mickey Mouse is an obvious example of anthropomorphism. A ship becomes "she", and guns are "he." Jerry Pournelle, a science fiction writer and columnist who often writes about computers, regularly refers to his IBM PC as "Lucy," and his older CP/M computer as "Zeke." In the movie, 2001, A Space Odyssey, H.A.L., the ship's computer, was treated as both machine and human, with disastrous consequences when the machine was asked to make human decisions and act as a human, that is, to lie.

Anthropomorphism is a sign of the human desire to communicate intelligently. Mickey Mouse is more real and more intelligent to us than a computer is, because he is more like us. Mickey Mouse has arms and legs, wears clothes, laughs and cries, smiles and frowns. Do we claim that Mickey Mouse is intelligent? No, but in the synergistic view of intelligence, Mickey Mouse is closer to being intelligent than any computer today.

"You win!" cry the scientists. They will build a human shaped container for their computer, with arms and legs, with a face that can smile or shed tears, with a mouth for talking and ears for hearing. This container will walk like a man, and talk like a man. Is it intelligent? -- and

if we shall say no, do we not feel that at least we have gotten a lot closer to the truth of the issue? Closer perhaps, but will it get tired? Will it bleed? "You're asking us to build a man!" the scientists accuse. "Yes" reply the synergists.

Issac Asimov is a prolific writer of science fiction and science fact books. He is famous for his series of stories and novels that deal with robots and their interaction with humanity. One story, "The Bicentennial Man" in particular deals with this issue of anthropomorphism. The plot of the story revolves around a robot that develops creativity. His owners allow him to keep the money from his efforts, and soon he is able to buy his freedom. The remainder of the story recounts the robot's efforts to be accepted by human society as an intelligent being. He begins to wear clothes, then he gets a skin-simulating covering with a face capable of human expression; and then starts replacing his internal mechanisms with the same artificial organs that humans of his time use for transplants. He takes his case to the Supreme Court, and finally makes the ultimate sacrifice. To be recognized as human, the robot must allow himself to die. The key issue here is that for the robot to be considered intelligent, he had to become human.

The synergistic view of humanity denies that intelligence can be separated from the human being.

The concept of "intelligence," whether fashioned by the broad-meaning camp or the narrow-meaning camp, only becomes coherent in light of what human being as complete human being is able to do. Intelligence is part of the synergy that is humanity, and for a machine to be intelligent, it must become human.

Mechanistic Humanity

The opposing philosophical viewpoint is the mechanistic view of humanity. A direct outgrowth of the mechanistic science in England and France in the seventeenth century, this position views humanity as composed of components in the same way a clock or a motor is composed of pieces. If we could assemble all the same parts correctly, we could build a man. Just as importantly, we could understand any of these components by studying it in isolation. To understand vision we study the organs of sight. To understand speech we study the voice box, the mouth, the lips, and the nerve paths to these areas. To understand the human mind, we study emotions, imagination, or intelligence.

The importance of this viewpoint to AI is that if intelligence is merely a component found in the human brain, then the outer form is not important. We could isolate whatever it is that intelligence is, and put it into whatever container we choose. If intelligence is a

component, it can be plugged into a machine as easily as into a human.

In the mechanistic view, communication is information exchange. Communication is valid, whatever the media. Humans communicate over display screens and talk on the telephone. If it is a computer on the other end of the wire, and we communicate with it, is it not as intelligent as we are? The written word is communication. If it is not, of what value are books and theses? To demand that communication requires human bodies in a physical, face to face meeting is to declare that all remote means of communications are false and useless. Why use the telephone, or radio, or the written word, if they are not means of communication?

The mechanistic view also denies that intelligence is completely a part of a synergistic relationship. It may be part of a synergistic whole, such as in the case of the human being, but it also possesses an independent existence apart from the synergistic whole, and retains all of its essential characteristics. This claim permits mechanists to divorce intelligence from the human body. It also permits refutation of the missing ingredient objection by claiming that any "missing ingredient" is not part of this component labelled "intelligence."

In terms of science fiction, the literature is full of references to the cyborg. The cyborg is a creature that

begins as a human being, but through whatever circumstances becomes both man and machine. In some instances all that is left of the human body is the brain. Yet the genre treats the cyborg as human, and intelligent, despite whatever body he possesses. In some novels, like Man Plus, by Poul Anderson, the form of the cyborg is extremely different from the human body, and yet the cyborg is treated both as intelligent and as a human being. The extreme case is illustrated by the old horror movies of the Forties and Fifties, and in the recent movie, The Man with Two Brains, where the brain can communicate, completely separated from the body. That it is both conceivable and believable that intelligence could be exhibited by beings of forms other than human, as illustrated by these examples, demonstrates the intuitive beliefs of a number of rational people.

The director of Carnegie-Mellon University's Mobile Robot Laboratory, Hans Moravec, has written a book entitled Mind Children: The Future of Robot and Human Intelligence. In this work, which lies somewhere between science fiction and futurist prediction, Moravec proposes that AI is the means to the next evolutionary step for humanity. Not only will we build machine intelligences, we will merge with them and become, in time, men-machines.

The test for intelligence becomes dependent upon the

philosophical position of the researcher. A synergist would state that intelligence must come in a human package, that the human form and intelligence cannot be reasonably separated without losing something essential. A mechanist would counter that intelligence may be part of the definition of humanity, but that humanity is not part of the definition of intelligence. Neither denies the possibility of machine intelligence. The synergist states that the machine must be human, while the mechanist declares the outer form to be unimportant, as long as intelligence is present.

Machine intelligence need not be equal or equivalent to human intelligence in nature or function. Considering "artificial" in its second usage, as meaning creation, at the very least, it is possible that humans could create intelligence. This man-made intelligence would have the same functionality and effects as some aspects of human intelligence. But it would differ in its nature. The internal processes, the possible subjective states of machine intelligence (barring possibly "organic" computers), will probably be radically different from how humans do what they do when they think. But this does not make that thinking machine any less intelligent. Computers do not and

probably never will possess human intelligence. How could they, when they are not human? Computers are machines, and one day they will possess machine intelligence.

CHAPTER 7
CONCLUSION

In examining artificial intelligence, we have attempted to find some answer to the philosophical and scientific question, "Can a computer possess intelligence?". Taking care to avoid vagueness and ambiguity, definitions for "intelligence," "sentience," "thought," "thinking" and "artificial" were examined to provide background into the issues of AI and to provide a foundation for reasonable discussion and exploration. This examination indicated a general split of AI researchers into two camps based on the type of definition they used for "intelligence" and for "artificial." The first group is, generally although not totally, optimistic about the future of AI, and use broad definitions that are based on a quantitative approach. The second group is using limited definitions for terms based on a qualitative approach. This second group is usually pessimistic about the possibility of a "true" machine intelligence.

The Turing Test holds a special place in the history of AI, and is still important to research today. Although its validity is disputed by opponents of AI, and some scientists question its methodology, it is held as a

standard for an objective test for machine intelligence. Weizenbaum's experimental program ELIZA was also pivotal in the development of AI, both in the realm of practical application to natural language interfaces and in the realm of theoretical dispute over definitions of intelligence, and the ethics of AI research.

There are several objections that voice negative answers to the question of computer intelligence. The Theological Objection, which claims the presence of a divine element within intelligence, goes unanswered and unapproached by the majority of AI researchers, for now. The Missing Ingredient Objection, which claims the presence of essential elements of intelligence that machines cannot possess, is answered either by the counterclaim of ad hoc argumentation, or by appeal to hardware or software improvement in the future. The Consciousness Objection is based on the claim of intentionality and self-awareness, qualities that machine have never exhibited. The reply is again that hardware and software development will eventually result in sentient computers.

The Moral Objection declares AI research to be fruitless and dangerous. It is a form of an older question from the philosophy of science. The question is whether scientific study of a theoretical area is morally justified when a major portion of its potential and actual applications can be used in immoral ways. Is the AI

researcher at MIT morally responsible if his theoretical work helps in the design of a weapon system that kills without a human to exercise human judgment and restraint? The researchers in AI answer as most scientists have throughout the ages. The pursuit of knowledge for its own sake is good in the moral sense. The search for theoretic knowledge must not be limited by potential abuse of applied theories.

Examining the positions of the two camps, we attempted to establish some criterion to determine success in creating machine intelligence. The first group, utilizing broad definitions, stated that computer will be capable of possessing intelligence soon. The second group with its limited definitions stated that machine intelligence will never be achieved. Areas of current research provide support for the optimists, with exciting developments in parallel processing, neural networks, and programming algorithms like heuristic programming in chess playing computers.

Viewing "artificial" to carry the meaning of "created by man" and "intelligence" to possess both qualitative and quantitative elements, a possible criterion that avoids some of the problems and objections to either of the major camps was proposed. The belief is that computers will never possess human intelligence, but that they will possess machine intelligence. This intelligence will possess the

qualities and characteristics of machines, distinguishing it clearly from human intelligence. This criterion is not sufficient by itself, however.

The philosophical issue of the humanity of intelligence presented two views on what form intelligence can be found in. The synergistic view of humanity and intelligence perceives intelligence as an integral part of humanity, and that the two cannot be separated. The mechanistic view of humanity and intelligence sees intelligence as a component of humanity that has existence that is independent of what body or form it exists within.

Adding the humanity factor to the first criterion proposed, the synergist in the broad-meaning camp would declare machine intelligence possible, but only if the machine became human. The mechanist in the broad-meaning camp would state that machine intelligence will be **machine** intelligence. The synergist in the limited-meaning camp would deny the possibility of "intelligent" computers at all, while the limited-meaning mechanists grant the possibility but not the probability of any achievement of AI.

Several works of a speculative nature have been written recently. Hans Moravec's book, with his proposition that AI is the means to the next evolutionary step for humanity, approaches this as factual and as a

positive step. Weizenbaum also sees this as a possibility for humanity. He fears that humanity will create its own replacement. He sees this as his worst nightmare coming true.

For us, these visions should serve as the catalyst for the search for the answers for the new questions of AI. Perhaps we should not only be asking "Can we?", but rather we should also begin asking "Should we?".

APPENDIX A

ELIZA PROGRAM LISTING

```

1 LPRINT TAB(26);"ELIZA"
2 LPRINT TAB(20);"CREATIVE COMPUTING"
3 LPRINT TAB(18);"MORRISTOWN, NEW JERSEY"
4 LPRINT:LPRINT:LPRINT
5 FOR X = 1 TO 5000: NEXT X
6 CLS
80 REM      INITIALIZATION
100 DIM S(36),R(36),N(36)
110 N1=36:N2=14:N3=112
120 FOR X=1 TO N1+N2+N3:READ Z$:NEXT X:REM SAME AS RESTORE
130 FOR X=1 TO N1
140 READ S(X),L:R(X)=S(X):N(X)=S(X)+L-1
150 NEXT X
160 LPRINT "HI! I'm Eliza. What's your problem?"
165 PRINT "HI! I'M ELIZA. WHAT'S YOUR PROBLEM?"
170 REM
180 REM      USER INPUT SECTION
190 REM
200 INPUT I$
201 I$ = " "+I$+" "
205 LPRINT I$
210 REM      GET RID OF APOSTROPHES
220 FOR L=1 TO LEN(I$)
230 IF MID$(I$,L,1)="'" THEN I$ =
LEFT$(I$,L-1)+RIGHT$(I$,LEN(I$)-L):GOTO 230
240 IF L+4 < LEN(I$) THEN IF MID$(I$,L,4)="SHUT"
THEN LPRINT "SHUT UP...":END
250 NEXT L
251 X$ = "PLEASE DON'T REPEAT YOURSELF!"
255 IF I$ = P$ THEN LPRINT X$:PRINT X$:GOTO 170
260 REM
270 REM      FIND KEYWORD IN I$
280 REM
290 RESTORE
295 S=0
300 FOR K=1 TO N1
310 READ K$
315 IF S>0 THEN 360
320 FOR L =1 TO LEN(I$)-LEN(K$)+1
340 IF MID$(I$,L,LEN(K$))=K$ THEN S=K:T=L:F$=K$
350 NEXT L

```

```

360 NEXT K
365 IF S>0 THEN K=S:L=T:GOTO 390
370 K=36:GOTO 570:REM WE DIDN'T FIND ANY KEYWORDS
380 REM
390 REM TAKE RIGHT PART OF STRING AND CONJUGATE IT
400 REM USING THE LIST OF STRINGS TO BE SWAPPED
410 REM
420 RESTORE:FOR X = 1 TO N1:READ Z$:NEXT X:
REM SKIP OVER KEYWORDS
430 C$=" "+RIGHT$(I$,LEN(I$)-LEN(F$)-L+1)+" "
440 FOR X = 1 TO N2/2
450 READ S$,R$
460 FOR L = 1 TO LEN(C$)
470 IF L+LEN(S$)>LEN(C$) THEN 510
480 IF MID$(C$,L,LEN(S$))<>S$ THEN 510
490 C$=LEFT$(C$,L-1)+R$+RIGHT$(C$,LEN(C$)-L-LEN(S$)+1)
495 L=L+LEN(R$)
500 GOTO 540
510 IF L+LEN(R$)>LEN(C$) THEN 540
520 IF MID$(C$,L,LEN(R$))<>R$ THEN 540
530 C$=LEFT$(C$,L-1)+S$+RIGHT$(C$,LEN(C$)-L-LEN(R$)+1)
535 L=L+LEN(S$)
540 NEXT L
550 NEXT X
555 IF MID$(C$,2,1)=" " THEN C$=RIGHT$(C$,LEN(C$)-1):
REM ONLY 1 SPACE
556 FOR L=1 TO LEN(C$)
557 IF MID$(C$,L,1)="!" THEN
C$=LEFT$(C$,L-1)+RIGHT$(C$,LEN(C$)-L): GOTO 557
558 NEXT L
560 REM
570 REM NOW USING THE KEYWORD NUMBER (K) GET REPLY
580 REM
590 RESTORE:FOR X=1 TO N1+N2:READ Z$:NEXT X
600 FOR X = 1 TO R(K):READ F$:NEXT X:REM READ RIGHT REPLY
610 R(K)=R(K)+1:IF R(K) > N(K) THEN R(K) = S(K)
620 IF RIGHT$(F$,1)<>"*" THEN LPRINT F$:
PRINT F$:P$=I$:GOTO 170
630 LPRINT LEFT$(F$,LEN(F$)-1);C$
635 PRINT LEFT$(F$,LEN(F$)-1);C$
640 P$=I$:GOTO 170
1000 REM
1010 REM PROGRAM DATA FOLLOWS
1020 REM
1030 REM KEYWORDS
1040 REM
1050 DATA "CAN YOU","CAN I","YOU ARE","YOU'RE","I DONT",
"I FEEL"
1060 DATA "WHY DONT YOU","WHY CANT I","ARE YOU","I CANT",
"I AM","IM "
1070 DATA "YOU ","I WANT","WHAT","HOW","WHO","WHERE",

```

"WHEN", "WHY"
1080 DATA "NAME", "CAUSE", "SORRY", "DREAM", "HELLO", "HI",
"MAYBE"
1090 DATA " NO", "YOUR", "ALWAYS", "THINK", "ALIKE", "YES",
"FRIEND"
1100 DATA "COMPUTER", "NOKEYFOUND"
1200 REM
1210 REM STRING DATA FOR CONJUGATIONS
1220 REM
1230 DATA " ARE ", " AM ", "WERE ", "WAS ", " YOU ", " I ",
"YOUR ", "MY "
1235 DATA " IVE ", " YOUVE ", " IM ", " YOURE "
1240 DATA " ME ", " !YOU "
1300 REM
1310 REM REPLIES
1320 REM
1330 DATA "DON'T YOU BELIEVE I CAN*"
1340 DATA "PERHAPS YOU WOULD LIKE TO BE ABLE TO*"
1350 DATA "YOU WANT ME TO BE ABLE TO*"
1360 DATA "PERHAPS YOU DON'T WANT TO*"
1365 DATA "DO YOU WANT TO BE ABLE TO*"
1370 DATA "WHAT MAKES YOU THINK I AM*"
1380 DATA "DOES IT PLEASE YOU TO BELIEVE I AM*"
1390 DATA "PERHAPS YOU WOULD LIKE TO BE*"
1400 DATA "DO YOU SOMETIMES WISH YOU WERE*"
1410 DATA "DON'T YOU REALLY*"
1420 DATA "WHY DON'T YOU*"
1430 DATA "DO YOU WISH TO BE ABLE TO*"
1440 DATA "DOES THAT TROUBLE YOU?"
1450 DATA "TELL ME MORE ABOUT SUCH FEELINGS."
1460 DATA "DO YOU OFTEN FEEL*"
1470 DATA "DO YOU ENJOY FEELING*"
1480 DATA "DO YOU REALLY BELIEVE I DON'T*"
1490 DATA "PERHAPS IN GOOD TIME I WILL*"
1500 DATA "DO YOU WANT ME TO*"
1510 DATA "DO YOU THINK I SHOULD BE ABLE TO*"
1520 DATA "WHY CAN'T YOU*"
1530 DATA "WHY ARE YOU INTERESTED IN WHETHER OR NOT I AM*"
1540 DATA "WOULD YOU PREFER IF I WERE NOT*"
1550 DATA "PERHAPS IN YOUR FANTASIES I AM*"
1560 DATA "HOW DO YOU KNOW I CAN'T*"
1570 DATA "HAVE YOU TRIED?"
1580 DATA "PERHAPS YOU CAN NOW*"
1590 DATA "DID YOU COME TO ME BECAUSE YOU ARE*"
1600 DATA "HOW LONG HAVE YOU BEEN*"
1610 DATA "DO YOU BELIEVE IT IS NORMAL TO BE*"
1620 DATA "DO YOU ENJOY BEING*"
1630 DATA "WE WERE DISCUSSING YOU— NOT ME."
1640 DATA "OH, I*"
1650 DATA "YOU'RE NOT REALLY TALKING ABOUT ME, ARE YOU?"
1660 DATA "WHAT WOULD IT MEAN TO YOU IF YOU GOT*"

1670 DATA "WHY DO YOU WANT*"
1680 DATA "SUPPOSE YOU SOON GOT*"
1690 DATA "WHAT IF YOU NEVER GOT*"
1700 DATA "I SOMETIMES ALSO WANT*"
1710 DATA "WHY DO YOU ASK?"
1720 DATA "DOES THAT QUESTION INTEREST YOU?"
1730 DATA "WHAT ANSWER WOULD PLEASE YOU THE MOST?"
1740 DATA "WHAT DO YOU THINK?"
1750 DATA "ARE SUCH QUESTIONS ON YOUR MIND OFTEN?"
1760 DATA "WHAT IS IT THAT YOU REALLY WANT TO KNOW?"
1770 DATA "HAVE YOU ASKED ANYONE ELSE?"
1780 DATA "HAVE YOU ASKED SUCH QUESTIONS BEFORE?"
1790 DATA "WHAT ELSE COMES TO MIND WHEN YOU ASK THAT?"
1800 DATA "NAMES DON'T INTEREST ME."
1810 DATA "I DON'T CARE ABOUT NAMES— PLEASE GO ON."
1820 DATA "IS THAT THE REAL REASON?"
1830 DATA "DON'T ANY OTHER REASONS COME TO MIND?"
1840 DATA "WHAT OTHER REASONS MIGHT THERE BE?"
1850 DATA "DOES THAT REASON EXPLAIN ANYTHING ELSE?"
1860 DATA "PLEASE DON'T APOLOGIZE!"
1870 DATA "APOLOGIES ARE NOT NECESSARY."
1880 DATA "WHAT FEELING DO YOU HAVE WHEN YOU APOLOGIZE?"
1890 DATA "DON'T BE SO DEFENSIVE!"
1900 DATA "WHAT DOES THAT DREAM SUGGEST TO YOU?"
1910 DATA "DO YOU DREAM OFTEN?"
1920 DATA "WHAT PERSONS APPEAR IN YOUR DREAMS?"
1930 DATA "ARE YOU DISTURBED BY YOUR DREAMS?"
1940 DATA "HOW DO YOU DO...PLEASE STATE YOUR PROBLEM."
1950 DATA "YOU DON'T SEEM QUITE CERTAIN."
1960 DATA "WHY THE UNCERTAIN TONE?"
1970 DATA "CAN'T YOU BE MORE POSITIVE?"
1980 DATA "YOU AREN'T SURE?"
1990 DATA "DON'T YOU KNOW?"
2000 DATA "ARE YOU SAYING NO JUST TO BE NEGATIVE?"
2010 DATA "YOU ARE BEING A BIT NEGATIVE."
2020 DATA "WHY NOT?"
2030 DATA "ARE YOU SURE?"
2040 DATA "WHY NO?"
2050 DATA "WHY ARE YOU CONCERNED ABOUT MY*"
2060 DATA "WHAT ABOUT YOUR OWN*"
2070 DATA "CAN YOU THINK OF A SPECIFIC EXAMPLE?"
2080 DATA "WHEN?"
2090 DATA "WHAT ARE YOU THINKING OF?"
2100 DATA "REALLY, ALWAYS?"
2110 DATA "DO YOU REALLY THINK SO?"
2120 DATA "BUT YOU ARE SURE YOU*"
2130 DATA " DO YOU DOUBT YOU*"
2140 DATA "IN WHAT WAY?"
2150 DATA "WHAT RESEMBLANCE DO YOU SEE?"
2160 DATA "WHAT DOES THE SIMILARITY SUGGEST TO YOU?"
2170 DATA "WHAT OTHER CONNECTIONS DO YOU SEE?"

2180 DATA "COULD THERE REALLY BE SOME CONNECTION?"
 2190 DATA "HOW?"
 2200 DATA "YOU SEEM QUITE POSITIVE."
 2210 DATA "ARE YOU SURE?"
 2220 DATA "I SEE."
 2230 DATA "I UNDERSTAND."
 2240 DATA "WHY DO YOU BRING UP THE TOPIC OF FRIENDS?"
 2250 DATA "DO YOUR FRIENDS WORRY YOU?"
 2260 DATA "DO YOUR FRIENDS PICK ON YOU?"
 2270 DATA "ARE YOU SURE YOU HAVE ANY FRIENDS?"
 2280 DATA "DO YOU IMPOSE ON YOUR FRIENDS?"
 2290 DATA "PERHAPS YOUR LOVE FOR YOUR FRIENDS WORRIES YOU."
 2300 DATA "DO COMPUTERS WORRY YOU?"
 2310 DATA "ARE YOU TALKING ABOUT ME IN PARTICULAR?"
 2320 DATA "ARE YOU FRIGHTENED BY MACHINES?"
 2330 DATA "WHY DO YOU MENTION COMPUTERS?"
 2340 DATA "WHAT DO YOU THINK MACHINES HAVE TO DO WITH YOUR
 PROBLEM?"
 2350 DATA "DON'T YOU THINK THAT COMPUTERS CAN HELP PEOPLE?"
 2360 DATA "WHAT IS IT ABOUT MACHINES THAT WORRIES YOU?"
 2370 DATA "SAY, DO YOU HAVE ANY PSYCHOLOGICAL PROBLEMS?"
 2380 DATA "WHAT DOES THAT SUGGEST TO YOU?"
 2390 DATA "I SEE."
 2400 DATA "I'M NOT SURE I UNDERSTAND YOU FULLY."
 2410 DATA "COME COME, ELUCIDATE YOUR THOUGHTS!"
 2420 DATA "CAN YOU ELABORATE ON THAT?"
 2430 DATA "THAT IS QUITE INTERESTING!"
 2500 REM
 2510 REM DATA FOR FINDING RIGHT REPLIES
 2520 REM
 2530 DATA 1,3,4,2,6,4,6,4,10,4,14,3,17,3,20,2,22,3,25,3
 2540 DATA 28,4,28,4,32,3,35,5,40,9,40,9,40,9,40,9,40,9
 2550 DATA 49,2,51,4,55,4,59,4,63,1,63,1,64,5,69,5,74,2,76,4
 2560 DATA 80,3,83,7,90,3,93,6,99,7,106,6

APPENDIX B

ELIZA RUN-TIME PRINTOUT

ELIZA
CREATIVE COMPUTING
MORRISTOWN, NEW JERSEY

ELIZA HI! I'm Eliza. What's your problem?
Interviewer I'M AFRAID OF THE DARK
ELIZA DID YOU COME TO ME BECAUSE YOU ARE AFRAID OF
. THE DARK
Interviewer YES
ELIZA ARE YOU SURE?
Interviewer I'M VERY AFRAID OF THE DARK
. HOW LONG HAVE YOU BEEN VERY AFRAID OF THE
. DARK
. SINCE I HAD THE BAD DREAM
. WHAT DOES THAT DREAM SUGGEST TO YOU?
. THAT I HAVEN'T GONE TO CHURCH ENOUGH
SAY, DO YOU HAVE ANY PSYCHOLOGICAL PROBLEMS?
I HATE GOING TO CHURCH
WHAT DOES THAT SUGGEST TO YOU?
NOBODY LOVES ME
ARE YOU SAYING NO JUST TO BE NEGATIVE?
I WANT TO BE LOVED
WHAT WOULD IT MEAN TO YOU IF YOU GOT TO BE
LOVED
I WOULD BE HAPPY
I SEE.
I WANT TO BE HAPPY
WHY DO YOU WANT TO BE HAPPY
I DONT WANT TO BE IN PAIN
DON'T YOU REALLY WANT TO BE IN PAIN
IT HURTS WHEN I'M IN PAIN
DO YOU BELIEVE IT IS NORMAL TO BE IN PAIN
IT IS NOT NORMAL TO BE IN PAIN
YOU ARE BEING A BIT NEGATIVE.
HAPPINESS IS NORMAL
WHY NOT?
I AM NOT HAPPY
DID YOU COME TO ME BECAUSE YOU ARE NOT HAPPY
YES
I SEE.
I WANT TO BE HAPPY

SUPPOSE YOU SOON GOT TO BE HAPPY
 MY LIFE WOULD BE GOOD
 I'M NOT SURE I UNDERSTAND YOU FULLY.
 I WANT A GOOD LIFE
 WHAT IF YOU NEVER GOT A GOOD LIFE
 I WOULD BE UNHAPPY
 COME COME, ELUCIDATE YOUR THOUGHTS!
 WHAT IS GOOD
 WHY DO YOU ASK?
 I WANT TO LIVE A GOOD LIFE
 I SOMETIMES ALSO WANT TO LIVE A GOOD LIFE
 CAN A COMPUTER LIVE
 DO COMPUTERS WORRY YOU?
 I'M SCARED OF COMPUTERS
 DO YOU ENJOY BEING SCARED OF COMPUTERS
 I ENJOY WORKING WITH COMPUTERS
 ARE YOU TALKING ABOUT ME IN PARTICULAR?
 YES
 I UNDERSTAND.
 DO YOU UNDERSTAND
 WE WERE DISCUSSING YOU— NOT ME.
 I WANT YOU TO UNDERSTAND ME
 OH, I TO UNDERSTAND YOU
 YES
 ARE YOU SURE?
 I WANT UNDERSTANDING
 WHAT WOULD IT MEAN TO YOU IF YOU GOT
 UNDERSTANDING
 I WOULD BE HAPPY
 CAN YOU ELABORATE ON THAT?
 I WANT TO BE HAPPY
 WHY DO YOU WANT TO BE HAPPY
 I WANT TO LIVE A GOOD LIFE
 SUPPOSE YOU SOON GOT TO LIVE A GOOD LIFE
 I WOULD BE LIVING MORALLY
 SAY, DO YOU HAVE ANY PSYCHOLOGICAL PROBLEMS?
 SHUT UP
 SHUT UP...

NOTES: Any grammatical or punctuation errors present in
 this transcript are a result of the actions of the program.
 The interviewer is required by the program to drop
 apostrophes from contractions.

SELECT BIBLIOGRAPHY

- Asimov, Issac. The Bicentennial Man and Other Stories. Garden City, NY: Doubleday & Co., 1976.
- Bulfinch, Thomas. Bulfinch's Mythology: The Age of Fable. Garden City, NY: Doubleday & Co., 1968.
- Clarke, Arthur C. "Are You Thinking Machines?" In Industrial Research (March 1969): 52-62.
- Dictionary of Philosophy and Psychology, ed. James Mark Baldwin. New York: Macmillan Co., 1918.
- Descartes, Rene. Discourse on the Method and the Meditations. Trans. F.E. Sutcliffe. New York: Viking Penguin Inc., 1968.
- Dorf, Richard C. Computers and Man. San Francisco: Boyd & Fraser Publishing Co., 1974.
- Evans, Christopher R. The Micro Millennium. New York: The Viking Press, 1979.
- Gasse, Jean-Louis. "This Idiot Savant: The Proper Study of Machine," ComputerLand Magazine 3, no. 2 (January/February 1988): 34-43.
- Haugeland, John. Artificial Intelligence: The Very Idea. Boston: MIT Press, 1985.
- Hunt, Morton. "What the Human Mind Can Do That the Computer Can't", In Philosophy and Contemporary Issues, ed. Milton Goldinger and John R. Burr. New York: Macmillan Publishing Co., 1984.
- Moravec, Hans. Mind Children. Cambridge, MA: Harvard University Press, 1988.
- North, Steve. ELIZA, in More Basic Computer Games, ed. David H. Ahl. New York: Workman Publishing, 1979.
- Reed, Fred. "Overlooking the Marvelous," Federal Computer Week, 3 October 1988, 12.

Searle, John R. "Minds, Brains and Programs", In Mind Design, ed. John Haugeland. Montgomery, VT: Bradford Books, Publishers, 1981.

Simpson, Robert L. "DARPA Pushes Frontiers of Machine Intelligence," interview by Neil Munro, Government Computer News, 22 January 1988, pp. 10-11.

Turing, Alan M. "Can a Machine Think?" In Mind 59, No. 236 (1950) 2099-2100.

Webster's New International Dictionary of the English Language, rev. ed. (1950).

Weizenbaum, Joseph. Computer Power and Human Reason: From Judgment to Calculation. New York: W. H. Freeman and Co., 1976.