DOCUMENT RESUME

ED 036 783 AL 002 062

AUTHOR Hays, David G.
TITLE Applied Computational Linguistics.
PUB DATE Sep 69
NOTE 19p.; Paper delivered at the International Conference Congress of Applied Linguistics, Cambridge, England, September 1969

EDRS PRICE EDRS Price MF-$0.25 HC-$1.05
DESCRIPTORS Analog Computers, *Applied Linguistics, *Communication (Thought Transfer), *Computational Linguistics, Computer Assisted Instruction, *Computer Programs, Digital Computers, Information Retrieval, *Information Storage, Linguistics, Machine Translation, Surface Structure
IDENTIFIERS *Action Patterns

ABSTRACT

Much work in computational linguistics, e.g. the preparation of concordances and text files, has dealt strictly with the surface of language, treating it as nothing more than strings of characters or phonemes. The "classical" scheme, developed as a result of dissatisfaction with the inability of such surface systems to deal with problems such as ambiguity, consists of surface processing, syntactic processing and semantic processing, with the object of obtaining an expression for the content of the input text; work with programming systems for generation of sentences with transformational grammar is representative of this tradition. It must be recognized, however, that the essential characteristic of language is its connection with information and that language is the external manifestation of the human capacity to process symbols in such ways that information is retained. This capacity should be the object of linguistics, and rules of grammar should describe those "action patterns" which underlie human symbol processing. Recent work in applied computational linguistics recognizes the importance of this conception and should therefore lead to wider computer applications, perhaps even to real man-machine conversations and the concomitant use of the computer as an imaginative consultant for a wide range of problems. (FWB)

APPLIED COMPUTATIONAL LINGUISTICS

David G. Hays
State University of New York at Buffalo

In the United States of America, applied linguistics was for a long time an elegant, perhaps somewhat overblown, term for the use of linguistic ideas and findings in the teaching of languages to nonnative speakers. The Center for Applied Linguistics, of Washington, D.C., has done, during more than a decade, excellent work in the promotion of English teaching abroad, and development of the teaching of Standard English to speakers of nonstandard dialects. But its decennial celebration was an orgy only for language teachers.

The congeries of activities presently given the collective name, computational linguistics, has developed around machine translation as first nucleus. Some would also mention information storage and retrieval, although the use of linguistic theories and data in this field has been rarer. After beginning with little theory, machine translation adopted as much linguistic theory as it could; those working on information retrieval have enjoyed their best successes by renouncing linguistics--and it is still fashionable, in some circles, to denounce linguistics as a frill and unprofitable luxury.

Machine translation was judged too costly to be useful, at least in the existing circumstances of linguistics and computation, by the Automatic Language Processing Advisory Committee (1966)--of which I was a member. If machine translation were the whole of computational linguistics, and if language teaching were the whole of applied linguistics (and if there were no new ideas since 1966), I should have little to say to you here. Neither condition is true, however; in fact, I must choose carefully what I report lest the details of the present situation prevent you from seeing what is more important, the probable shape of the future.

Language can be taken at many levels, according to the purposes of the taker. It is wholly natural that work on the more superficial levels be most common. Of that work, I shall say least--but I do not wish to neglect it entirely.

A system for linguistic computation is inevitably complex, particularly if the designers of the system intend it to penetrate to the deepest levels of language. Hence it is natural that much of the literature be devoted to the components--the working out of details. I shall say almost nothing about this literature, although the proper treatment of details is essential to satisfactory operation of any linguistic system.

My principal theme is the growth of a new conception of linguistics, which seems to be progressing more rapidly among developers of applied linguistic computing systems than among theoreticians. As we all know, the view that one practices an applied science by taking that science and applying it is not always valid. The science can appear as the result of applied research. Several continuing efforts give me reason to hope that applied linguistics will bring us important new ways of comprehending the human nature of language.

In closing, I shall discuss briefly some possible applications that, I think, have not hitherto been considered but which have at least this advantage: that they reveal the depth to which the computer, as a language processor, can become involved in human affairs. The propsect of ever-deepening collaboration between man and computer is the strongest motivating force we could have for continuing applied research; if the prospect is clear enough, economic arguments against present applications need not restrain us.

SURFACE SYSTEMS AND CLASSICAL SYSTEMS

A system works on the surface of language if it deals with language as strings of characters or phonemes. I see nothing improper in this manner of working; it is economical when it is effective.

In the early stages of grammatical study of a new language, concordances are extremely useful to linguists. The Summer Institute of Linguistics operates a system to make concordances of texts collected in Mexico (Grimes et al., 1968). The linguist takes a typewriter with him to the field, and sends typescript to the computing center. There the text is punched, computation is performed, and from there the concordance is mailed back to the linguist in the field. The shortest time reported for this cycle is 35 days; the average quoted is three months.

In the preparation of teaching materials, too, the concordance, the word list, and such tools are convenient. The Language Centre of the University of Essex (Haden and Kay, 1968) has collected spoken Russian, supplied rich grammatical annotations by hand, and is making concordances, etc.

Many of the typing errors in input--for whatever purpose text is punched--can be eliminated by automatically matching text words with dictionary entries in a way that does not rely on perfect spelling. Szanser (1969) reports work at the National Physical Laboratory (England) toward such matching procedures.

Work on the ancient and classic languages is in progress in many places; see the review by Waite (1968). Durham and Rogers (1969) use concordances, arranged in ways to suit phonological analysis, as tools in reconstruction of protolanguages in the Romance family.

At Zagreb, contrastive analysis of English and Serbo-Croat is proceeding with computer-made concordances. In part, the research uses tapes of English prepared at Brown University, with matched translations (Bujas, 1969).

Information storage and retrieval consists in gathering and keeping documents, accepting questions, and finding either answers or documents probably containing answers, in the collection. A typical plan, on receiving a document, is to form some short statement--in words or numbers--to represent it, then store the statement in a computer, the document on a shelf. When a question arrives, it is matched (perhaps after translation into appropriate terminology) with the statements. Documents corresponding to statements that match the question are delivered. There are systems in which reduction of documents to statements is done by persons; in others, computers do it. The reformulation of questions is also done both ways.

Last year, Gerard Salton published a major report on the SMART system for information retrieval which he has been developing for several years at

Harvard and Cornell (Salton, 1968). This very elaborate system provides for input of text, consultation of a dictionary, syntactic analysis, recognition of statistically significant phrases, consultation of a hierarchically-organized thesaurus of concepts, and matching of requests against stored representations of documents. As one reviewer noted, this system has the great advantage that any new element imaginable can be added to it and tested.

Nevertheless, Salton would, I think, agree that his system belongs among those that work at the surface. He did syntactic analysis of input and found it presently not as helpful in retrieval as the statistical methods otherwise used.

The SMART system is being redesigned for on-line operation. When the new version is ready, the user will be able to submit a search request, see the results (or a sample of them), and reformulate the request. The system allows this interaction, but on a longer time scale. The user, in fact, need only mark some of the documents offered him as relevant to his problem to improve the working of the system, which can then revise its selection criteria to give him more documents similar to the ones he prefers. An innovation is to regroup documents in the file on the basis of the users' relevance judgments as they accumulate.

An astonishing fact about information retrieval is the goodness of the effect that can be obtained by using only surface information. The Medlars system, operated by the U.S. National Library of Medicine, employs persons with special training both to index documents and to formulate search requests in the terms the system can use; Salton uses automatic procedures for both. Yet, in an experiment with 18 questions and 273 documents in biomedicine, the SMART system produced results comparable with those of the Medlars system: SMART found approximately the same number of truly relevant documents, but also selected slightly more irrelevant documents (Salton et al., 1968).

In yet another experiment, Salton and his colleagues employed a file consisting of 1095 English abstracts and 468 German abstracts. They used

48 questions written in English and translated into German. In the thesaurus, German words were added beside the English words attached to concepts. Then each question was converted by means of the thesaurus into a list of numbers, and each abstract was converted in the same way. Hence, a question could be matched with a document, no matter what the original language of either. The cross-language runs produced results about as good as those with questions and documents in the same language. However, runs against the German document collection were inferior, apparently for the simple reason that not as many German as English words could be matched; the thesaurus in English is fairly complete; that in German, just begun (Salton, 1969). Experiments of this kind are encouraging to those with the practical goal of making the world's scientific literature more widely available to scientists; the language barrier seems not quite so high.

It is possible now to draw diagrams--for example, maps--under the control of a computer. Francis, Svartvik, and Rubin (1969) are using this facility to plot maps of dialectal variation. The possibility of combining map making with statistical operations is fascinating, since it means reduced labor and therefore, probably, greater sophistication in future dialect surveys.

Files of text are part of the life of every scholar. Many systems have been developed to help the scholar--and also the lawyer and the business manager--to handle his files more conveniently. A system at the MITRE Corporation has many different capacities; some are surface-level: the user can look at his text on a cathode-ray tube, annotate lines or other parts of it, and create his own dictionary of synonyms as he goes. He can search the file, using as he sees fit the system's ability to extract stems from word forms (Walker, 1969).

The early machine-translation systems were surface systems, which would be no fault had they worked. Unhappily, they did not work to the satisfaction of all concerned. Rather early, a scheme was proposed that I consider classic; with variations, it is a scheme for translation, information retrieval, reading correspondence in an office, tutoring in schools, and so on.

The classic scheme, then, consists of surface processing, syntactic processing, and semantic processing. The object of these three steps, in order, is to obtain an expression for the content of the input text. Of course, the input text is itself an expression of its own content, but the expression that the classic system intends to obtain has such advantages as absence of ambiguity, perspicuity with respect to cognition, or even universality (independence of the original natural language).

Machine translation, classically, consists of the classic scheme run forward (analysis) from the input text, and then in reverse (synthesis) to obtain the output text. Abstracting might consist of the classic scheme, followed by logic to find the main points. For automatic tutoring, the classic scheme could be applied to a student's response, and logic applied to determine whether the response conformed to the predetermined answer required.

A system of the classic variety is being studied by Simmons (1968). He proposes to keep definitions in storage, and to construct deduction rules to determine whether the definitions of terms used by the student are equivalent to those of the system's terms. When the test is made, the organization of the response--the syntactic and semantic relations among its parts--is known. A deduction rule can specify those relations among several terms; it is not bound to one term at a time.

Rapidly, now, let us examine some other activities of this kind. Pratt and Pacak (1969) are working toward automatic analysis of pathologists' reports. They work on classic lines to get the site of a disease, the cause, the structural changes it has produced, and its physiological manifestations out of plain text.

Wu and Harper (1969) are developing a paragraph generator. Friedman programmed a system for generation of sentences with transformational grammars; she reports (1969) uses of it with three (modern) English grammars, one of French, one of Swahili, and one of Alfredian English. Some of these tests are small, but the main purpose is to find errors--from the most trivial to those that can be fixed only by revision of the theory--and both sorts have appeared. Vauquois and colleagues (1969) at Grenoble, working on

Russian-to-French translation, are far beyond syntactic analysis, and are engaged with design of a pivot language. Schank and Tesler (1969) describe a conceptual dependency parser. With a verb such as "go" they note that the verb may be accompanied by "with" and a noun naming any movable physical object. Arsent'eva, Balandina, and Krasovskaya (1969) report on computer programs to generate all Russian expressions from lexico-syntactic structures described earlier by A. K. Zholkovsky and I. A. Mel'chuk in several papers. Like Friedman's transformational system, this work seems at present most significant for its uses in research on the grammar of a natural language.

Smith (1969) has written a program to apply historical rules of phonological change. With 21 rules and 500 reconstructed Proto-Indo-European forms, he obtained some--but not nearly a majority--of the modern Russian forms he wanted. Klein, Kuppin and Meives(1969) are continuing Klein's work in the simulation of language history. Persons, represented in the simulation by their grammars, carry on conversations and alter their rules to parse what they hear.

A Kolloquium Maschinelle Sprachverarbeitung was held at Mannheim in 1968; more than a dozen papers were delivered, showing vigorous activity in West Germany in many areas of the classic paradigm (Übersetzerdienst der Bundeswehr, 1969). von Glasersfeld and Pisani (1968) report a new parser based on Ceccato's theory of language. Reich (1968) reports a simulator for the kind of network required in Lamb's newest version of stratificational grammar.

We could continue this tour d'horizon around the world, and it is unfair not to, but time is limited; I hope I have convinced you that the classic tradition maintains its vigor, and that the usefulness of surface approaches still commends them for limited purposes.

A NEW CONCEPTION

My definition of linguistics is neither one of the classic definitions, nor--as best I can judge--the Chomskian definition. For me, the first step is to think about information. The relation of information to language is intimate, although we cannot say that only language carries, supports, or conveys information. Paintings and sculpture are comprehended visually in a way that a page of printed text is not. To speak of a language of pictures is metaphorical. Even music, which is produced (in part) by the same vocal apparatus that utters speech, and is perceived by the ears and auditory tract that perceive speech, is only language-like, not language. There is information in visual and auditory images, and we work on that information without putting it into words. Nevertheless, the connection between information and language is intimate.

I am even tempted to say that the essential characteristic of language is its connection with information; in other words, that the primary function of language--which may have a dozen others--is to support, convey, etc., information. You are likely to misunderstand, and disbelieve, such a remark. The emotional life of man is perhaps biologically more important than his intellectual life; hence the emotive function is primary. Or, the solidarity of the social group is essential to human survival, and therefore the phatic, that all but vacuous function of language, is primary. And so on. But none of these claims--for which I do not vouch--contests the primacy of information in the sense I intend. Let me put it this way: if the emotive function is primary, it is the primary function of the human information system. Thus I put information in as part of the definition of language: Human language is the external manifestation of the uniquely human capacity to process symbols in such ways that information is retained, revised, and recounted.

Specialists in computation distinguish between analogue and digital computers. Within an analogue computer, something changes in proportion to each input datum; by coupling two of those somethings, and deriving a third, something can be made proportional to a sum, product, or other function of two inputs. Of course, a digital computer is not like that at all. A digital computer is a symbol manipulator, embodying rules formation and transformation of symbol sequences.

The human organism, I think, also has analogue and digital capacities; by symbol processing, I mean the work we do digitally. And I am tempted to assert that all digital processing in the human being is most intimately linked with language, and that analogue processing is linked with vision, audition, and other kinds of information.

Bear with me, now, while I inquire into the nature of the human information processor. One of its capacities is that of symbolic storage. We can, I believe, retain symbolic formulae. For example, we can memorize and recite poems; we can also understand a discourse, retain a summary of it, and retell it "in our own words." The words of the poem, in their fixed order, constitute a symbolic formula. The elements of the summary of a discourse are not, I think, words; perhaps they are sememes, or perhaps we should refrain from naming them until we know a little more about them. In some fixed order, which need not be sequential, a group of sememes is a symbolic formula.

A second capacity we have is to retain action patterns. The universally accepted illustration is riding a bicycle. Another is typing a word without thinking about its spelling. Another is deducing a conclusion from premises; every person can do it:

She: "Mother is coming to visit."

He: "I'll move my papers out of the guest room."

Each of us uses action patterns of this kind with such frequency and naturalness as not to notice them, and to believe that drawing conclusions from premises is something learned only by those who take academically oriented programs in schools.

Are action patterns stored digitally or analogously? I do not know; perhaps in both ways. The answer to this question has great interest. At any rate, action patterns are not necessarily simple; psychologists and linguists have considered how one action pattern can be constituted of others. A recursive system of action patterns is a powerful scheme for expounding the enormous behavioral capacity--infinitely flexible and creative--of the human being.

But for me the most remarkable capacity of the human organism is the capacity to convert a stored symbolic formula into an action pattern. For example, one person can teach another to ride a bicycle by telephone. Sometimes words fail us, and we discover that our language is not subtle and refined enough to describe an action pattern; then we can only perform the action and ask our pupils to imitate us. This is how pronunciation of a new language is taught, or high skill in musical performance, or ciruus acrobatics.

Among the action patterns, as I have said, there are symbol-processing patterns. And some of those patterns can be described in words, and even taught. Our capacity to process symbolic information--our faculté de langage--consists of these action patterns and the ability to store symbolic information.

Linguistics is the study of the faculté de langage.

This is my definition. Suppose, for example, that a linguist enunciates a rule of grammar. He gives it as a tagmemic formula, or a context-free rewrite rule, or a transformational rule, or a sign pattern (Lamb's term), or in some other form. Either he intends his rule to be the description of an action pattern, or--by my definition--he is not really doing linguistics. For an older generation of linguists, my position is psychologizing, and as anathematical as psychologizing in philosophy. For others, what I call for belongs to the distant future, and is as unrealistic as if, during the natural-history phase of biology, someone had called for molecular biology forthwith. For a few, my call is merely a trite repetition of their views.

A digital computer realizes rules of formation and transformation of symbol sequences. The realization may be by means of transistors, vacuum tubes, or little plastic devices through which puffs of air flow. For the designer of the computer, the choice does not matter. He can write logical formulae that describe the storage capacities of the machine, and its action patterns, without knowing the details of the realization. True, he may be able to arrange for more efficient or economical operation if he knows what realization will be used. However, the designer specifies certain details of the realization that the programmer does not. For the computer programmer, the machine is a device that can add, move information, make logical decisions and alter its course of operation accordingly. For the designer, the machine is a device that adds in a certain way, moves information according to certain routines, etc.

The linguist's rule of grammar can be understood as part of a program, in which case it is truly independent of the human realization of the faculté de langage, or it can be understood as part of a design, in which case the linguist is saying something about the realization--although he is not responsible for the details. Just as the designer can economize if he knows the details, however, the linguist can presumably come closer to the truth--at his own level--if he knows the details of biological realization.

Where do action patterns come from? Some say they are learned; some say they are inherited. Both, of course, are correct. The distinction is no doubt too clear and precise to do justice to the facts. One way to acquire a new pattern is to imitate; before that can happen, the pattern of imitation--including some patterns of perception and comparison--must be available. To escape this regress, we deliver the final stages into the hands of the phylogenists, who can always invoke chance mutations! Beyond whatever patterns are inherited, the three obvious ways to acquire patterns are imitation, being taught, and inventing. It seems probable to me that a great many patterns are reinvented by every organism that possesses them; they are easy to invent, the ones that take only a recombination of existing elements in a manner that the environment often causes.

One of the kinds of invention that Hockett has often mentioned is analogy. Invention of an action pattern by analogy is a kind of partial replication; here I have a pattern that works with certain inputs and gives certain outputs, and I copy it altering minor features to fit a different class of inputs, and get a different class of outputs. Oddly enough, several members of the transformationalist school find this concept odious. It seems to me altogether probable that linguistic invention by analogy happens with great frequency.

Most linguistic inventions, naturally, are discarded at once. Given a try, they prove noncommunicative--they are not understood. Of course, I am talking about action patterns, not new sentences. When I do put together a sentence, I may like it and retain its outline--some of its main features--as a symbolic formula. Later I can build another sentence on that model, and if I do so repeatedly--just as in learning to ride a bicycle by following instructions--I can acquire the model as an action pattern.

Why do I not call action patterns habits? Because, for one thing, habit has the connotation of simplicity; it is a unit that cannot be broken up, cannot grow complex. For another thing, it has the connotation of learning by repetition. Neither suits me.

The position I am upholding differs sharply from that of, for example, Katz. During the last decade, many linguists have worked on the assumption that a dictionary is a finite list of entries, each characterizing a word--or, more precisely, some underlying lexical unit. The characterization for each unit is a closed, even a small, object concocted of non-word-like elements. The problem of semantics, as seen by these linguistis, is to decide on a list, preferably short, of elements to be used in semantic characterizations in the dictionary (or lexicon), to formulate a grammar of lexical entries, and to propound one or more rules by which to create semantically acceptable sentences or to assess the semantic acceptability of a given sentence (of known syntactic structure).

But let us consider instead the possibility that a person's lexical entries are not independent of one another. As a first example, we can take the system described by M. Ross Quillian (1968; from his dissertation submitted in 1966). "The central question asked in this research has been: What constitutes a reasonable view of how semantic information is organized within a person's memory?" (Quillian, 1968, p. 216). Taking the ordinary dictionary as a starting point, Quillian constructed a computer system in which each definition was a unit that could be isolated, but each content word in a definition was only implicity present. What was actually present was a cross reference to the definition of the word required. If, for example, the definition of 'plant' includes the word 'living' then a pointer in the definition of 'plant' to the definition of 'live' implicitly includes the definition of the defining term in the definition of the defined. Does this not make the whole dictionary the definition of each word in it? It does; according to Quillian, "a word's full concept is defined in the memory model to be all the nodes that can be reached by an exhaustive tracing process . . ." (Quillian, 1968, p. 227). The meaning of a word is the speaker's knowledge of his language seen from the point of view of that word. The elements of a definition are content words and in addition something non-word-like; the relations among elements within a definition were not shown, in Quillian's computer model, as in a printed dictionary, by word order and other grammatical clues, but rather by specific relators of limited variety--although he did treat some prepositions as content words.

The current version of Quillian's system is called a Teachable Language Comprehender (TLC). In it, units and properties are distinguished. "A unit represents the memory's concept of some object, event, idea, assertion, etc. Thus a unit is used to represent anything which can be represented in English by a single word, a noun phrase, a sentence, or some longer body of text." (Quillian, 1969, p. 462). Every unit is identified in a special way as a particular kind of some other unit; the fact that a special provision is made for this relationship is striking, since it may not be the case that every concept is agreeably defined by giving genus and differentia. But Quillian does provide for the absence of a pointer to the genus. The differentia are given as properties of the unit; each property is stated in memory as a predication, and the predicates that can be used are those of the language itself.

Quillian gives the example of 'client': "a person by whom a professional is employed." Now, 'client' seems clearly to be a relational concept; one cannot be a client by virtue of one's own nature, but only by virtue of a relationship with a professional. The semantic memory of TLC provides a format, so that when a specific person is mentioned as client and another person, perhaps a lawyer, is mentioned as professional, TLC can understand the statement by fitting it to the format.

Given a text to understand, TLC attempts to build a model for its content, on the same general plan as its permanent memory. Working on the instructions of a person as tutor, TLC can add to its permanent memory what it learns by understanding a text. Quillian speaks (1969, p. 473) of an odd problem in that connection. Were it not for the tutor, TLC would not know where to store the new information it acquires: "is 'Battleships in World War I had 16-inch guns,' about battleships, about World War I, about guns, or about, perhaps, naval history? Or is it about all of these?" The answer would be all, except that knowledge in this model has to be organized under units; if every possible unit is taken for every sentence that comes along, all knowledge in TLC will be stored in multiple copies. What seems odd about this system at the moment is the strictness of the sense in which every fact has a place. We saw earlier that a client is a person who employs a professional; if we wanted to know immediately, when reading about a professional, that he is employed by clients, the fact would have to be stored, so to speak, inside out.

This oddness presumably arises out of the use of pointers. The fact that there is a sign in the bus station pointing to the airport does not help me find the bus station when I am at the airport, and the fact that "employs a professional" is stored at 'client' and points to the entry for 'professional' does not help when TLC consults the latter entry. Here the computer model is far short of human capacity, it seems, since I can remember that a thumb tack is a round thing almost, if not quite, as easily as I can remember that roundness is a property of thumb tacks.

Both of Quillian's models employ an exploratory operation in their networks. Start at any pair of nodes; go out one step from each of them on every possible path; mark the nodes you reach; now proceed from all of the nodes marked in the first step; and so on. This exploration could, of course, cover the entire memory; in fact, any two nodes selected because the words associated with them occur together in text are likely to be connected by short paths. When the search comes to a marked node, it has found a path from one of the original nodes to the other, and the search can perhaps stop.

In the new model, the search helps in such tasks as eliminating ambiguity. Suppose 'client' has more than one meaning; then it has two main nodes in the memory. If it occurs in context with 'lawyer', search starts at both nodes for 'lawyer'. The first node reached on a path from either meaning of 'client' and also on a path from 'lawyer' is likely to be 'professional'. The search stops, and the appropriate meaning of 'client' is chosen. The genus of both 'client' and 'professional' is presumably 'person'; tests prevent paths with wrong kinds of links from giving satisfaction is the search.

STUDENT, a computer system to understand and solve algebra word problems, was accepted as a thesis in 1964 (Bobrow, 1968). It used rewrite rules to parse the input (word problems), conversion rules to obtain equations, and an algebra routine to obtain the answer. This is an example of action patterns; the natural-language work was a small part of the system, and the knowledge that permitted good parsing was knowledge of algebra, not of ownership, apples, and oranges-- the concepts that arise in this context.

An early system of this type (the thesis describing it was accepted in 1964) was called SIR (Raphael, 1968). It stored items such as 'boy', 'hand', 'finger', and 'person'; and for each item it stored a list of properties, including set membership and part-whole relations. The program included functions to mark set inclusion, test set membership, and so on; these are simple examples of the action patterns I think are needed for verbal or nonverbal thought in human beings and in simulations of it by machine.

SIR had no way of acquiring a new action pattern, but obviously a programmer could write one and add it to the system. As a language in which the programs are written becomes increasingly like natural language, programming action patterns will increasingly resemble teaching a skill to a person. Learning from a teacher is, in many activities, the normal mode. As Minsky observes, "The usual objection is that 'it didn't learn it; it was told.' I hope the reader does not flatter himself by believing that he 'figured out' the difference between energy and momentum. He was told; Newton figured it out, Galileo didn't, and there was no one to tell him." (Minsky, 1968, p. 14). Yet to assume that programming will become more like teaching, without explicating understanding of natural language, is to beg the question. How a computer program can learn for itself, say by trial and error or by insight, that energy and momentum must be distinguished is a different, and inordinately more difficult question than how a computer program can learn by being taught to use certain formulae on certain occasions. The latter is a question that has to be asked in the near future.

Robert F. Simmons published a review paper on Natural Language Question Answering Systems: 1969 (Simmons, 1969), and I see no reason to copy his whole review here. He discusses half a dozen systems in some detail, and mentions others. He does not mention the BIT system, developed at Minsk by Skorokhod'ko et al., (1968), but this system might have fallen within his purview. Like the others, it includes syntactic and semantic interpreters of input text. Some of the systems reviewed by Simmons have stores in the form of networks, like the Quillian semantic memory. Techniques for matching a question with a stored fact include deduction, abstraction, and so on.

Schwarcz (1969) examines the problem of semantics in natural language with great care and insight. He observes, much to the point I wish to make, that 'people carry in their memories not one model but many, corresponding to the many different situations that they have knowledge of. Thus, a message must refer either to a specific model, to a specific range of models, or generically to all models in which the specified intensions have nonempty extensions." (Schwarcz, 1969, p. 9). By presuppositions, which can differ among words that otherwise have the same effect, and by including something given as well as something new in each statement, the speaker helps the listener find his place in memory. Schwarcz discusses linguists' theories of semantics, computer programs with significant semantic elements, and logico-theoretic problems. He gives enough specifications for natural-language processors to show how he would attempt to build one, and the reader has some confidence that the outcome would be worth the expense.

My optimism about applied computational linguistics rests, most firmly, on the attention several research groups are giving to the storage of encyclopaedic knowledge, and the definition of useful action patterns. I do not renounce surface systems, which are economical, if limited. The classic scheme, I now believe, can be revised to take better advantage of stored knowledge. The result, I think, will be reduced responsibility for linguistics - syntax and semantics - but only a reduction to such a point that linguistics, at last, becomes possible.

Likewise, a new kind of application should become possible. If we conceive of a computer in which facts are stored, and action patterns for matching new facts with old ones, a certain limited kind of conversation with the machine is possible. Add to that what Minsky calls physical theories, what Schwarcz calls models; a popular name is paradigm, my private term is anecdote; these are sketches, at varying levels of abstraction. Add, also, action patterns for matching new facts against old paradigms. At that stage, I think, real conversation is possible. The computer can help the user think through a problem by offering him paradigms to fit his problems. Perhaps, in 10 or 20 years, we shall find the computer as helpful as an imaginative consultant for a wide range of problems.

BIBLIOGRAPHY

Arsent'eva, N.G., N.A. Balandina, and A.I. Krasovskaya. O Mashinnoj Realizatsii Sistemy Perifrazirovaniya, Institut Prikladnoj Matematiki, Moscow, 1969.

Automatic Language Processing Advisory Committee. Language and machines computers in translation and linguistics. National Academy of Sciences - National Research Council, Washington, 1966.

Bobrow, Daniel G. Natural language input for a computer problem-solving system. In: Marvin Minsky, Ed. (1968), pp. 135-215.

Bujas, Z. Computers in the Yugoslav Serbo-Croat/English, Contrastive analysis project. ICCL ≠ 26.

Durham, S.P., D.E. Rogers. An Application of Computer Programming to the reconstruction of a proto-language. ICCL ≠ 5.

Francis, W. Nelson, Jan Svartvik, and Gerald M. Rubin. Computer-produced representation of dialectal variation: initial fricatives in Southern British English. ICCL ≠52.

Friedman, Joyce. Application of a computer system for transformation grammar. ICCL ≠14.

Grimes, Joseph E., James W. Sweeney, Robert W. Shields, Larry F. Core. Linguistic concordance system manual, University of Oklahoma, Norman, 1968.

Haden, R. and J.B. Kay. Contemporary Russian Language Analysis Project and Foreign Language Recordings Project, Directors' Reports for 1968. Language Centre, University of Essex, 1968.

ICCL. Preprints, International Conference on Computational Linguistics. KVAL, Stockholm, 1969.

Klein, Sheldon, Michael A. Kuppin, and Kirby A. Meives. Monte Carlo simulation of language change in Tikopia and Maori. ICCL ≠ 21.

Minsky, Marvin. Semantic information processing. The MIT Press, Cambridge, Mass., 1968.

Pratt, A.W. and M.G. Pacak. Automatic processing of medical English. ICCL ≠ 11.

Quillian, M. Ross. Semantic memory. In: Marvin Minsky, ed. (1968), pp. 216-270.

Quillian, M. Ross. The teachable language comprehender: a simulation program and theory of language. Communications of the ACM, 12:8, August 1969, 459-476.

Raphael, Bertram, SIR: semantic information retrieval. In: Marvin Minsky, ed. (1968), pp. 33-134.

Reich, Peter A. The relational network simulator. Report RNS. Linguistic Automation Project, Yale University, New Haven, 1968.

Salton, Gerard. Automatic Information organization and retrieval. McGraw-Hill, New York, 1968.

Salton, Gerard, et al. Reports on analysis, search and iterative retrieval. Scientific Report No. ISR-14. Department of Computer Science, Cornell University, 1968.

Salton, G. Automatic processing of foreign language documents. ICC ≠ 4.

Schank, Roger C., and Larry Tesler. A conceptual dependency parser for natural language. ICCL ≠ 2.

Schwarcz, Robert M. Towards a computational formalization of natural language semantics. ICCL ≠ 29.

Simmons, Robert F. Linguistic analysis of constructed student responses in CAI. Computation Center, The University of Texas at Austin, 1968.

Simmons, Robert F. Natural language question answering systems: 1969. Computation Center, The University of Texas, 1969.

Smith, Raoul N. Automatic simulation of historical change. ICCL ≠ 9.

Szanser, A.J. Automatic error-correction in natural languages. ICCL ≠ 7.

Ubersetzerdienst der Bundeswehr, Kolloquium Maschinelle Sprachverarbeitung (Mannheim, 26-27, November 1968), Mannheim, 1969.

Vauquois, B., G. Veillon, N. Nedobejkine, C. Bourguignon. Une notation des textes hors des contraintes morphologiques et syntaxiques de l'expression. ICCL ≠ 17.

von Glasersfeld, Ernst, and Pier Paolo Pisani. The multistore system, MP-2. Georgia Institute for Research, Athens, 1968.

Waite, Stephen V.F. Computers and the classics. Computers and the Humanities, 3 : 1, September 1968, 25-29.

Walker, Donald E. Computational linguistic techniques in an on-line system for textual analysis. ICCL ≠ 63.

Wu, Stanley S.Y., and Kenneth E. Harper. A directed random paragraph generator. ICCL ≠ 13.