Computational Linguistics

D. G. BOBROW, Editor

# Contextual Understanding by Computers

JOSEPH WEIZENBAUM
*Massachusetts Institute of Technology, Cambridge, Mass.*

**A further development of a computer program (ELIZA) capable of conversing in natural language is discussed. The importance of context to both human and machine understanding is stressed. It is argued that the adequacy of the level of understanding achieved in a particular conversation depends on the purpose of that conversation, and that absolute understanding on the part of either humans or machines is impossible.**

We are here concerned with the recognition of semantic patterns in text.

I compose my sentences and paragraphs in the belief that I shall be understood—perhaps even that what I write here will prove persuasive. For this faith to be at all meaningful, I must hypothesize at least one reader other than myself. I speak of *understanding*. What I must suppose is clearly that my reader will recognize patterns in these sentences and, on the basis of this recognition, be able to recreate my present thought for himself. Notice the very structure of the word "recognize," that is, know again! I also use the word "recreate." This suggests that the reader is an active participant in the two-person communication. He brings something of himself to it. His understanding is a function of that something as well as of what is written here. I will return to this point later.

Much of the motivation for the work discussed here derives from attempts to program a computer to understand what a human might say to it. Lest it be misunderstood, let me state right away that the input to the computer is in the form of typewritten messages—certainly not human speech. This restriction has the effect of establishing a narrower channel of communication than that available to humans in face-to-face conversations. In the latter, many ideas that potentially aid understanding are communicated by gestures, intonations, pauses, and so on. All of these are unavailable to readers of telegrams—be they computers or humans.

Work reported herein was supported (in part) by Project MAC, an MIT research program sponsored by the Advanced Research Projects Agency, Department of Defense, under Office of Naval Research Contract Number Nonr-4102(01).

Further, what I wish to report here should not be confused with what is generally called content analysis. In the present situation we are concerned with the fragments of natural language that occur in conversations, not with complete texts. Consequently, we cannot rely on the texts we are analyzing to be grammatically complete or correct. Hence, no theory that depends on parsing of presumably well-formed sentences can be of much help. We must depend on heuristics and other such impure devices instead.

The first program to which I wish to call attention is a particular member of a family of programs which has come to be known as DOCTOR. The family name of these programs is ELIZA. This name was chosen because these programs, like the Eliza of Pygmalion fame, can be taught to speak increasingly well. DOCTOR causes ELIZA to respond roughly as would certain psychotherapists (Rogerians). ELIZA performs best when its human correspondent is initially instructed to "talk" to it, via the typewriter, of course, just as one would to a psychiatrist.

I chose this mode of conversation because the psychiatric interview is one of the few examples of categorized dyadic natural language communication in which one of the participating pair is free to assume the pose of knowing almost nothing of the real world. If, for example, one were to tell a psychiatrist "I went for a boat ride" and he responded "Tell me about boats," one would not assume that he knew nothing about boats, but that he had some purpose in so directing the subsequent conversation. It is important to note that this assumption is one made by the speaker. Whether it is realistic or not is an altogether separate question. In any case, it has a crucial psychological utility in that it serves the speaker to maintain his sense of being heard and understood. The speaker further defends his impression (which even in real life may be illusory) by attributing to his conversational partner all sorts of background knowledge, insights and reasoning ability. But again, these are the speaker's contribution to the conversation. They manifest themselves inferentially in the interpretations he makes of the offered responses. From the purely technical programming point of view, the psychiatric interview has the advantage that it eliminates the need for storing explicit information about the real world.

The human speaker will contribute much to clothe ELIZA's responses in vestments of plausibility. However, he will not defend his illusion (that he is being understood) against all odds. In human conversation a speaker will make certain (perhaps generous) assumptions about his conversational partner. As long as it remains possible to interpret the latter's responses to be consistent with those assumptions, the speaker's image of his partner remains undamaged. Responses which are difficult to so interpret may well result in an enhancement of the image of the partner, by additional rationalizations which then make more complicated interpretations of his responses reasonable. When, however, such rationalizations become too massive and even self-contradictory, the entire image may crumble and be replaced by another ("He is not, after all, as smart as I thought he was".) When the conversational partner is a machine (the distinction between machine and program is not useful here), the idea of credibility may well be substituted for that of plausibility in the above.

"When in conversation with ELIZA, the user types in some statement or set of statements in natural language using normal punctuation and sentence structures."[1] "The user's statement is terminated by a double carriage return which serves to turn control over to ELIZA. ELIZA then analyzes the user's statement and generates some response which it types out. Control is then again with the user.

"A typical conversation is the following:

Men are all alike.
IN WHAT WAY
They're always bugging us about something or other.
CAN YOU THINK OF A SPECIFIC EXAMPLE
Well, my boyfriend made me come here.
YOUR BOYFRIEND MADE YOU COME HERE
HE says I'm depressed much of the time.
I AM SORRY TO HEAR YOU ARE DEPRESSED
It's true. I am unhappy.
DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY
I need some help, that much seems certain.
WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP
Perhaps I could learn to get along with my mother.
TELL ME MORE ABOUT YOUR FAMILY
My mother takes care of me.
WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU
My father.
YOUR FATHER
WHAT RESEMBLANCE DO YOU SEE
You are not very aggressive but I think you don't want me to notice that.
WHAT MAKES YOU THINK I AM NOT VERY AGGRESSIVE
You don't argue with me.
WHY DO YOU THINK I DON'T ARGUE WITH YOU
You are afraid of me.
DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU
My father is afraid of everybody.
WHAT ELSE COMES TO MIND WHEN YOU THINK OF YOUR FATHER
Bullies.
DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE

"The capitalized lines are the machine responses.

I described the original ELIZA program in some detail in [6]. For the present it is sufficient to recall that ELIZA's conversational powers are governed by a *script*. A script consists of a set of key words and key phrases each of which is associated with a set of so-called decomposition and reassembly rules. A decomposition rule is a data structure which searches a text for specified patterns, and, if such patterns are found, decomposes the text into disjoint constituents. A reassembly rule is a specification for the construction of a new text by means of recombinations of old and possible addition of new constituents. From one point of view, an ELIZA script is a program and ELIZA itself an interpreter. From another perspective, ELIZA appears as an actor who must depend on a script for his lines. The script determines the contextual framework within which ELIZA may be expected to converse plausibly.

"The gross procedure of the program is quite simple; the text is read and inspected for the presence of a key word. If such a word is found, the sentence is transformed according to a rule associated with the key word; if not, a content-free remark or, under certain conditions, an earlier transformation is retrieved. A rule-cycling mechanism delays repetition of responses to identical keys as long as possible. The text so computed or retrieved is then printed out."[1]

One of the principle aims of the DOCTOR program is to keep the conversation going—even at the price of having to conceal any misunderstandings on its own part. We shall see how more ambitious objectives are realized subsequently. In the meanwhile, the above discussion already provides a framework within which a number of useful points may be illuminated.

By far the most important of these relates to the crucial role *context* plays in all conversations. The subject who is about to engage in his first conversation with the DOCTOR is told to put himself in a role-playing frame of mind. He is to imagine that he has some problem of the kind one might normally discuss with a psychiatrist, to pretend he is actually conversing with a psychiatrist, and under no circumstances to deviate from that role. While some of the responses produced by the program are not very spectacular even when the subject follows his instructions, it is remarkable how quickly they deteriorate when he leaves his role. In this respect, the program mirrors life. Real two-person conversations also degenerate when the contextual assumptions one participant is making with respect to his partner's statements cease to be valid. This phenomenon is, for example, the basis on which many comedies of error are built.

These remarks are about the *global* context in which the conversation takes place. No understanding is possible in the absence of an established global context. To be sure, strangers do meet, converse, and immediately understand one another (or at least believe they do). But they operate

[1] The cooperation of the editors of the *Communications of the ACM* in permitting the extensive quotations from the paper "ELIZA," Vol. 9, No. 1, January, 1966, by the author is hereby gratefully acknowledged.

in a shared culture—provided partially by the very language they speak—and, under any but the most trivial circumstances, engage in a kind of hunting behavior which has as its object the creation of a contextual framework. Conversation flows smoothly only after these preliminaries are completed. The situation is no different with respect to visual pattern recognition—a visual pattern may appear utterly senseless until a context within which it may be recognized (known again, i.e., understood) is provided. Very often, of course, a solitary observer arrives at an appropriate context by forming and testing a number of hypotheses. He may later discover that the pattern he "recognized" was not the one he was intended to "see," i.e., that he hypothesized the "wrong" context. He may see the "correct" pattern when given the "correct" context. It doesn't mean much to say that the pattern "is" such and such. We might, for example, find a string of Chinese characters beautiful as long as we don't know what they spell. This, an apparent impoverishment, i.e., really a broadening, of context will enhance the esthetic appeal of a pattern. Similarly, many people think anything said in French is charming and romantic precisely *because* they don't understand the language.

In real conversations, global context assigns meaning to what is being said in only the most general way. The conversation proceeds by establishing subcontexts, sub-subcontexts within these, and so on. It generates and, so to speak, traverses a contextual tree. Beginning with the topmost or initial node, a new node representing a subcontext is generated, and from this one a new node still, and so on to many levels. Occasionally the currently regnant node is abandoned—i.e., the conversation ascends to a previously established node, perhaps skipping many intermediate ones in the process. New branches are established and old ones abandoned. It is my conjecture that an analysis of the pattern traced by a given conversation through such a directed graph may yield a measure of what one might call the consequential richness of the conversation. Cocktail party chatter, for example, has a rather straight line character. Context is constantly being changed—there is considerable chaining of nodes—but there is hardly any reversal of direction along already established structure. The conversation is inconsequential in that nothing being said has any effect on any questions raised on a higher level. Contrast this with a discussion between, say, two physicists trying to come to understand the results of some experiment. Their conversation tree would be not only deep but broad as well, i.e., they would ascend to an earlier contextual level in order to generate new nodes from there. The signal that their conversation terminated successfully might well be that they ascended (back to) the original node, i.e., that they are again talking about what they started to discuss.

For an individual the analog of a conversation tree is what the social psychologist Abelson calls a *belief structure*. In some areas of the individual's intellectual life, this structure may be highly logically organized—at least up to a point; for example, in the area of his own profession. In more emotionally loaded areas, the structure may be very loosely organized and even contain many contradictions. When a person enters a conversation he brings his belief structures with him as a kind of agenda.

A person's belief structure is a product of his entire life experience. All people have some common formative experiences, e.g., they were all born of mothers. There is consequently some basis of understanding between any two humans simply because they are human. But, even humans living in the same culture will have difficulty in understanding one another where their respective lives differed radically. Since, in the last analysis, each of our lives is unique, there is a limit to what we can bring another person to understand. There is an ultimate privacy about each of us that absolutely precludes full communication of any of our ideas to the universe outside ourselves and which thus isolates each one of us from every other noetic object in the world.

There can be no total understanding and no absolutely reliable test of understanding.

To know with certainty that a person understood what has been said to him is to perceive his entire belief structure and *that* is equivalent to sharing his entire life experience. It is precisely barriers of this kind that artists, especially poets, struggle against.

This issue must be confronted if there is to be any agreement as to what machine "understanding" might mean. What the above argument is intended to make clear is that it is too much to insist that a machine understands a sentence (or a symphony or a poem) only if that sentence invokes the same imagery in the machine as was present in the speaker of the sentence at the time he uttered it. For by that criterion no human understands any other human. Yet, we agree that humans do understand one another to *within acceptable tolerances*. The operative word is "acceptable" for it implies *purpose*. When, therefore, we speak of a machine understanding, we must mean understanding as limited by some objective. He who asserts that there are certain ideas no machines will ever understand can mean at most that the machine will not understand these ideas tolerably well because they relate to objectives that are, in his judgement, inappropriate with respect to machines. Of course, the machine can still deal with such ideas symbolically, i.e., in ways which are reflections—however pale—of the ways organisms for which such objectives are appropriate deal with them. In such cases the machine is no more handicapped than I am, being a man, in trying to understand, say, female jealousy.

A two-person conversation may be said to click along as long as both participants keep discovering (in the sense of uncovering) identical nodes in their respective belief structures. Under such circumstances the conversation tree is merely a set of linearly connected nodes corresponding to the commonly held parts of the participants' belief structures. If such a conversation is interesting to either participant, it is probably because the part of the belief structure being made explicit has not been consciously verbalized before, or has never before been attached to the higher level node to which it is then coupled in that conversation, i.e., seen in that context, or because of the implicit support it is getting by being found to coexist in someone else.

Backtracking over the conversation tree takes place when a new context is introduced and an attempt is made to integrate it into the ongoing conversation, or when a new connection between the present and a previous context is suggested. In either case, there is a need to reorganize the conversation tree. Clearly the kind of psychotherapist initiated by the DOCTOR program restricts himself to pointing out new connectivity opportunities to his patients. I suppose his hope is that any reorganization of the conversation tree generated in the therapy session will ultimately reflect itself in corresponding modifications of his patients' belief structures.

I now turn back to the program reproduced earlier. I hope the reader found the conversation quoted there to be smooth and natural. If he did, he has gone a long way toward verifying what I said earlier about the investment a human will make in a conversation. Any continuity the reader may have perceived in that dialogue—excepting only the last machine response—is entirely illusionary. A careful analysis will reveal that each machine response is a response to the just previous subject input. Again with the exception of the last sentence, the above quoted conversation has no subcontextual structure at all. Nor does the description of the program given in [6] give any clues as to how subcontexts might be recognized or established or maintained by the machine.

To get at the subcontext issue, I want to restate the overall strategy in terms somewhat different from those used above. We may think of the ELIZA script as establishing the global context in which subsequent conversations will be played out. The DOCTOR script establishes the Rogerian psychiatric context. The key word listings are guesses about what significant words might appear in the subject's utterances. The appearance of any of these words in an actual input tentatively directs the system's attention to decomposition rules associated with the identified word. These rules are hypotheses about what an input containing such words might mean. An important property of the system is that just as the entire set of key words need not be searched in the analysis of every input sentence, neither does the whole list of hypotheses about what *any* input might mean have to be tested in order to generate a response to a specific sentence. Should the set of hypotheses associated with the currently regnant key word be exhausted without success, there is a way (although not here described) to go to decomposition rules associated with lower ranking key words that may have appeared in the input string.

Much is known about an input sentence once it matches a decomposition rule. If, for example, the rule

(O MY (/FAMILY) DRINKS O)

matches a sentence, then we know that the subject matter is a family member, and that this person drinks. (The notation "(/ A B C)" means that in that position a word "tagged" with any of the symbols A or B or C is looked for. Words are tagged as part of the script. This mechanism is not further discussed here.) A reassembly rule can subsequently identify the family member in question, e.g., father. The successful matching of a decomposition rule—hence, the verification of a hypothesis—is therefore an opportunity not only to generate a plausible response but also to establish a subcontextual node. The laying down of such a node does not mean merely remembering what is now being talked about, but what is more important, making hypotheses about what is likely to be said next within the new contextual framework.

The realization of such a mechanism in an ELIZA script involves attaching a substructure to decomposition rules which says, in effect, that if a so modified decomposition rule matches an input sentence and a suitable subcontext exists, then control is transferred to a set of decomposition and reassembly rules appropriate to that subcontext. The details of mechanization are not important here. I will, however, show its effect by exhibiting the following conversational fragment from the DOCTOR program:

Because you are so stubborn.
IS THAT THE REAL REASON
Maybe
WHAT IS A BETTER REASON THEN
My stomach hurts.
TELL ME ABOUT YOUR PAINS IN GENERAL
My head bothers me, too.
YOU SEE SOME PSYCHOLOGICAL REASON WHY YOUR HEAD SHOULD PAIN YOU
Maybe.
YOU THINK COMING HERE MIGHT HELP ALLEVIATE YOUR HEAD PAIN
Certainly.
LET'S TALK ABOUT YOUR FEELINGS WHEN YOU THINK OF YOUR HEAD AND ITS PAINS

We are here in the middle of a dialog. The first sentence is an input to the program, the next a machine response, and so on in alternation.

The word "because" was the key that triggered the first machine response. At the same time the *reason* subcontext, so to speak, was established. Notice that the next subject input, i.e., the word "maybe", was interpreted as meaning "no," as is appropriate in that particular context, and that a contextually correct machine response was produced. The next input serves to establish a "hurt" subcontext. I believe the conversation is self-explanatory from that point on. Notice, however, that in the new subcontext the word "maybe" which was earlier interpreted to mean "no" is now seen as meaning "yes." Under still other circumstances it would be understood as indicating uncertainty.

My secretary watched me work on this program over a long period of time. One day she asked to be permitted to talk with the system. Of course, she knew she was talking to a machine. Yet, after I watched her type in a few sentences she turned to me and said "Would you mind leaving the room, please?" I believe this anecdote testifies to the success with which the program maintains the illusion of understanding. However, it does so, as I've already said, at the price of concealing its own misunderstandings. We all do this now and then, perhaps in the service of politeness or for other reasons. But we cannot afford to elevate this occasional tactic to a universal strategy. Thus, while the DOCTOR program may be useful as an instrument for the analysis of two-person conversations, and while it is certainly fun, its aim must be changed from that of concealment of misunderstanding to its explication.

Another difficulty with the system currently under discussion is that it can do very little other than generate plausible responses. To be sure, there are facilities for keeping and testing various tallies as well as other such relatively primitive devices, but the system can do no generalized computation in either the logical or numerical sense. In order to meet this and other deficiencies of the original ELIZA system, I wrote a new program, also called ELIZA, which has now replaced its ancestor.

The ELIZA differs from the old one in two main respects. First, it contains an *evaluator* capable of accepting expressions (programs) of unlimited complexity and evaluating (executing) them. It is, of course, also capable of storing the results of such evaluations for subsequent retrieval and use. Secondly, the idea of the script has been generalized so that now it is possible for the program to contain three different scripts simultaneously and to fetch new scripts from among an unlimited supply stored on a disk storage unit, intercommunication among coexisting scripts is also possible.

The major reason for wishing to have several scripts available in the core (i.e., high speed) memory of the computer derives from the arguments about contexts I made above. The script defines, so to speak, a global context within which all of the subsegment conversation is to be understood. We have seen that it is possible for a single script to establish and maintain subcontexts. But what is a subcontext from one point of view is a major (not to say global) one as seen from another perspective. For example, a conversation may have as its overall framework the health of one of the participants but spend much time under the heading of stomach disorders and headache remedies.

In principle one large, monolithic ELIZA script could deal with this. However, such a script would be very long and extremely difficult to modify and maintain. Besides, long exposure to computer programming should at least instill a healthy respect for subroutines, their power and utility. In the new ELIZA system scripts are in fact very much like subroutines. One script plays the role of the so-called "main program," i.e., it is the one for calling in and replacing subscripts. This is the one which, in an important sense, governs the computer's role in the conversation. Such a script may, for example, play the role of an initial interviewer in a hospital, i.e., a person who knows just enough about the field of medicine and the organization of the hospital to be able to determine, by means of appropriate questions, to whom the incoming patient is to be directed. A more highly specialized script then plays the role of that person. It may call on still further levels of scripts—and each script, except for the "main" one, can return the issue up one or more levels with information "explaining" that action.

We saw earlier how the word "maybe" was differentially interpreted in what might be called different micro contexts. The hierarchy of scripts now under discussion makes much more nearly global contextual distinctions possible. For example, the word "tone" probably means muscle tone to the internist but audio tone to the ear specialist. The occurrence of that word as a key word would establish an entirely different contextual mode, or call in a different subscript, depending on whether it was identified in an internist or an ear specialist script.

Of course, another important advantage of this kind of subscripting is that scripts (sub or otherwise) can presumably be written by specialists in a very narrow field. The fact that scripts now tend to be rather short means that a specialist might actually find the time to produce one as part of a larger cooperative effort.

The conversations reproduced below were governed by a script in the new ELIZA system. This script was designed to illustrate the power of the evaluator to which ELIZA has access. The key word and other such mechanisms are not very different from those described above. The important departure is that expressions can be turned over to the evaluator, evaluated, and the results displayed and stored.

The area of a circle is pi times the radius of the circle squared.
The area of a sphere is pi times the diameter of the sphere squared.
The area of a square is the side of the square squared.
The diameter of an object is twice the radius of the object.

The radius of a ball is 10.
What is the area of the ball.

THE ASSUMPTION THAT BALL IS CIRCLE IS CONSISTENT—SHALL I GO ON? no

THE ASSUMPTION THAT BALL IS SPHERE AND OBJECT IS CONSISTENT—SHALL I GO ON? yes

IT'S 1256.6359

What is the diameter of the ball.

IT'S 20

The radius of a dish is 10.
What 's the area of the dish.

THE ASSUMPTION THAT DISH IS CIRCLE IS CONSISTENT—SHALL I GO ON? yes

IT'S 314.15899

The radius of a globe is 10.
A globe is a sphere. A sphere is an object.
What is the area of the globe.

IT'S 1256.6359

The side of a plate is 10.
What is the area of the plate.

THE ASSUMPTION THAT PLATE IS SQUARE IS CONSISTENT—SHALL I GO ON? yes

IT'S 100

When I say board I mean plate.
What is 3 times the side of the board.

It's 30.

Please call that x.
Let f(z) be z squared minus z.
Compute f(x).

It's 870.

When I say ich sage I mean I say.
When ich sage meine ich I mean I mean.
Wenn ich sage wieviel ist meine ich what is.
Wenn ich sage quadrat meine ich squared.
Nun, sagen sie mir bitte, wieviel is x quadrat.

I DIDN'T UNDERSTAND YOU.

Ja, tut mir leid, ich sollte fragen wieviel IST x quadrat.

IT'S 900.

Thank you.

YOU'RE QUITE WELCOME, COME AGAIN

The particular example shown here is one in which the system is asked not merely to entertain but to perform some useful service. The script is one which establishes a context in which a few terms can be interpreted as giving directions about the evaluation of certain expressions. The number of keys which actually lead to decomposition rules is very small. They are as follows:

| | |
|---|---|
| I mean | |
| Is | (* a an) |
| Of | (* a an the) |
| What | (* the is) |
| Where | |
| Call | |
| Let | |
| Define | |
| Compute | |
| = | |
| Also | |
| Thanks | |

(A key of the form "Is (* a an)" will match either "Is a" or "Is an".) In addition, there are 1/ substitution keys like "Twice = 2*". The effect of that particular one is to substitute the symbols "2*" whereever the word "twice" appeared in the input text. Of course, the evaluator sitting behind ELIZA, so to speak, recognizes words like "SQRT", "LOG", etc. The function of this script is to interpret the user's wishes with respect to the evaluation of expression, perform certain translation functions on these expressions, and control the traffic between the input/output system of ELIZA and that of the evaluator.

Consider the dozen keys shown above. The sentence "*Let* me try to *define what the call of the* sea means" contains five of these keys. It could perhaps be understood by the DOCTOR but not by the program we are now considering. It would reply "I didn't understand you."

I call attention to this contextual matter once more to underline the thesis that, while a computer program that "understands" natural language in the most general sense is for the present beyond our means, the granting of even a quite broad contextual framework allows us to construct practical language recognition procedures.

The present script is designed to *reveal*, as opposed to *conceal*, lack of understanding and misunderstanding. Notice, for example, that when the program is asked to compute the area of the ball, it doesn't yet know that a ball is a sphere and that when the diameter of the ball needs to be computed the fact that a ball is an object has also not yet been established. Appropriate questions are therefore asked. But when, later on, the area of the globe is required, all the parameters are known and no questions are asked. The transcript also makes clear that the program now has memory and, of course, that it can compute.

I can imagine at least two questions that might have arisen in the mind of the reader in response to the above material. First, but least important, is what role the kind of psychiatric program shown here has in serious psychotherapy. My answer to that must be that psychotherapy is not my field and that I have no intention to pursue any therapeutic implications of this work. It is clear that there is considerable psychological material to be exploited here, material that may conceivably lead to some therapeutic developments. There may also be some use in this kind of technique in simulation of the neurotic processes. K. M. Colby, a practicing psychoanalyst, does propose to use conversational programs as psychotherapeutic tools vis-a-vis real patients. I believe the unsolved technical problems in the computer realm alone are so enormous as to guarantee the postponement of the realization of such a goal for yet a very long time.

The second question concerns itself with the implications of this work on so-called teaching machines. Here I feel on much firmer ground. It is my belief that the kind of programmed teaching that is based on multiple choice testing of the student is best done without the use of computers. With the techniques described here, it becomes possible to instruct the student in a corner of some field and then, the contextual framework having been established, to have the student interact with the machine in natural language.

The following conversation, based on a script written by Dr. Edwin Taylor of the Education Research Center of MIT, makes the point very well.

PLEASE CONSIDER HOW WE MIGHT SYNCHRONIZE TWO CLOCKS TO READ THE 'SAME TIME'. ONE OF THESE CLOCKS IS IN SAN FRANCISCO, THE OTHER IS IN NEW YORK. AS YOU THINK ABOUT THE PROBLEM, ASK ME ANY QUESTIONS THAT OCCUR TO YOU. WHEN YOU ARE READY TO PROPOSE A SYNCHRONIZATION METHOD, TYPE 'READY'. THEN I WILL QUIZ YOU ABOUT YOUR PROPOSED METHOD.

what do you mean by synchronize

FOR THE PRESENT MAKE USE OF THE EVERYDAY MEANINGS OF SUCH PHRASES AS 'SAME TIME' AND 'SYNCHRONIZE'. LATER WE WILL DISCUSS THE MEANING OF THESE TERMS MORE CAREFULLY

may i use telegraph signals

CERTAINLY, FEEL FREE TO USE TELEGRAPH

ready

NOW I WILL QUIZ YOU ON YOUR SYNCHRONIZATION METHOD DOES YOUR METHOD OF SYNCHRONIZATION INVOLVE THE PROPAGATION OF SOME SIGNAL—OR MATERIAL OBJECT—FROM ONE PLACE TO ANOTHER

yes

WHAT KIND OF PROPAGATING SIGNAL OR TRAVELING OBJECT DO YOU PROPOSE TO USE

a watch

WILL YOUR WATCH MOVE AT A SPEED VERY MUCH SMALLER THAN THE SPEED OF LIGHT

of course

GOOD. TRAVELLING CLOCKS CAN BE USED TO SYNCHRONIZE DISTANT CLOCKS TO ARBITRARILY HIGH ACCURACY IF THEY TRAVEL SLOWLY ENOUGH. YOU MAY LEAVE THE PROBLEM BY PUSHING THE 'QUIT' BUTTON TWICE. IF YOU WOULD LIKE TO EXPLORE FURTHER THE MATTERS OF PRINCIPLE ASSOCIATED WITH OTHER METHODS, CHOOSE ANOTHER WAY TO COMMUNICATE BETWEEN THE DISTANT CLOCKS
WHAT KIND OF PROPAGATING SIGNAL OR TRAVELING OBJECT DO YOU PROPOSE TO USE

light

WELL, OF COURSE THE LIGHT SIGNAL WILL NOT FOLLOW THE CURVATURE OF THE EARTH. SO LET US ASSUME THAT SOME SYSTEM OF RELAY TOWERS IS USED. NOW LET ME ASK, WILL THE LIGHT SIGNAL TRAVEL DIRECTLY BETWEEN NEW YORK AND SAN FRANCISCO OR WILL THE LIGHT SIGNAL TRAVEL TO OR FROM SOME INTERMEDIATE POINT

it will travel from st louis in both directions

IS YOUR INTERMEDIATE POINT EQUIDISTANT FROM NEW YORK AND SAN FRANCISCO

let's say it is, yes

MY UNDERSTANDING IS THAT YOU EXPECT TO USE TWO SIGNALS AND THAT THESE SIGNALS WILL EITHER CONVERGE ON OR BE SENT OUT FROM A POINT EQUIDISTANT FROM NEW YORK AND SAN FRANCISCO. IS MY UNDERSTANDING ON THIS POINT CORRECT

yes

WILL THIS PAIR OF SIGNALS CONVERGE ON THE INTERMEDIATE POINT OR BE SENT OUT FROM THIS POINT
they will be sent from the point
WILL THE TWO SIGNALS LEAVE THE INTERMEDIATE POINT AT THE SAME TIME

certainly

AHA, I'VE GOT YOU. YOU ARE TRYING TO SYNCHRONIZE THE NEW YORK AND SAN FRANCISCO CLOCKS TO READ THE 'SAME TIME'. HOW CAN IT BE VALID TO ASSUME A 'SAME TIME' FOR THE DEPARTURE OF A PAIR OF SIGNALS FROM THE MIDPOINT

because the pair of signals leave from the same point

RIGHT
THE PHRASE 'SAME TIME' HAS AN UNAMBIGUOUS MEANING WHEN APPLIED TO EVENTS THAT OCCUR NEAR THE SAME POINT IN SPACE. NOW YOU ARE GOING TO CONTINUE YOUR ARGUMENT BY SAYING THAT THE TWO SIGNALS THAT START FROM THE MIDPOINT AT THE SAME TIME WILL PRESUMABLY REACH NEW YORK AND SAN FRANCISCO AT THE SAME TIME. THESE SIGNALS CAN THUS BE USED TO SYNCHRONIZE THE NEW YORK AND SAN FRANCISCO CLOCKS. IS THIS THE WAY YOUR ARGUMENT WILL CONTINUE

yes

ONE FINAL QUESTION. WHAT ASSUMPTION ARE YOU MAKING ABOUT THE SPEEDS OF THE TWO SIGNALS THAT TRAVEL IN OPPOSITE DIRECTIONS

that these speeds are equal

VERY GOOD. YOU HAVE DEVISED A VALID METHOD FOR SYNCHRONIZING TWO CLOCKS, ONE IN NEW YORK AND ONE IN SAN FRANCISCO. I AM NOW RETURNING YOU TO THE CTSS SYSTEM

I have chosen to skew this discussion in the direction of "understanding" and have used that difficult word almost as if it were synonymous with "recognition." Actually, recognition is a prerequisite to understanding. The latter provides a test of the former. Even though no test of understanding can be defined with rigor, we do feel we are understood when the person we're talking to asks what we believe to be relevant questions and when he finally answers our question or solves the problem we posed. By such criteria, the conversations displayed above justify the assertion that the programs that produced the responses recognized textual patterns and understood what they were being told.

Finally, I wish to call attention to three pieces of work, two of which predate the programs discussed here and to whose authors I owe a considerable intellectual debt. The last is more recent but nevertheless highly relevant to my own current line of attack.

The SIR program of Raphael is capable of inferential data acquisition in a way analogous to that displayed in the ELIZA ball and sphere conversation displayed above. Notice that in that conversation the program had to infer that a ball was a sphere and an object. Once that inference was affirmed, the program retained the information by, in this case, associating with ball the fact that it is a sphere and an object and with sphere and object that ball is an instance of each, respectively. SIR is a program which specializes in establishing such relationships, remembering and invoking them when required. One of its principal aims was to establish methodology for formalizing a calculus of relations and even relations among relations.

Bobrow's program STUDENT is capable of solving so-called algebra word problems of the kind that are typically given in high school algebra texts. He uses a mechanism not very different from an ELIZA script. Its chief task is to transform the input text, i.e., the natural language statement of an algebra word problem, into a set of simultaneous linear equations that may then be evaluated to produce the desired result. A particular strength of his program is its power to recognize ambiguities and resolve them, often by appeal to inferentially acquired information but sometimes by asking questions.

The work of Quillian is mainly directed toward establishing data structures capable of searching semantic dictionaries. His system could, for example, decide that the words "work for" in the sentence "John works for Harry." mean "is employed by", while the same words appearing in the sentence "That algorithm works for all even numbers that are not perfect squares." mean "is applicable to."

Each of the computer papers referenced below represents an attack on some component of the machine understanding problem. That problem is not yet solved.

Received April, 1967

REFERENCES

1. Bobrow, D. G. Natural language input for a computer problem-solving system. Ph.D. Thesis, MIT, Dept. of Mathematics, Cambridge, Mass. 1964.
2. Colby, Kenneth Mark. Computer simulation of change in personal belief systems. Paper delivered in Section $L_2$, The Psychiatric, Sciences, General Systems Research, AAAS Berkeley Meeting, December 29, 1965. To appear in *Behav. Sci.*, 1967.
3. Quillian, M. R. Semantic memory. Ph.D. Thesis, Carnegie Inst. of Technology, Pittsburgh, Pa., 1966.
4. Raphael, B. SIR: A computer program for Semantic Information Retrieval. Ph.D. Thesis, MIT, Dept. of Mathematics, Cambridge, Mass., 1964.
5. Rogers, C. *Client Centered Therapy: Current Practice, Implications and Theory*. Houghton Mifflin, Boston, 1951.
6. Weizenbaum, Joseph. ELIZA—a computer program for the study of natural language communication between man and machine. *Comm. ACM 9*, 1 (Jan. 1966), 36–45.