Spring 1984

Mind over matter : considerations of understanding and artificial intelligence

John C. Brice
University of Richmond

Philosophy Honors Thesis;

MIND OVER MATTER: CONSIDERATIONS OF UNDERSTANDING AND ARTIFICIAL INTELLIGENCE

By
John C. Brice

Philosophy Department
University of Richmond

April 12, 1984

Contents

PART I: INTRODUCTION

There are troubles encountered in approaching an understanding of what intelligence is in stipulative terms without attempting an empirical investigation of the phenomenon. Such processes will necessarily entail an investigation of the behavior of those entities which we agree exhibit such properties which are sufficiently indicative of an intelligent presence (intelligence being not something we anticipate sensing physically but rather a label wich we apply to a system which exhibits a certain family of behavior patterns which we consider to be indicative of a rationally designed approach to its world). We must be limited to observational statements in identifying those criteria which are indicative of intelligence. The term "intligence" will thus refer to a style of behavior rather than something hidden away inside our heads.

If we ask, then, in an age in which robots and computers have begun to drastically alter the face of our work environment, what it is about computers that makes them so like us, or, how smart can these machines get, it seems that an equally relevent question to ask is directed at what it is about humans that computers (so far) cannot imitate. What are the limits of Artificial Intelligence? Are these limits the limits of our own brains in designing intelligent systems or is there something special about human intelligence which discrete state devices are incapable of duplicating?

So far machines have proven quite capable of duplicating the mechanical motions of the human body, thus the prevalence of robots in factories. Computers can solve mathematical problems (by plugging in formulas) and can type manuscrips. Basically, computers can do just about anything one cares to tell them to do (provided one knows how). And that's the difference. Anything a computer does, it does because it's explicitly told, in some form of programming language, to do so. What we herald as the great thing about humans is that nobody has to (in priciple) tell them to do anything and yet they still work. And (partly for the same reasons) they are unpredictable. A computer's actions, if one cares to take the time to analyse its program, are totally predictable. All a computer's actions can be traced to specific orders within the computer's programming. The computer has no say in what it does -- it doesn't even know what (or that) it does. A computer's actions are merely responses to electromagnetic stimuli which fall precisely within the laws of nature.

But can we not say the same thing, on a molecular level, of human beings? After all, if our bodies and brains were not constructed in accordence with the laws of nature, we wouldn't (couldn't) exist.

Being a self, alive in the world, man has a basic, call it wired-in, motivation to maintain this status of his self as alive in the world. Is the difference, then, between men and computers that the computer has no sense of self to which the things that it encounters in its world can be said to have any relevence or meaning? If this is the case, then is the self something (a process or a set of processes) which can be written into a computer program?

There are three classes or words that, when applied to intelligent systems have everything to do with human beings and, ordinarily, nothing to do with computers. These are "self," which is the agent who can be said to have an "understanding" of things in the world which when contrasted to the self by the self are said to have "meaning." These concepts are clearly interdependent and therefore any discussion of one imples the existence and situational inclusion of the other two.

On this level, we see that part of understanding is something which happens when the self, or subject, becomes aware of the object under scrutiny as being distinct from the self. But understanding is more than mere recognition. Perceiving simply the otherness of an object does not inform one in any way which would indicate an appropriate attude to adopt in approaching some alien object. The quality of this otherness that is perceived forms the meaning which the object holds for the subject. It is therefore the ability to detect meaning in or assign meaning to the elements of one's environment (or a topic of discourse either spoken or written) which constitues an essential part of understanding which computers so far do not have. The extent to which this process can be duplicated in a computer program should play a large part in determining the extent to which machines may be said to be intelligent.

To date, virtually all of the attempts at simulating understanding in an artificial system have concentrated on written texts. There are several reasons for this. First, when we speak of understanding, we generally have in mind the understanding of communication of an idea expressed in natural language (as opposed to computer language). It is commonly accepted, however, that the technical problem of understanding spoken language have little bearing on the basic problems involved with getting a computer to understand the meaning conveyed by language. By concentrating on written text, full attention can be directed toward this more basic problem.

It is commonly accepted that the task of understanding a written text becomes the standard analog of the major task which stands before those working in Artificial Intelligence -- the establishment of synthetic understanding in general. The problem here is that an artificial intellignece should be versatile enough (as much as human intelligence) to understand the wide variety of texts which are possible in natural language. The difference between natural language and computer language is that the former is capable of expressing meaning (any given meaning) in an almost unlimited number of ways, often without even explicitly stating the precise meaning, while the latter is confined to very limited expressions of meaning, and is composed of elements in which all meaning must be precise and explicit. There is exactly no ambiguity in comptuer language. This can be very useful -- no questions should ever arise concerning what is meant by a particular text, however this also establishes a very rigid structure for expression and some ideas (for example, religious ideas) which depend on metaphor become impossible to express adequately in such a rigid format. What this means is, that if an artificial, potentially intelligent system is to be bound by a fully explicit linguistic system, much of what we take to be meaningful in our everyday discourse must be rejected as meaningless by its computer language. This is certainly unsatisfactory.

If it is therefore accepted that one mark of an artificially intelligent system will be the ability to interact with humans (or other, perhaps artificial, intelligences) through the use of natural language, then it is not enough that the computer utilize a language which employs, say, English words (any human language of discourse, i.e., natural language, would do here). The system must also be capable of both understanding and generating text in an unrestricted realm (with regard to subject matter) defined only by the grammar of the natural language itself.

One might now inquire as to what are the relevent differences in this context between natural language (the languages of man) and artificial languages (languages of logic and computers). One difference may be that as humans, we are able to associate meanings and build networks of connotations for words in ways which so far have nt been simulated by machines. This means that as humans, we do not invest all of the meaning of a word in the symbol itself. Part of the meaning evidently arises out of its conjunction with other words. For example, John may tell Joe that his (Joe's) cat is on his (John's) mat. If Joe understands this utterance, and knows also that he is not on good term with John, then he may conclude that John may be advising him to get his cat off the mat. Furthermore, the representation of possible and actual events (the cat being on or off the mat) is possibly and probably linguistic. Perhaps Joe is even now looking at the cat on the mat. He can see part of the meaning (the obvious part) of John's utterance -- the cat, he sees, certainly is on the mat. He can also visualize the same mat minus the cat and this possible situation's probable implications. Even if the two were speaking over the telephone, it is likely that Joe would be visualizing the meaning of John's utterence. It is not implicit that any computer would. Because he is capable of such visualizations he has access to information about the situation and the relationships which define the situation to which the computer does not.

Even if Joe has never seen his cat (John's cat has just had kittens, one of which is, by prior contractual obligation now Joe's and is, incidentally, making a mess on John's mat) he is capable of visualizing a cat (any cat) on some mat. The representation of meaning for him is not restricted to a linguistic one, especially if he has had prior concrete, physical experience with what is meant by all the words involved. It is not so important here that meaning be expressed specifically or exclusively in visual or any other kinds of symbols, but only that it not be bound to simply linguistically codified terms. Rather, it seems that the function of language for an intelligent system should be (and is, for humans) simply an expression of stored meaning -- independent of however such meaning is stored. This same meaning has been expressed in different words by Roger Schank:

> No matter how events described by sentences are stored in memory, they cannot be stored in terms of whatever words may have been used in those sentences. There are a great many ways to describe an event, but the event remains the same regardless of the description. Similarly, memory's encoding of that event is the same regardless of the description. Thus, any meaning representation used by mind must be non-lingual, involving pure meaning elements only. No matter what lexical form is chosen to express a particular thought, there will be only one conceptual meaning representation stored in memory. This explains why people who speak two or more languages fluently sometimes forget which language they are using. The concepts they wish to express are stored in neither language in their memories. (1)

There is more, however, to intelligence and understanding than mere memory. If a system is to be said to be intelligent (hereafter implying the possession of the ability to understand natural language) then we would expect the system to be capable of demonstrating its understanding through its ability to engage in discourse within natural language. This means that an intelligent machine must, for our purposes, be able to understand written text and respond to and perhaps formulate questions about the text in natural language. This task should be far less demanding than the achievement of the corresponding task through spoken language since often much of the meaning of discourse is contained at least as much in the tone of voice and perhaps in body language. The limitation to written language forces such expression to be either reduced to words of forfieted.

If meaning is to be stored non-linguistically then the system must be capable not only of translating linguistic representations of meaning into non-linguistic representations (reasoning) and ultimately constructing linguistic representations for the purpose of communication with other minds.

We would also expect the system to maintain some kind of goal directing mechanism. The system must maintain a model of the discourse and the goals toward which its reasoning is to be directed. To put it crudely, the system ought to have some idea of what its talking about. It should know roughly where the discourse is headed. Thus, not only should it be capable of reasoning it should know when to reason and about what that reasoning should be. Toward this end, it should have a sense for what is relevent to the topic of conversation, what needs to be explicitly stated, and what is taken for granted. It should know or be capable of anticipating the mind of its partner in discourse. This does not mean that the computer should actually know what the other person is thinking, only that is should be maintaining a model of that other person. It ought to know roughly what the partner in discourse is thinking about. It should make predictions or gursses not only as to what the other intelligence is driving at but also about what it doesn't know or ought to know.

For example, John may instruct Joe to open the door. However, Joe may, justifiably, assume that John's instructions are based on the hypothesis that someone is knocking on the door. If Joe has been rapping on the table, he may suspect that John has mistaken this rapping for a knock at the door; thus his request. Instead of answering the door, Joe would likely respond, "No, that was just me rapping on the table."

An intelligent participant in a dialogue maintains models both of the topic of conversation and of the mental sets of those involved in the conversation. These models are updated as the dialogue or conversation progresses.

Perhaps the most significant implication of this maintenance of models is that the subject must maintain a model of self; self awareness. This self must be distinguished from those of the other participants in the conversation. It is not entirely certain that all humans recognize other participants in conversation as selves equivalent to their own selves however most humans do recognize others at least as distinct from their selves.

Insofar as these other entities, as participants in conversation exhibit qualities which distinquish them from inanimate objects (which are also recognized as other than the self), among these qualities being the very ability to participate in conversation or other forms of communication, then the intelligent system must be expected to maintain models of such entities. By maintainance I mean not only the establishment of a representational model, but also its updating as such updating becomes possible.

This ability implies that the intelligent system be capable of conducting various forms of reasoning about these other mentalities (as well as the topic of discourse). When John instructed Joe to answer the door, Joe was able to reason about the situation, and about utterances about their two perceptions of reality. Through such reasoning, Joe was able to fill certain blanks about John's goals and perceptions which affected his response to John's request. John might still insist on Joe's answering the door if he is aware of Joe's nervous habit but also believes he heard a separate and distinct knock on the door. If John so insisted, Joe would need to update his model either of John's perceptions, their relation to reality, reality itself, or all of these in relation to his own perceptions.

In order for an artificial intelligence to have a "complete" understanding (by "complete" I only mean "comparible to that of humans") of discourse in the real world, it would need a source of experience similar to that of those who exist in the real world. It needs the sensory input available through a body as a vehicle of experience in the world. Much of understanding has to do with physical relationships or qualities which sometimes may not be explicit or even implicit through language alone. Understanding the world without a body through which to experience it might be better appreciated if we were to consider trying to explain chairness to a disembodied brain which never had a lap to benefit from chairness nor had ever experienced anyone else deriving such benefit. It might be able to manipulate the concept but would never know if it was sitting, in some container, on a chair.

It is the body which brings the mind in contact with the external world and so any attempt to understand this world must presuppose the existence of the intermediary agency of a body. We may never speak of a computer capable of understanding event in the world which was not itself linked to the world through the agency of a body. In what follows, I shall present a review of some of the more interesting results of the current research in Artificial Intelligence, to be followed by a discussion of Alan Turing's work along with some elaboration of my own thoughts.

PART II: REVIEW OF THE CURRENT RESEARCH

ELIZA AND THE DOCTOR

Joseph Weizenbaum developed a program called DOCTOR which plays the game of Psychiatrist with its user. The user initiates the program by typing in a statement. The system then scans the statement for key words that it recognizes and formulates a responding question based on that key word. If the input contains no recognizable key words, a "content-free remark" is returned. Key words are ranked so that if more than one key word is found in the input, it utilizes the one with the highest rank. Also, when the scanner comes to a comma or a period, it stops scanning if it has found a ranked key word. If it finds no key words, the system picks up a key word from some previous input.

The system tries to identify "some minimal context within which the chosen key word appears; for example, if the key word is 'you', it is followed by the word 'are' (in which case an assertion is probably being made)." (2)

Most of what the system does is just text manipulation. Although the program in operation may appear to be intelligent, this is only an illusion. No attempt is made here of understanding. Beyond the few key words which the program recognizes, all unfamiliar words are dealt with using certain transformation formulas based on the location of known words in the input and the numbers of unknown words which surround them. For example, if the input consists of the sentence "You've never liked me." and the only familiar words are 'You' and 'me' then the computer breaks the sentence up into corresponding chunks; "1) You've 2) never liked 3) me." Although the computer has no idea what phrase (2) means, it doesn't matter. The computer merely changes the "You" to an "I" and the "me" to a "you" and applies the sentence to a template; "1) What makes you think 2) I've 3) never liked 4) you?" The system selects decomposition rules according to whatever key words it finds.

The success of this program depends upon the normal assuptions and expectations that the user brings with him to any conversation. The program capitalizes on the fact that most people interpret the ability to manipulate language as understanding of that language. However, as Weizenbaum points out, as soon as the user breaks from the role of Psychiatrist's Patient, the illusion of understanding breaks down. Although it seems possible that a good enough pretender may one day force us to recognize it as intelligent, I would insist that the quantities of data which would need to be preprogrammed for such a device would make any attempt using the current techniques inherently less than adequate. As long as information must be explicitly encoded for the machine there will remain relatively creative aspects of human intelligence which will ot have been emusated.

Weizenbaum contrasts this program with a second program in the ELIZA series which, within an apparently narrower range of subject matter, goes much farther towards true understanding. The program has a·very limited vocabulary (about a dozen or so words in addition to the usual mathematical manipulators). However, through this vocabulary it is able to learn new words when they are defined in terms of the ones it already knows. Whereas the DOCTOR program tried to conceal its ignorance, the newer program asks questions when it encounters a crucial word it doesn't know.

Although the system seems mainly desinged to understand limited natural language input (in English) and perfom basic mathematical calculations, its capacity for learning enabled Weizenbaum to teach it the German equivalents of the words it knew so that it was soon answering questions (in English) that were asked in German.

With respect to the Doctor Program, the essential point is that there is no semblance of semantic analysis going on in this program. The words which the system recognizes are not attached to any data stores or meaning matrices, but rather to a prefabricated response. Any intelligence which the system may appear to display is actually a reflection of the cleverness of its author. This structural shortcoming becomes blatent when the user turns the situation around and begins asking the DOCTOR questions. Since it doesn't have any clue with respect to the meaning of the input, there's really no way it can even recognize that a question has been asked, let alone formulate an intelligent response.

The success of this program is like the success of a master of slight of hand. It is a success which relies on the gullibility of the oberver. It doesn't matter that it never really even tries to do the things that it pretends to do, just as long as it's fairly good at pretending.

The ELIZA program, although its potential range of discourse available is far less than that of DOCTOR, seems to come quite a ways closer to understanding its world. The greatest value of this program is that it demonstrates the ability of machines to learn. This is certainly an essential element of any intelligent system. We must grant that this learning occurs at best on avery formal level and that the computer only "understands" the German which it learned as much as it ever understood the English that it uses. Since the only things this program can really talk about are mathematical operations (either in English or in German) the world which it needs to understand is a very stripped-down, logical world the truths of which are a priori. In order to establish a broader base of understanding and learning, we would need to establish some kind of mechanism by which the computer might establish isomorphisms between the things which it "learns" and the world which it perceives. Thus when it learns that a chair is something people sit in, it ought to be able to associate this knowledge with the perception of at least one kind of chair in its perceived world.

As the ELIZA program demonstrates, the ability to learn and create associations may provide an alternative to explicit programming.

DEJONG'S FRUMP SYSTEM

The FRUMP (Fast Reading Understanding and Memory Program) system takes as its input the text which comes in over the United Press International news wire. These stories are subsequently "read" and summarized by FRUMP. Because of the effectively random quality of the kinds of stories which are input, there is really no way that the system can be "primed" for this input other than through a general anticipation of the way in which news stories are normally written.

Whereas most previous programs "have a greater or lesser ability to draw inferences and reason," (3) the FRUMP uses its "understanding" of the text which it has read so far to make specific "predictions" about what ought to come later in the article. If FRUMP doesn't find what it is looking for, it starts again with a different interpretation.

Actually, DeJong's use of the word "prediction" is somewhat misleading as he uses it, although a better word is hard to find. The FRUMP system uses a system of scripts which are called upon when clues are found in the text indicating the probable subject matter of the text. The "sketchy script" then provides a list of important events to look for which are then included in the printed summary. DeJong provides an example of a "sketchy script", this one for demonstrations.

```
PREDICTED EVENT 1:
     THE DEMONSTRATORS ARRIVE AT THE DEMONSTRATION

LOCATION.
PREDICTED EVENT 2:
     THE DEMONSTRATORS MARCH.
PREDICTED EVENT 3:
     POLICE ARRIVE ON THE SCENE.
```

PREDICTED EVENT 4:
THE DEMONSTRATORS COMMUNICATE WITH THE TARGET OF

THE DEMONSTRATION.
PREDICTED EVENT 5:
THE DEMONSTRATORS ATTACK THE TARGET OF THE

DEMONSTRATION.
PREDICTED EVENT 6:
THE DEMONSTRATORS ATTACK THE POLICE.
PREDICTED EVENT 7:
THE POLICE ATTACCK THE DEMONSTRATORS.
PREDICTED EVENT 8:
THE POLICE ARREST THE DEMONSTRATORS. (4)

DeJong points out that the previous natural language understanding systems kept semantic analysis portions of their operations separate from stored information about the context of what has already been said. The FRUMP system, however, uses information extracted from the previous text to guide the parsing of the current context. The current context is the representation in memory of the previously processed text from the article. When a prediction is substantiated, it is added to the current context. Thus substantiated predictions can be used for the basis of further predictions.(5)

Based on the script provided, the SUBSTANTIATOR searches the text for certain kinds of phrases that will satisfy the current prediction. If the prediction is substantiated then this information increases the probability of accuracy (or appropriateness of the script) and the system continues with the next prediction. If, however, the prediction is not substantiated, then the system reassesses the situation and may call upon another script.

There are three ways that the FRUMP system is able to call up a script: 1) by Explicit Reference; certain words contain within their definition accessed by the computer the order to call upon the corresponding script. For example, if the system looks up the word "arrest" (in any of its forms), part of the definition may be understood as the actual script which guides the understanding of the rest of the article, 2) by Implicit Reference; actually, another word, such as "crime", may explicitly refer to its own script. However, because scripts are often interactive, the crime script may contain an activation of the arrest script, 3)Event-Induced Activation; in this case, although the script name may not be explicitly (or implicitly) called, the story may use words which actually contain essentially the same menaing as the script name. Evidently, the definitions for these words or phrases simply contain a directive to consult the appropriate sketchy script or are recognized as part of that definition to begin with. Precisely how this connection is made is not made clear by DeJong. Once a script is selected, it provides a context for interpreting ambiguous words, or words which may have radically different meanings depending on context in order to satisfy predictions.(6)

The system uses syntax only to determine where in the sentence it needs to perform semantic analysis. If it is trying to satisfy a prediction a prediction, it will only consider the meanings of words likely to satisfy predictions. "The semantic predictions determine exactly which word and word sense will be used to build the predicted conceptual item. This way the system doesn't need to consider the meanings of all the words it encounters. It knows what it's looking for and where it's likely to find it." (7)

Although this system may by its output give an impression of understanding, there are certain shortcomings which indicate that a true understanding has not been achieved. First, as DeJong himself points out, the FRUMP system cannot "understand" situations for which it has no script. DeJong attempts to rationalize this shortcoming by noting that even humans who are not familiar with "what goes on in political conventions and why, for example, will have trouble understanding a news article about one" (8). Still, it seems that the human would be able to derive enough meaning from the article to ask questions about it so that he might then construct his own sctipt to be used the next time he reads about a political convention. The FRUMP system has no such script for creating new scripts. Furthermore, I would insist that no future version will be able to adequately imitate this human ability because it involves that mysterious ability of the self not only to create such a script where none previously existed, but to, if need be, create a script for creating scripts, or a script for creating scripts for creating scripts, etc., etc.. As long as we rely on explicit programming techniques we will not produce a machine that can generalize in such an adaptive fashion for the levels of such a regression are potentially infinite and the infinite cannot be represented explicitly in a finite computer program of the kind curently being written.

In addition, since the system's processing of news stories is dependent upon the degree of explicitness provided by the script, the system is incapable of paying attention to facts which might otherwise be interesting but which are not accounted for within the script.

The predictions which DeJong attributes to the program are actually performed by the author. The most active part the program plays is the confirmation or refutation of those predictions based on the kinds of words encountered in the text. This program uses the logic of computing machinery to interface a prewritten script with a text whcih hopefully describes a random instantiation of that script.

Furthermore, the range of applicability of this program is limited exclusively (by its inherent inability to create new scripts) to newspaper accounts of events. Even editorials are beyond its intellectual grasp. It would seem that if a true understanding of the words the program reads is achieved then something ogf the format of an editorial ought not pose any major problem for the understanding mechanism. After all, both are written in the same language. The very techniques which make this such a powerful program within its domain limit it from stepping over into other domains. The FRUMP system only applies a template script to a given text. As there is no adaptive provision for assembling new scripts to fit new situations, we msut agree that although it may seem to understand some of what it reads, the same words combined in different (yet still meaningful) ways may create situations which it cannot understand. Can we call such "understanding" intelligent?

The actual shortcomings of the program make themselves known when it is out playing in its own yard. News items for which FRUMP may actually have a script may conceal more complexity than that with which the corresponding script was designed to cope. Important or interesting facts which may make the story unique and perhaps highly controversial may be passed over by FRUMP. Its selection of the importance of facts is not based on an undestanding of the story but on a prefabricated, generic outline of what the story is probably about. We may forgive the system for not anticipating such things, but not for ignoring them.

ROGER SCHANK AND THE SEARCH FOR CONCEPTUAL CONTENT

In "Finding the Conceptual Content and Intention in an Utterance in Natural Language Conversation." Roger Shank discusses a system by which he posits that a computer may enhance its understanding process by imitating a human element of expectation. By making predictions about what may follow conceptually given a certain context, the machine may come to build expectations which enable the computer "to discover not only the conceptual content of an utterance, but also the intention of that utterance in context."

Schank asserts that previous studies of grammaticality are misguided as the proper study with respect to communiction and artificial intelligence, he says, is that of the meaning which is communicated and not (9) so much the grammaticality of the communication. Much that communicates meaning occurs outside standard rules of grammar. Still we can derive meaning from ungrammatical utterances by referring to the context in which they occurred.

Schank enumerates six levels on which these expectations may be fomulated (10). The first level on which expectations may occur is the grammatical level. Based on our knowledge of the language, we can anticipate what part of speech may concievably come next in an utterance. This idea is very similar to Douglas Hofstadter's discussion of well-formed-formulas in his p q system. With a basic knowledge of the grammar of a language we can point out strings which, although they may incorporate legitimate symbols of the language, combine them in a way that is meaningless. For Schank this implies that as we read a string (or sentence) we apply some internalized version of these rules as we move through the string. At each step we know what kinds of parts of speech are legal within the system. We have expectations. Similarly, at the conceptual level do we maintain such expectations. Given the conceptual context of an utterance, we maintain certain expectations concerning what kinds of concepts may follow and make sense. Closely related to this level, but slightly more sophisticated, is an expectation of those concepts among the sensible concepts which, given the emotional and situational context, are more likely than the others. On the conversational level, Schank points out that people speak to communicate. The speaker, by speaking, implicitly intends to achieve some desired effect in the mind of his listener. Therefore we may derive implicit information about the speaker's intentions.

Expectations may be formulated on a fifth level which draws from one's constructed world model based on information applicable to the given situation derived from memory. This includes one's mental model of the speaker. One may compare the speaker's words with one's historical model of that speaker to construct an expectation of how seriously to consider the given utterance.

Schank describes a sixth level at which these expectations are operative. It is these expectations which when applied to one's cultural norms determine the listener's course of action. If, as in Schank's example, the speaker is seriously contemplating the execution of a muder, then there may be certain preventative measures which the listener will likely consider as the consequence of his expectations.

Schank proceeds to present an analog of his method of representation of conceptual content carried within a natural language utterance (11). For Schank, the meaning which occurs in our minds and the language we use to express this meaning are clearly distinct. Schank holds that many different sentence forms are capable of expressing the same meaning. As he claims, "evidence indicates that people cannot remember the surface forms of sentences and only remember their meanings" (12). Schank insists that the meaning which is common to the many alternative sentence forms is expressed in the mind in only one form. Therefore, the present section of his paper deals with reducing the input text to a universal form which would look the same no matter what any particular version of the vast number of possible sentece formations from which it was derived looked like. He calls this method "Conceptual Dependency" (13).

In this method, the realtionships within an utterance are expressed by a system of arrows which seem to have a grammar of their own. In this way, a sentence comes to look more like a map or diagram. The kind of arrow used depends on the kind of relationship it needs to express.

In processing the input, the computer first scans the sentence for a verb. When it finds one, it classifies it according to the kind of action it represents. Each kind of action will require certain other kinds of words to provide a meaningful context. The identification of the verb then opens up a list of the kinds of contextual support words which need to be found in the text. The hardest part, of course, is finding the verb. This is done through a syntactic parse. Mainly through a process of elimination, all words that can't be verbs are found and the verb is pinpointed based on grammatical relationships dictated by these other words (14).

Once the verb is established, the system establishes what Schank calls "Actor <-> Actcombinations " (15). The appropriate types of words are searched for in order to establish who is doing what to whomever or whatever. When candidates are found, they are tested against the "Verb-Act dictionary" to determine whether the relationship established is a plausible one. If it's not, it starts over. If it is plausible, it is included in the network which will be the graphic representation of the meaning of the sentence.

Once the act is determined, it serves as a set of predictions about what is happening in the rest of the sentece.

and directs the subsequent parsing of the sentence. At this point, the work is essentially a matter of filling in the blanks.

According to Schank, the next step in understanding is that of "extracting the presupposed information implicit in an utterance". (16) Most of the meaning that is implicit can be made explicit through an appeal to the contextual meanings of the words which are used. Schank uses the example, "I like books." To the computer, this makes no sense at first because the category of things to be liked excludes inanimate objects. What is implicit here is that the speaker likes what it is that he can do with the inanimate objects. So the computer looks up "books" and finds that it is one of a class of things designed explicitly to be read by humans. When the computer returns to the sentence, it does find a human present which is conceptually capable of maintaining the appropriate relationship of reading books. The computer then understands that it is the reading of these books which "I" likes. (17)

Words may have implicit meaning within the context of their utterance. But the utterance may also have implicit meaning within the broader context of the social and geographical environment in which it is uttered. Therefore, an utterance must be analysed with respect to the intentions of the speaker whithin his social context. Using the same graphic representation system, Schank is able to attach the consequences of actions to their meanings, thus providing a glimpse at possible or likely reasons for the actions. Thus all actions include as part of their definitions their universal consequences, such as satiation for the verb of nutritional consumption. By matching the consequences of the acts discussed within an utterance with the relevant elements of the social context, we may understand the implications of the utterance within its context.

Schank does not provide any examples of the results of such programming techniques. Judging by what he does talk about, the system would be designed mainly for answering questions about input text. Perhaps even making predictions about what will happen in a text before all the text is input.

Schank continually refers to his method as one of representing meaning "non-linguistically." However, his method utilizes words and logical operators. These symbols are assigned meaning not by the computer but by the humans who study them. As these symbols only assume meaning as the result of an isomorphism established between these words and the real world, all symbols become linguistic by the very virtue of their utilization by the system. Furthermore, it is not likely that any system would be capable of manipulating pure meaning elements in the absence of words or other linguistic, symbolic devices. The system may be able to reduce words (or expand them) and manipulate them in ways which would seem to human observers to indicate understanding, but unless the system is capable of establishing the same (or similar) isomorphisms as those established bv humans, we must not consent to an assertion of artificial understandinq.

Such a view assumes that the encoder has direct access to the structure of realitv. It assumes further that lanquaqein some wav caotures a comoletelv isomorohic relationshio with realitv. However, it mav be arqued that as realitv is onlv accessed indirectlv throuqh sense imoressions that lanquaqe can onlv be isomorohic to these imoressions and not realitv itself. Furthermore, I would hold that in fact lanquaqe helos shaoe these sense imoressions. It divides them uo into units and entities whoselinquistic svmbolizations mav be mentallv manioulated inwavs unavailable to the linquisticallv deorived. Schank asmes there is onlv one realitv to be reoresented and that hece all lanquaqe must contain analoques of some of its strucure. But as such a realitv is inherentlv inaccessable tous, we must entertain the oossibilitv of an infinite numberof realities, some of which are oerceived throuqh a conjunction of lanquaqe and a oersonal ooint of view.

The oresent svstem works onlv on input text. Thus, all meaninqs are orefabricated for the svstem's consumotion. This svstem manioulates the orefabricated meaninqs which are stored in memorv. But as these meaninqs are not matched to anv elements of the outside world, we mav conceive of this svstem beinq able to manioulate words in wavs which we miqht interoret as indicative of an understadinq but which, in their isolation from realitv, do not in the memorv of the comouter share anv corresoondence with the outside world in such a wav as to provide the computer with any intellectual abilities commensurate with the demonstrative manipulation of elements of the physical world such as we might expect from an intelligent system.

I would hold that unless any system is capable of creating its own meaning in a situation (i.e., applying words to events and situations in the outside world) then it is incapable of truly understanding any text in a way comparable to humans. Understanding ought to require the establishment of isomorphism between the symbolism of whatever "language" it uses and the corresponding elements of the real world. Therefore, computer understanding requires that the machine be endowed with the ability to interact with its physical environment, as do humans.

WINOGRAD'S BLOCK-HEAD COMPUTER

Terry Winograd rejects the notion of achieving machine understanding by concentrating on syntactic structures at the expense of semantic structures. He maintains that much of what happens in language is the result of both an intelligent speaker and hearers communicating in a setting. He stresses the importance of the setting which includes not only the context of the communication situation, but the models which each of the participants maintains of each other and their immediate environment.

Recognizing the impossibility of duplicating the world-knowledge which normal human participants bring to a dialogue, Winograd created a micro-world about which his program could have extensive familiarity. Within the context of this world, Winograd developed a program that could carry on a conversation in natural English and demonstrate its understanding by manipulating the three-dimensional geometric forms which populated the computer's micro-world. By limiting the scope of the world about which the computer needs to be knowledgeable, Winograd was able to concentrate on the problems of natural-language understanding.

According to Winograd, there are three main domains within the program; "There is a syntactic parser which works within a large-scale grammar of English; there is a collection of semantic routines that embody the kind of knowledge needed to interpret the meanings of words and structures; and there is a cognitive deductive system for exploring the cosequences of facts, making plans to carry out commands and finding the answers to questions. There is also a comparatively simple set of programs for generating appropriate English responses. In designing these pieces, the main emphasis was on the interaction of the three domains." (18)

The three main areas of the program do not operate sequentially, first parsing the input, then analysing this for meaning, translating the meaning into a procedural program and then generating responses to questions. Instead, all the areas interact constantly. "As soon as a piece of syntactic program is called to see whether it might make sense, and the resultant answer can direct the parsing. In deciding whether it makes sense, the semantic routine may call deductive processes and ask questions about the real world" (19)

The system keeps track of what has gone on before, and so when figuring the answers to questions checks to see if the user already knows the answer. If he does, then chances are that the comptuer misunderstood the question and so goes back and refigures what may be meant. Also, the program may offer information that isn't specifically asked for if it seems that such information is relevant.

Winograd's program is perhaps the most impressive attempt at machine understandinq to date. It has demonstrated its abilitv to understand natural lanquaqe inout throuqh its manioulations of the block-world. What this means is that, for all oractical ourooses, it has constructed isomorohic relationshios between its communicative lanquaqe and that domain which it considers to be the "real world." When we talk to it about a biq red block or a small qreen ovramid, it can demonstrate that it reallv does know what those words mean (rather than just blindlv orocessinq them) bv oointinq out and manioulatinq the corresoondinq realities.

Or does it? Because the block-world exists onlv on the screen (or in the memorv banks) we miqht consider this oarticular world to be somethinq of a mental construct maintained bv the comouter. Because this world exists fundamentallv as descriotions withinq the oroqram's data banks, it mav in fact feiqn understandinq. The oroqram is, in effect, omniscient about its world because the totalitv of that world exists within its oroqramminq. It does not need to trv to understand; all oroblems mav be resolved through an aooeal to the svstem's loqic functions.

Althouqh it is caoable of manioulatinq structures whose svmbolizations it also manioulates linquisticallv, these structures exist inside the comouter - not outside. Thsvstem mav demonstrate the abilitv to learn but it still lacks the more difficult abilitv of matchinq its words to structures which it oerceives to exist outside of itself. Were this svstem caoable of the same oerformance with resoect to the actual ohvsical universe, then we could consider it intelligent. Still, we must grant that the system's ability to process natural language input and generate appropriate responses based on some reality is certainly an impressive advance in artificial intelligence research. It is likely that the model world which it maintains in its memory will prove crucial to an eventual understanding of a more concrete reality.

PART III:

SOME PHILOSOPHICAL CONCERNS IN ARTIFICIAL INTELLIGENCE

WHAT IS TO BE DONE?

It seems that, in order to avoid the oroblems of interactinq with the infinite environment that is the real, ohvsical world, researchers have concentrated their efforts lnstead on understasdinq a written text which is in some wav deoendent uoon the real world. Those who write these texts oresuooose a familiaritv with or understandinq of the real world on the reader's oart. How then can we exoect a machine to understand a written text before it can understand its ohvsical environment?

In all orevious Artificial Intelliqence research, the qreatest obstacle to an authentic understandinq of a written text, which is the inabilitv of the machine to establish and manioulate isomorohism between the elements of the tooic of discourse and the corresoondinq events and thinqs in the oerceived world is its lack of access to sensorv inout mechanisms which themselves must be seen as necessarv but not sufficient for machine understandinq. Robotic techniques, the soecifics of which are bevond the scooe of this oaoer, must be combined with soecific oroqramminq methods. I shall trv to suggest some of the basic structural elements of such programming in what is to follow. If I may repeat myself, "What is needed is not simply an understanding of the meaning of a text, but an understanding of that text as a text. The machine must understand the text in relation to the machine itself" - (19).

ALAN TURING AND THE TURING TEST

During the 1930's, Alan Turing was concerned with whether or not it would be possible to build a machine that was capable of computing all the propositions within mathematics which were previously accepted as true and recognizing which ones were undecidable. He ultimately concluded that such a machine could never exist, that the notion was self contradictory, or, in other words, that the system of mnathematics could never be completely formalized. Mathematics must, in principle, remain incomplete. His proof will prove useful to our situation, as the ability to recognize undecidable propositions is at the core of the problem of artificial intelligence. My argument will parallel his.

Let us suppose that for each machine man can build there is a corresponding number, much like a Godel number, which encodes not only the machine's identity, but its function as well. Turing hypothesized a machine whose function would be to read the, I'll call it the "Turing Number," of any machine and subsequently adopt the function of that machine. It must be recalled that Turing had this idea before computers had been invented. In fact, his plan for a Turing Machine had been rejected by researchers in favor of what we now know as computers. In any event, a computer program is to a computer much as this Turing Number would be to a Universal Turing Machine. Now the twist which Turing offered to this idea was to feed the machine its own number.

Turing supposed that his special, "intelligent" program would be one designed to read any other program for any other machine and execute it.

Since whatever happens within the "mind" of the computer is specified within its programming, the crystalization of Artificial Intelligence, the ability to recognize undecidable propositions, to understand them as such, and jump out of an infinite regress, would have to be encoded withing the computer program or sequence of commands.

Now, if we should decide to let our intelligent program "contemplate" (run within its own running) "itself" (its own program), it ought to, in some sense, achieve an understanding of what its own program is -- what it does -- its identity. Humans have a difficult time doing this, but what would happen if a computer tried? Since this event would have to occur within the structure of the system's program, we should be able to imagine something of the general appearance of what would happen.

We have a computer which is being told, in essence, to READ AND IMITATE THE FOLLOWING PROGRAM, which is (READ AND IMITATE THE FOLLOWING PROGRAM: (READ AND IMITATE.... The system gets caught in an infinite regression which is marked (or rather, not marked) by the absence of closing parentheses which would return control to the next line of instruction in the previous level and ultimately in the original program. However, since it is continually being referred back upon itself, it is incapable of reaching such a point.

Humans are capable of extracting a crystal of meaning from this procedure which is something more than the simple meaning of the instruction to READ AND IMITATE... We understand the very fact that such an enterprise is, in such a situation, an infinite regression. It is this ability which we wish to capture and incorporate into what would be an intelligent system.

The problem is that even efforts to tack a recognition clause onto the program are doomed to lead to an infinite regression for the recognition clause would need to include the encoding of the identity of the program which in turn would include the recognition clause which would include the identity clause and so on.

Any device that would recognize an undecidable proposition would have to operate external to and independent of the entire program which is engaged by the undecidable proposition. Therefore, Artificial Intelligence cannot be achieved using the current technology of discreet state devices which read computer programs. Either a change must occur in the very nature of the hardware, or we must consider alternative methods of "programming."

In 1950, Alan Turing developed what has come to be known as the "Turing Test." His criterion for judging artificial intelligence was very simple. He would place a man and a computer in separate rooms, each room being linked with a third room in which a judge or judges sat with a teletype. The only means of communication between the three rooms would be the teletype, and the judges would not know in which room sat the man nor in which room was the computer. It would be the task of the judges to ask questions of each subject and, judging by the responses received, determine the identity of the occupant of each room. If the computer could fool the judges then it would be, as far as Turing was concerned, intelligent. And that was it.

Although his test made no explicit references to any criteria which he might have supposed to be necessary for intelligence, it was made implicit within the structure of the test that anything which might be essential for intelligence would present itself in the course of the test. If the computer could not fool the judge into thinking it was human then it failed.

The test has been criticized by those who feel that the test is not one of intelligence by of how well a computer can imitate a human. They emphasize that a computer which is not truly intelligent but which is very good at imitating human responses might pass the test. Although this seems in principle to be possible, considering the nature of the problems which have plagued researchers since Turing invented the test, I would hold that such a scenario is highly unlikely. The test is loose enough and has the potential for covering such a broad range of subject matter that any imitating system would break down were it not also truly intelligent. On the other hand, the judges may not be smart enough to tell the difference but this is not an in principle fault of the test.

It may be interesting to note that the initial criticism implies an even more likely flaw in the test. It cannot be denied that the test does indeed only test the computer's ability to imitate being human. This is the goal of the computer in taking the test. It is entirely likely that a different form of intelligence, although truly intelligent, might fail the Turing Test because it was no good at playing human. However, such failure must not be construed as an iron clad indication of a lack of intelligence. The Turing Test depends, to a large extent, upon an intelligence's access to real world knowlege. A computer which lacks complete world knowlege, or a least knowlege comparable to that of a normal man (within the cultural context of the Judge) would likely fail the test even though it may possess an acute degree of intelligence. Real world knowlege and intelligence are not the same thing.

We may rearrange the situation by substituting in the role of the computer another, non-human form of intelligence. One which is obviously non-human and which we know to be highly intelligent is the dolphin. If we assume for our new situation the existence of an automatic translating device so that the dolphin might hear the questions in its own language and be able to deliver its responses in English through the teletype, it is not entirely certain that even the dolphin would pass the test, although studies of dolphin brains have indicated that they may be even more intelligent than humans. The dolphin has no (or extremely little) experience with the world of men. The world of the dolphins has very little intersection with the world of men.

If we intend to recognize the intelligence of the dolphin, and the evidence suggests that we should, then we must test machines for the processes and abilities which those intelliegences of which we are aware exhibit. We should not mistake a large store of data about the world of men or the possession of scripts about situations in that world for those aspects of intelligence which gave rise to such a world in the first place. Intelligence seems to be more of a potential for acquiring knowledge and manipulatig it in the right ways than a simple possession of or access to large quantities of data.

The reason we would want a dolphin to pass an intelligence test is not because of any degree of familiarity he has with the world of men but because the dolphin exhibits the potetial for understanding something about certain aspects of that world more so than a dog or cat, (creatures which certainly have much more familiarity with our world but which obviously lack the intelligence to understand it).

Although the goal of the computer in taking the Turing Test is to fool the judges, the good thing about the test is that toward this goal, the test is open. The Judges need not restrict themselves to these species-specific kinds of questions and are free to develop questions and situations which they feel are actually genuine tests of intelligence.

OBJECTIVITY AND SUBJECTIVITY: THE BOUNDARIES OF THE SELF

One asoect of human oersonalitv which we would not normallv exoect a comouter to exhibit is emotion. But what is the connection between emotion and that asoect of understandinq which so far eludes Artificial Intelliqence research?

According to Robert Solomon, it is emotion which assiqns meaninq to the relationshios between the self and the oeople and things which the self encounters in the world. Thev are an exoression of teh value of the difference between the self and the world and between the self and one's ideal self. Savs Solomon, "Everv emotion is an act of self-creation" (20), and, "In constitution ourselves, we constitute each other and our relationshios with each other, which in turn define us" (21).

Emotions, now, seem to olav a crucial role in intelliqence. In defininq the self with resoect to the worls, emotions create meaninq for the thinqs which the self encounters in the world. This is one thinq which so far comouters cannot do -- create meaninq.

So far, Artificial Intelliqence research has concentrated on the comorehension of orefabricated meaninqs. But now we must ask, "Is it likelv that a machine which is incaoable of creatinq its own meaninq for the thinqs in the world will be caoable of sufficientlv understandinq meaninq qiven to it bv someone who is capable of such creation?"

Much of understanding, for example, the actions of another intelligent human, requires that the understander participate in the process of verstehen -- that he put himself in the shoes, so to speak, of the other person whose actions he is trying to understand.

When we speak of "understanding", it is never in isolation. Understanding may only be contemplated within the context of an object of understanding and a subject who engages in understanding that object. In considering "understanding" then, we cannot confuse it with other things which are purely objective or subjective. It is rather an intermeshing of objective reality with subjective reality.

Understanding is closely tied with meaning and the debate on the location of meaning is far from resolved. But if we exclude textual material and concentrate on finding meaning in the things and events of one's everyday world, we must agree that as there is often no discernable intent in nature, any meaning which is percieved cannot be totally objective. But neither can it be totally subjective as it is rather spawned from the interaction of the observer with his environment.

In this sense, understanding is inherently never objective or subjective to the exclusion of the alternative. We could say that in its infinity, nature contains the referent of all possible meaning, that meaning is as infinite as nature itself. But the way we normally use the term "meaning" is more subjective. Meaning is always held relative to oneself. It does not exist in and of itself but is the product of the relationship between subject and object.

Conversely, meaning is always rooted in the objective world - it is things and events which are said to hold meaning for us. Because of this dual nature of both meaning and understanding, we can not approach them as objective phenomena. In considering "understanding" we must sumultaneously consider both its subjective and objective implicants.

We cannot avoid the realization that understanding is a phenomenon which is both subjective and objective. Understanding is actually a byproduct of this very dichotomy. Where there are no objects, there could be no subjects (and vice versa) as these are relative terms. Thus the very possibility of subjectivity implies the contingent necessity of objectivity and this fact implies the reverse.

Understanding is contingent upon this dialectic. But if Subjectivity implies Objectivity, and Objectivity implies Subjectivity, then in some sense the two must be equivalent. This is not to say that a subject must be equated with its object, but rather that the terms of subjectvity and objectivity denote different aspects of the same relationship. The foundation of this relationship, the definition of the limits of the self, is crucial to an unerstanding of the world but, relative to the possibility of any inherent status or qualities of these subjective and objective elements of the world, this relationship of subjectivity and objectivity is totally arbitrary. Subjectivty and objectivity are not inherent properties of elements of the real world, but ratheer they are relational poperties induced by the presence of awareness.

The self enjoys a unique status within the dichotomy of objectivity and subjectivity as it is capable of adopting both aspects simultaneously, as it ought for the reader at this moment (at once it may be both the object and the subject (agent) of understanding). In considering the self, one shifts the tables so to speak and makes something which is normally the agent of understanding the object. At the same time, one attemps to defy this transition by maintaining one's subjective role as the agent of understanding. The self, then, is something which seems naturally to ignore the dread dichotomy. How well it does so might best be guaged, however, by the self's ability to understand itself.

Even as it can conceive of its objectivity, the self cannot escape its subjectivity if it desires to achieve understanding. When it realizes that both objectivity and subjectivity are arbitrary it is faced with an even more striking question; Is there a true nature of reality and if so, what is the relationship of the self to that reality? Or, to put it another way, if the self as a subjective unit is only relative then is the self real at all? Are there more fundamental grounds which define selfhood?

Perhaps a practical discussion of one of the unique functions of self-hood may bring us closer to the point. It is the very elusive nature of the self which enables it to stand outside of its own activity. It is the unique property of the self that it can consider itself both objectively and subjectively that enables it to step outside of something like an infinite regression and recognize not simply the facticity of the infinite nature of this regression but to interpret the meaning of this fact with respect to itself and the value of its time and labor against the value potentially (or impotentially) derivable from such an infinite enterprize. It is precisely this elusive nature of the self which provides it with the ability to recognize that continued pursuit of an undecidable proposition is a waste of time. And this is the key to intelligence which is sought. We could not reasonably call anything intelligent which was not capable of this realization.

Now what is the self that naturally exists outside the dichotomy of our everyday existence and yet uses this dichotomy, even though it is only arbitrary, to understand its world? If the objectivity of the world is not itself an objective property of that world, and furthermore, if we have not access to anything which may be an objective propert of that world, then it is ullusory to believe that wht we ever do understand is reality itself. But if what we perceive as the objective world is illusory then all attempts at creating machines which may share such mistaken impressions (or come up with entirely new, yet still mistaken impressions of its own) seems a little off base.

Now, if emotion is an evaluation which defines the self and the world, then it would seem that that which it is a recognition of or an assignment of must stand outside both of these. If it is that which defines both subject and object it cannot be contained by either class. We cannot assign quality to an object, and we cannot recognize quality in an object; but rather, quality, as Robert Pirsig, in Zen and the Art of Motorcycle Maintenance says, is the relationship of the subject with an object. Objects have no quality until they are recognized by a subject. The emotion we experience, then, must be a reaction to this recognition of an object. "The very existence of subject and object themselves is deduced from the Quality event. The Quality event is the cause of the subjects and objects which are then mistakenly presumed to be the cause of Quality!" (22). Later, he says:

> At the cutting edge of time before an object can be distinguished, there most be a kind of non-intellectual awareness, which he (Phaedrus) called awareness of Quality. You can't be aware that you've seen a tree until after you've seen the tree, and between the instant of vision and the instant of awareness there must be a time lag...
>
> Reality is always the moment of vision before the intellectualization takes place. There is no other reality. This preintellectual reality is what Phaedrus felt he had properly identified as Quality. Since all intellectually identifiable things emerge from this preintellectual reality, Quality is the parent, the source of all subjects and objects. (23)

Our recognition of things in the world as specific things in the world is an objective interpretation of the perception of Quality. Our emotional reactions to those objects are our subjective interpretations of Quality. Once we are aware of things in the world, a distinction has been made. The result of this distinction is one form of meaning.

But if Quality is the true reality, then any mechanism which can perceive things in the world can perceive Quality. Even if any system were capable of perceiving Pirsig's Quality, such non-distinction seems a relatively useless function to assign to a machine. It is the application of linguistic distinctions to this Quality which forms the basis for what we normally consider to be understanding. Therefore, it is the systematic application of linguistic "isomorphism" to the raw, sensory experiences resultant from a subjective relationship with reality that we interpret as intelligent understanding of the world. If it is capable of creating a model of self then emotions and values can be portrayed as they are with humans; subjectively. It does not matter if this perspective is an illusion; the illusion is practical.

"An emotion is a basic judgement about our Selves, and our place in the world, the projection of the values and ideals...according to which we live and through which we experience our lives" (24). The basic value we, as humans, project onto our world is the value of life as opposed to death. As living organisms, we must satisfy certain prerequisites to maintain our living status. Those things which contribute positively to such maintenance are considered to have "good" quality, those which detract are considered to have "bad" quality.

However, such evaluative impressions must occur at a higher level of mental processing than that which we need to consider at the moment. These evaluative terms, being linguistic themselves, denote concepts which may be learned along with a wealth of other concepts. At the foundation of mental processing I anticipate a sorting of raw data in very simple and highly empirical terms (x is greater than or equal to y, etc.). Such basic values as mentioned earlier might then emerge from the programming of a robot which would need to find its own sources of energy in the world just as we must find food.

Such a system, says Pirsig, "would seek analogues, that is images and symbols from its previous experience, to define the unpleasant nature of its new environment and thus 'understand' it" (225). It is assumed here that the understanding of one's environment as less than optimal with respect to a desired, possible configuration of the worl may constitute a motivation for the exhibition of intellgent behavior. If the divice is, on the other hand, plugged into a never-ending source of energy, in its isolation from the world it will find no fundamental motivation for action. It is assumed that one driving force in the evolution of intelligence is the relative scarciy of species-relevant foodstuffs. The scarcer food becomes, the more intelligent or powerful must be those who seek it. Although in this context Pirsig was concerned with organic organisms, I believe that a program which employed associative memory techniques and which had to find its own way in the world and which was endowed with this basic value system and the ability to learn from its errors would similarly employ and therefore be capable of understanding an ever increasing multitude of events.

The quality of mind which is sought in Artificial Intelligence research must reside not in the programming which constitutes its brain's capabilities (its potential) nor in anything outside this potential (anything which can be the object of the attention of this potential) but in the interaction between the two. In this sense, mind is a product of the interaction between the brain and its environment through the agency of the body. Because we choose to name it, it is a construct of the self in understanding its interaction with the world.

The body provides the brain, or the self, with the ability to change its perspective, to adopt an infinite number of alternative perspectives, and when it learns that there are certain aspects of things in the world which remain constant, that is, that all things, by virtue of being things, have some quality or qualities which, by not changing as the perspective changes, (that Quality which is the identity of the reality; for example, the sphere's being equidimentional) come to be known as the identity of those elements of reality. When the self realizes this it begins to associate its created meanings with those available to it through language. Some are subjective, but some, like that of the sphere, are objective. It is by scanning and comparing such tags that metaphor is made possible. In order to understand new things in its environment, the robot compares the qualities which it does perceive with other things in its memory which share these qualities, perhaps as universals, perhaps as particulars, dependant upon the perspective.

Further, it is this ability to adopt different perspectives which will enable such a device to understand the actions and discourse of others who are engaged in similar activities. It may not understand all things completely at first, perhaps very few, but it will learn.

The relationship between learning and understanding is complex when we consider that much of what happens in both is a product of the relationship between language and reality. I would hold that although there may be learning outside of language, an understanding of anything on the level required by the Turing Test necessitates the ability to manipulate representations of reality in natural language.

Language can never be entirely isomorphic to all of reality because any such instantiation would neccessitate that such a system include an element which is completely isomorphic to itself. In order to be isomorphic to all of reality, language would neccessarily stand outside of reality. This is clearly impossible.

There must be more to language than isomorphism. On a more fundamental level, language must respect logic. The supremacy of logic penetrates the very concept of language. If language is isomorphism, then at the foundation of language is the assumption that once a system of symbolization is adopted, within that system there are proper and improper ways of establishing isomorphism with reality. The essence of language is that some combinations of symbols will be ismorphic with some corresponding portion of reality. Language assumes that there are correct and incorrect symbolizations injust this sense. Prior to language then is the understanding of this basic law which describes all of reality and is symbolized by every language (by virtue of being language). That is that X implies not-X. In other words, if X is finite, then there must be some space in which X is not. We may assume that since we as agents of perception have physical limits (that since we in fact exist) that all other, non-transcendental things which exist are neccessarily finite, otherwise we could not exist in the form in which we do.

PART IV: CONCLUSION

By way of conclusion, I will attempt to apply some of the principles which I have developed here to a specification of the kinds of robotic techniques which I believe may be profitable in Artificial Intelligence Research. To this end let us consider that we must assume this principle of contradiction in interpreting the universe. Language, insofar as it implies isomorphism, must also imply non-isomorphism. In fact, it is the express function of language to draw such distinctions. Any utterance naturally divides its real-world implicants up into that which is isomorphic and that which is not.

If all sensory input can be encoded within a system, the lowest level of which uses symbolism which is isomorphic to the basic dichotomy of the principle of negation (i.e. the binary code), then any situation could be encoded in either one of only two possible ways (as each form of the symbolization would imply the alternative symbolization). Since each symbolization implies its alternative, only one need be recorded.

We can then create a second code which would be isomorphic to the first by replacing repetitious sequences of symbols with a number which would correspond to the number of repetitions of that symbol. A subnomial would retain the value of the original set of symbols being replaced. Such a system could be applied to an optical information processing system.

The visual information contained on one such grid could then be checked against memory in order to see if the holistic pattern has already been named. The basic problem with this is that the higher the resolution of the grid, the higher the number of possible grids in memory agains which the present grid must be checked.

One way around this might be to eliminate a large number of irrelevent possibilities by starting off the search in a low resolution. In this way, the highest resolution need only be used in the tricky cases and grids in lower resolutions would come to be tagged with connotations or expectations of the most frequently accessed high resolution grid which shares the same basic patterns.

In communications then, understanding may be established in the form of an isomorphism between patterns (or names of patterns) in memory and those called up by the input text. Put another way, understanding can be seen as the input message calling up isomorphic patterns in memory which can then be combined with other facets of meaning in a contingent, associative, lexical memory. Undersatanding on the former level can be affirmed by the positive or negative quality of correspondence between the symbolization created by the computer in its historv and that which is recognized as a foreign symbolization. It can then be suoolemented through its associative, lexical memory.

If we can get a machine to establish a linquistic exoression on anv level which is at least oartiallv isomoroic with the realitv it oercieves, a statement that its world is one way and not another, then we ought to be able to get it, on another level, to establish an isomorohism between its reoresentation of realitv and that of another oartv such that we could teach it our own language. We might then sav that it understood our language to the same extent taht we would sav a man understands anv language.

It seems to me that anv machine which is to exhibit artificial intelligence would require a minimum of orogramming as it is now understood. Rather, an intelligent machine would be equiooed most cruciallv with the caoacitv for learning and any further orogramming would take a form similar to behavioral conditioning until such a ooint at which sufficient linguistic comorehension is exhibited so as to allow a transition to more efficient modes of communication.

Endnotes

[1]Roger Schank, Reading and Understanding: Teaching from the Perspective of Artificial Intelligence (Hillsdale: Lawrence Erlbaum Associates, 1982), p.3.

[2]Joseph Weizenbaum, "Contextual Understanding by Computers", in Communications of the Association for Computing Machinery (New York: Association for Computing Machinery, 1967) p. 173.

[3]Gerald DeJong, "An Overview of the Frump System", in Strategies for Natural Language Processing, ed. by Wendy G. Lehnert and Martin H. Ringle (London: Lawrence Erlbaum Associates, 1982) p. 150.

[4]Ibid., pp. 150,1.

[5]Ibid., p. 157.

[6]Ibid., p. 164.

[7]Ibid., p.164.

[8]Ibid., pp. 174,5.

[9]Roger Schank, "Finding the Conceptual Content and Intention in an Utterance in Natural Language Conversation", International Joint Conference on Artificial Intelligence: Advance Papers (London: The British Computer Society, 1971) p. 444.

[10]Ibid., p. 444,5.

[11]Ibid., p. 446.

[12]Roger Schank, Reading and Understanding: Teaching from the Perspective of Artificial Intelligence (Hillsdale: Lawrence Erlbaum Associates, 1982) p. 67.

[13]Roger Schank, "Finding the Conceptual Content and Intention in an Utterance in Natural Language Conversation", International Joint Conference on Artificial Intelligence: Advance Papers (London: The British Computer Society, 1971) p. 447.

[14]Ibid., p. 448.

[15]Ibid., p. 448.

[16] *Ibid*., p. 449.

[17] *Ibid*., p. 449.

[18] Terry Winograd, "A Procedural Model of Language Understanding", in *Understanding Natural Language* (Artificial Intelligence Lab, MIT: Edinburgh University Press, 1972) p. 154.

[19] *Ibid*., p. 182.

[20] Robert Solomon, *The Passions: The Myth and Nature of Human Emotions* (Garden City: Anchor Press, 1976) p. 84.

[21] *Ibid*., p. 84.

[22] Robert Pirsig, *Zen and the Art of Motorcycle Maintenance* (New York: Bantam Books, Inc., 1974) p. 215.

[23] *Ibid*., p. 221,2.

[24] Robert Solomon, *The Passions: The Myth and Nature of Human Emotions* (Garden City: Anchor Press, 1976) p. 186.

[25] Robert Pirsig, *Zen and the Art of Motorcycle Maintenance* (New York: Bantam Books, Inc., 1974) p. 225.

Bibliography

Anderson, Alan Ross. Minds and Machines. Englewood Cliffs, N.J.. Prentice-Hall, 1964.

Barrow, H.G.."Representation and Use of Knowledge in Vision", Natural Research Review v. 29, April 1976 3-17.

Clark, Eve V.. "Non Linguistic Strategies in the Aquisition of Word Meanings", Cognition 2, 161-182 '73.

Coffey, Enrice J. "Computers and Intelligence", Journal of the British Interplanetary Society v. 33, July 1980 271-278

De Champeaux, Dennis. "Machine Understanding", Cirpho 3, 11-31 Spr. 1975,6.

Dreyfus, Hubert L.. What Computers Can't Do: The Limits of Artificial Intelligence. New York. Harper & Row, 1979.

Dyer, Micheal George. In-Depth Understanding. Cambridge, Mass: MIT Press, 1983.

Fink, Donald G.. Computers and the Human Mind. Garden City, N.Y.. Anchor Books, 1966.

Fogel, Lawrence J., and Others. Artificial Intelligence Through Simulated Evolution. Wiley, 1966.

Gregg, Lee W., Ed.."Knowledge and Cognition", Symposium on Cognition, 9th Pittsburgh, 1973. London. Wiley, 1974.

Haugeland, John. Mind Design: Philosophy; Psychology; Artificial Intelligence. Bradford Books, MIT Press, Cambridge, 1981.

Haugeland, John. "Understanding Natural Language", Journal of Philosophy 76, 619-632 N '79.

Hilts, Phillip J. "The Dream of Artificial Intelligence", Psychology Today, v. 17, Jan. 1983: 28-31,3.

Hite, David. "In Our Own Image: Why Computers Will Never Be As Smart as We Are", Washington Post Magazine. Mar. 21, 1982 24-31.

Hofstadter, Douglas R.. Gödel, Escher, Bach: An Eternal Golden Braid. New York: Random House, Inc., 1980.

Huggins, Hoyt W.. "Computers Alive?", University Review. v. 27, May-June 1976: 24-32.

Joseph, Earl C.. "Defense Computers and Software - What's Ahead for AI?", Concepts. v. 5, Autumn 1982: 141-147.

Kinnucan, Paul. "Artificial Intelligence: Making Computers Smarter", High Technology, v. 2, Nov.-Dec. 1982: 60-63, 66, 68-70.

Lao Tsu. Tao Te Ching, trans. by Gia-fu' Feng and Jane English. New York: Random House, 1972.

Lehman-Wilzig, Sam N. "Frankenstein Unbound: Towards a Legal Definition of Artificial Intelligence", Futures: Journal of Forcasting and Planning. v. 13, Dec. 1981: 442-457.

Martinez, Julia. "Robots and Consciousness", Dialogue (PST) 15, 81-92 MY '73.

Minsky, Marvin Lee, Ed.. Semantic Information Processing. MIT Press '68.

Mountain, Marian. The Zen Environment: The Impact of Zen Meditation. New York: Bantam Books, 1982.

Pirsig, Robert M.. Zen and The Art of Motorcycle Maintenance: An Inquiry into Values. New York: Bantam Books, 1982.

Preston, Glenn W.. "The Very Large Scale Integrated Circuit", American Scientist. v. 71, Sept.-Oct. 1983: 466-472.

Puccetti, Roland. "Pattern Recognition in Computers and in the Human Brain", Brit. J. Phil. Sci. 25, 137-154 JE '74.

Raphael, Bertrom. The Thinking Computer: Mind Inside Matter. San Francisco. W.H. Freeman '76.

Restak, Richard M. "Smart Machines Learn to See, Talk, Listen, Even 'Think' for Us", Smithsonian. v.10, Mar. 1980:48-56.

Ringle, Martin, Ed.. Philosophical Perspectives in Artificial Intelligence. Atlantic Highlands, N.J.. Humanities Press, 1979.

Schank, Roger J. "A Conversation with Roger Schank", Psychology Today. v. 17, April 1983: 28-30, 32, 34-36.

Schank, Roger C., and Colby, Kenneth Mark, Eds.. Computer Models of Thought and Language. Freeman, 1973.

Schank, Roger C..Conceptual Information Processing. NewYork: American Elsevier, 1975.

Schank, Roger C.. Dynamic Memory. New York: Cambridge University Press, 1982.

Bibliography Cont.

Schank, Roger C.. *Reading and Understanding.* Hillsdale: Erlbaum Associates 1982.

Schank, Roger C., and Abelson, Robert P.. *Scripts, Plans, Goals and Understanding: An Inquiry Into Human Knowledge Structures*. Lawrence Erlbaum Associates, 1977.

Simon, Herbert A.. "Artificial Intelligence Research Strategies in Light of AI Models of Scientific Discovery", *Naval Research Reviews*. v. 34, no.2. 1982:2-15.

Solomon, Robert. *The Passions: The Myth and Nature of Human Emotions*. Garden City: Anchor Press, 1976.

Stauffer, Lee. "Recognizing the Intelligent Computer", *SW Phil Stud.* 3, 70-74 AP '78.

Waltz, David L.. "Atificial Intelligence", *Scientific American*. v. 247, Oct. 1982: 118-124, 126, 128-133.

Wilks, Yorick. "Machine Translation", *New Scientist*. v. 76, Dec. 15, 1977.