NATURAL SEMANTICS IN ARTIFICIAL INTELLIGENCE

Jaime R. Carbonell
Allan M. Collina
Bolt Beranek and Newman Inc.
Cambridge, Massachusetts

Abstract

This paper discusses human semantic knowledge and processing in terms of the SCHOLAR system. In one major section we discuss the imprecision, the incompleteness, the open-endedness, and the uncertainty of people's knowledge. In the other major section we discuss strategies people use to make different types of deductive, negative, and functional inferences, and the way uncertainties combine in these inferences.

Keywords

Semantics, inference, cognitive processes, natural language processing, human memory, question-answering systems, deduction, analogy

1. Introduction

In this paper we will discuss how to represent and process information in a computer in ways that are natural to people. This does not mean doing away completely with representations and procedures which computers have traditionally used, but adding new representations and procedures which they have not used.

People often store and communicate imprecise, incomplete, and unquantified information; they often assert truth or falsity in relative terms; and they seldom seem to use rigorous logic in their inferential processes. Because of these conditions, people seem to have an almost infinite information processing capacity, with inference making and problem solving abilities more refined and far more flexible than any existing computer program.

How can we study these human capabilities in order to make our machines show similar performance? A combination of approaches is perhaps best. Observation of people's behavior, introspection, some experimentation, protocol analysis, and synthesis of computer programs can all be valuable techniques. A recent paper (Collins,Carbonell and Warnock[6]) discusses a technique for combining protocol analysis with program synthesis as applied to tutorial dialogues. The synthesis directs what to analyze, and the strategies observed in the analysis are evaluated by synthesis, in a kind of feedback loop. We have been using the SCHOLAR system in this way as a vehicle for experimentation with natural semantics.

Before we discuss some of the major problems in natural semantics, we will briefly describe the SCHOLAR system, since it is the environment for our research. A word of caution though: we are only trying to develop some insights, without attempting to be exhaustive. More questions will be raised than answers provided. There are many observable things people do that we do not know how to simulate.

2. The Scholar System as an Environment to Study Natural Semantics

In this section we will discuss, very briefly, some pertinent aspects of SCHOLAR, a mixed-initiative instructional system. More detailed discussions are provided in Carbonell[2,3,4] and Warnock, Carbonell, and Carbonell[14]. Two data bases are currently being developed: one is about the geography of South America, and the other is about the ARPA network. SCHOLAR'S knowledge about any subject matter is in the form of a static semantic network of facts, concepts, and procedures. This is a modified and extended network a la Quillian[12] and has a rich internal structure with a well-defined syntax.

Dialogue with SCHOLAR takes place in a subset of English that is limited mainly by SCHOLAR'S currently primitive syntactic capabilities. In tutorial fashion, the system uses its semantic network to generate the material it presents, the questions it asks, and the corrections it makes. At any time SCHOLAR accepts questions and requests from the student, and generates responses based on its semantic network, making calculations and inferences of different types that it has been programmed to handle. The dialogue is unanticipated, and is dependent on the student's responses, questions,and requests. Figure 1 presents a sample protocol of an interaction with SCHOLAR; the person types after the asterisk and returns control to the system with a second asterisk.

Figure 1
A Sample Dialogue Between
SCHOLAR and a Student

{Student inputs are enclosed by asterisks)

WHAT IS THE LANGUAGE OF ARGENTINA?

♦PROBABLY SPANISH*

VERY GOOD.

APPROX WHAT IS THE AREA OF ARGENTINA?

•TELL ME SOMETHING ABOUT PERU*

PERU IS A COUNTRY.
IT IS LOCATED IN SOUTH AMERICA.
THE CAPITAL OF PERU IS LIMA.

NOW ANSWER THE QUESTION YOU DIDN'T ANSWER BEFORE.

THE AREA OF ARGENTINA IS 1,100,000 SQUARE MILES

APPROXIMATELY TRUE, BUT A MORE EXACT VALUE IS 1,200,000.

THE EASTERN BORDERING COUNTRY OF PERU IS ----.

CATHOLICISM

WRONG. CATHOLICISM IS NOT A COUNTRY BUT A RELIGION. THE CORRECT ANSWER IS: BRAZIL

WHERE IS PERU?

IS ASUNCION HOT?

SINCE ASUNCION IS PART OF PARAGUAY, YES.

WHAT IS THE MOST IMPORTANT DIFFERENCE BETWEEN PERU AND PARAGUAY?

THE PRIMARY DIFFERENCE IS: THE POPULATION OF PERU IS LARGER BY A FACTOR OF 7.8.

Figure 2 shows some excerpts of SCHOLAR's semantic network. Properties, none of which are obligatory, can have as values a set of LISP atoms (usually English words defined elsewhere in the network), different types of lists, and other properties. Attributes are usually English words, but there is a set of special attributes for important relations, like SUPERC (for superconcept or superordinate, e.g., Lima is a city and a capital), SUPERP (for superpart, e.g., Lima is a part of Peru and South America), SUPERA (for superattribute, e.g., fertile refers to soil and soil refers to topography), APPLIED/TO (color applies to things, and capital to countries and states), case-structure attributes like agent and instrument (see Fillmore[8]), and various others.

Figure 2
Four Partial Entries from
SCHOLAR's Data Base

```
[RPAQQ CAPITAL (((CN CAPITAL))
        NIL
        (SUPERC NIL CITY
           (PLACE NIL (OF NIL GOVERNMENT)))
        (APPLIED/TO (I 4) COUNTRY STATE)
        (EXAMPLES (I 2) ($EOR BUENOS/AIRES
           LIMA MONTEVIDEO BRASILIA
           GEORGETOWN CARACAS BOGOTA
           QUITO SANTIAGO ASUNCION LA/PAZ
           WASHINGTON)
(RPAQQ FERTILE (((ADJ FERTILE)
        (CONTRA (BARREN)))
        NIL
        (SUPERA NIL SOIL)))
[RPAQQ PERU (((XN PERU))
        NIL
        (SUPERC NIL COUNTRY)
        (SUPERP (I 6)
           SOUTH/AMERICA)
        (LOCATION NIL
           (IN NIL (SOUTH/AMERICA NIL
              WESTERN))
           (ON NIL (COAST NIL (OF NIL
              PACIFIC))
           (LATITUDE (I 4)
                 (RANGE NIL -18 0))
           (LONGITUDE (I 5)
                 (RANGE NIL -82 -68))
           (BORDERING/COUNTRIES (I 1)
              (NORTHERN (I 1)
                 ($L COLUMBIA ECUADOR))
              (EASTERN (I 1)
                 BRAZIL)
              (SOUTHEASTERN (I 1)
                 BOLIVIA)
              (SOUTHERN (I 2)
                 CHILE))
        (POPULATION (I 2)
           (APPROX NIL 11 000000)
              (LANGUAGE (I 2)
               (($L PRINCIPAL OFFICIAL)
               NIL SPANISH)
               (INDIAN (I 2)
                 ($L QUECHUA AYMARA )))
        (CAPITAL (I 1)
           LIMA)
        (CITIES (I 2)
           (PRINCIPAL NIL ($L LIMA
              CALLAO ARAQUIPA TRUJILLO
              CHICLAYO CUZCO)))
(RPAQQ LIMA (((XN LIMA))
        NIL
        (SUPERC NIL CITY CAPITAL)
        (SUPERP (I 6)
           PERU SOUTH/AMERICA)
        (LOCATION NIL (IN NIL PERU)
           (NEAR (I 1)
              PACIFIC]
```

The entry for location under Peru in Figure 2 illustrates an important aspect of SCHOLAR's semantic network called embedding. Under the attribute location there is the value South America plus several subattributes among which is bordering countries. But under bordering countries there are subattributes like northern and eastern, some of which have several values. Embedding describes the ability to go down as deep as necessary to describe a property in more or less detail.

In the data base there are also tags, such as the Nil after location and the (I 1) after bordering countries. These tags are called irrelevancy or importance tags, and they vary from 0 (Nil) to 6. The lower the tag, the more important the piece of information is (except that I 6 is used to suppress printing, as with SUPERP). The tags add up as you go down through lower embedded levels. One of the ways SCHOLAR uses I-tags is to decide what is relevant to say at any given time.

In the rest of this paper, we will discuss how we are using SCHOLAR to cope with some of the problems in natural semantics. However, there are still many natural-semantics problems we have not touched.

3. Natural Semantic Information

In this section we discuss some aspects of natural semantic information and its relation to artificial intelligence.

3.1 Imprecision or Fuzziness

Imprecise language is an essential characteristic of human communication. As Lyons[10] says, "Far from being a defect as some philosophers have suggested, referential 'impreciseness'... makes language a more efficient means of communication." Talking about a tall person or a blue-green object does not require precise specification of height or spectral characteristics. The imprecision may occur either in communication or storage. If we say that a colleague receives a large salary, we may or may not know the figure.

SCHOLAR currently stores areas and populations in numerical form, but it can respond to the fuzzy question "Is Montevideo large?" with a pertinent answer like: "It is not one of the largest cities in South America, but it is the largest city in Uruguay". Here SCHOLAR has found two superparts, South America and Uruguay, and then compared Montevideo to other cities in each with respect to population.

However, it is more common for people to store values that are imprecise or 'fuzzy', what Zadeh[19] calls 'linguistic' variables. This is the case with values like 'large', 'red', 'hot', 'rich', etc. It seems to us that one must be able to store either precise values or fuzzy values interchangeably. (In fact, SCHOLAR has fuzzy values as well as precise values stored, e.g., that the Brazilian Highlands has a large population.) Furthermore, the procedures that act upon these values must be flexible enough to deal with either.

3.2 Incompleteness, Embedding, and Relevancy

Imprecise statements are often motivated by incomplete specification. Since all specifications can be refined, they are essentially incomplete. We store what is necessary, and even if we store more, we only communicate what is pertinent. SCHOLAR does this through its I-tags. If it is asked "Tell me about Peru", it only gives a few salient facts.

Further specification can be added by refining existing values. For example, instead of 'blue', we can have 'Navy blue', or 'quite dark Navy blue', etc. Further specification can also be added by giving new properties with attributes somewhat orthogonal to previous ones. An example of this is 'tall man' versus 'tall, heavy man wearing glasses'. Properties can be specified to any level of detail by embedding, an inherent quality of SCHOLAR-type semantic networks.

3,3 The Reference Problem and Context

Somewhat related to incompleteness and relevancy is the reference problem (see Olson[11]), Referring to a colleague, we may 'define' him as the father of Jack and Jill, or the author of that paper on self-referential statements, or the tall thin fellow with glasses. We decide on some specification depending on the context, including our assumptions about the person we are talking to. People usually specify only to the degree that is needed. In this sense, every partial specification is a •definition'.

The problem of context pervades natural semantics. Definitions and specifications, anaphoric references, what and how to answer, all depend on context. Furthermore, there usually co-exist a range of contexts from overall context to short-term running contexts. For example, at a given time, SCHOLAR may have the contexts South America, Argentina and Buenos Aires, each with some dynamically adjustable life. What is relevant at any given time depends on this contextual hierarchy,

A start toward making references specific to the listener is possible in a SCHOLAR type system by using I-tags (see Collins, Carbonell and Warnock[6]). The likelihood that another person will know about any concept is roughly proportional to the importance of the concept, as measured by the I-tags, with respect to the overall context. Therefore, it is possible to estimate the sophistication of a person based on the level of tags of the concepts he mentions in his conversation. This estimate than can influence the description one uses in referring to some concept. For example, to an unsophisticated listener one might refer to the "capital of Argentina" rather than "Buenos Aires", because the I-tags for the concepts "capital" and "Argentina" are lower than those for "Buenos Aires", as measured from a context such as geography.

In the future we want to have adjustable contexts in SCHOLAR, so that it can talk about the ARPA network, say, "from a communications point of view" to one person and "from a programming point of view" to another person. What this entails is a temporary alteration of the relative values of 1-taqs throughout the semantic network. Those concepts that are referred to under the concept "communication" (such as message capacity, bit-rate, etc.) should be temporarily increased in importance wherever they occur in the data base, for the person interested in communication. A correspondinq change must be made for the person interested in programming or any other concept or set of concepts. This kind of sensitivity to the interests and background of the person, and the kind of sensitivity (described above) to the sophistication of the person may be the two major elements in the way people adapt what they say to the listener.

3.4 Closed verauB Open Worlds

In some realms of discourse such as an airline reservations system (Woods [17]), a blocks world (Winograd[15]), or a lunar rocks catalogue (Woods, Kaplan, and Nash-Webber***}, there is a closed set of objects, attributes, and values to deal with. However, in most real world domains such as those faced by SIR (Raphael[13]), TLC (Quillian[12]), or SCHOLAR (Carbonell[2]), there are open sets of objects, attributes, and values. It turns out that the procedures and even the rules of inference that can be applied are different in closed and open worlds.

The distinction between closed and open sets is one of exhaustiveness and not one of size. For example, the set of states (e.g., Iowa), which is a closed set for most people, is probably larger than the set of cattle breeds (e.g., Holstein), which is an open set. However, open sets tend to be larger in general than closed sets.

The distinction is important in a variety of ways. For example, if there are no basaltic rocks stored in a closed data base, then it makes sense to say "No" to the question "Were any basaltic rocks brought back?" But if no volcanoes are stored for the U.S., it does not follow that the answer should be "No" to the question "Are there any volcanoes in the U.S.?" A more appropriate answer is "I don't know". Furthermore, it makes sense to ask what the smallest block in a scene is or the rock with least aluminum concentration, but it makes no sense to ask what is the smallest city in Brazil or the least famous lawyer in the U.S. It would be an approriate strategy for deciding how many flights from Boston to Chicago are nonstop, to consider each flight and count how many make 0 stops. But it would not be an appropriate strategy to consider each person stored in a limited data base (such as humans have),in order to answer the question "How many people in the U.S. are over 30 years old?" Within open worlds there are closed sets, so that a question like "How many states are on the Pacific?" makes sense whereas "How many cities are on the Pacific?" does not. SCHOLAR deals with this by distinguishing exhaustive sets from non-exhaustive sets.

We will discuss in Section 4 how SCHOLAR begins to deal with open world semantics. The essential point here is that the well-defined procedures that are appropriate for a closed world simply do not carry over to an open world. Unfortunately, most of human knowledge is open-ended, and so people have complex strategies for dealing with uncertainty and facing problems such as how to apply new attributes or values to objects where they haven't applied in the past.

3.5 The True-False Dichotomy and Quantification

The two-valued logic that underlies the propositional calculus and related approaches to inference cannot encompass natural semantics. The trouble arises because truth varies in degree, in time, in range, in certainty, and in point of view of the observer, when it is applied to real-world objects. We will briefly examine some of the implications of the multivalued nature of truth for natural semantics.

Symbolic logic uses quantification to distinguish between the universal and the particular, e.g., between "All men are mortal" and "Some men have warts". But there is no allowance made for the degrees of truth as between say "Some men have warts" and "Some men have ears", even though only a fraction have warts and almost all have ears. People will infer that Newton had ears (given no information to the contrary as with Van Gogh), but will not infer that Newton had warts. The inference in the former case treats the particular like the universal, because almost all men have ears. The more generally true a statement is, the more certainty people assign to such an inference. There just are not many universal truths to be found out in the cold, cruel world, and so people make the best of it.

Degree of truth varies not only with respect to fu2zy variables (see Section 3.1) and quantification, but also in other respects. The sky is blue, but not all the time. The yellow of a lemon is less variable than the yellow of corn, which sometimes borders on white. Boston is cold in the winter, but it is not so cold from the point of view of an Eskimo. Nixon told us that he didn't know about the cover-up of Watergate, but one is only more or less certain that he didn't know. What these examples are designed to show is that people are uncertain about the truth of any proposition for a variety of reasons. Sometimes people seem to merge all the many sources of uncertainty together, but sometimes they can distinguish different aspects of their uncertainty with respect to a single proposition.

SCHOLAR does not now have any means for representing uncertainty, but the natural way to add such information is in tags stored along with the I-tags. Just as with I-tags, U-tags can apply at all embedded levels of the data base. Because we have started on programming uncertain inferences (discussed below), it has become desirable to represent the underlying uncertainty in the data base as well, in order to evaluate how certain any inference may be.

4. Natural Inferences

We classify human semantic inferences into four major types: deductive, negative, functional, and inductive inferences. The various types are discussed in somewhat greater detail in Collins and Quillian[7] and Collins, Carbonell, and Warnock[5]. We do not argue that these describe all the inferential strategies that people use, but only some of the major varieties. Each of the different strategies described is being implemented as a specific subroutine in SCHOLAR to work on either the geography data base or the ARPA network data base. While we think that people have a large set of such strategies, the number is probably less than one hundred. Therefore, despite the inelegance of such an approach, we do not regard it as an endless task to encompass the bag of inferential tricks a person uses.

In Figure 3 we have included excerpts from tape-recorded dialogues between human tutors and students to illustrate some of the more complicated strategies people use, and the ways they combine together. We will discuss the examples individually below.

4.1 Deductive Inferences

There are several transitive relations that people use frequently to infer that a property of one thing may be a property of the other. These include superordinate, superpart, similarity, proximity, subordinate, and subpart relations.

Of the above types SCHOLAR now handles only superordinate and superpart inferences, which are the most common. For example, if asked "Does the Llanos have a rainy season?", SCHOLAR will first look under Llanos and failing to find the information there, will look under Llanos' SUPERC {for superordinate), which is savanna, and its SUPERP (for superpart), which is Venezuela and Colombia. A rainy season is a property of savannas and so the superordinate inference provides the answer. The superpart inference is less general because it is restricted to certain attributes such as climate, language, and topography. One would not want to conclude that the capital of Massachusetts is Washington D.C., just because Massachusetts is part of the United States. Because most properties of a superordinate or superpart are only generally true, and not universally true, exceptions must be stored to preclude an incorrect inference (Raphael[13]).

Similarity and proximity inferences parallel the superordinate and superpart inferences, but they carry less certainty. An example of a person using a proximity inference is shown in the latter part of the tutor's response in Example 1 of Figure 3. The tutor first said that a savanna could not be used for growing coffee, but then he backed off this conclusion because of the proximity of the large Brazilian savanna to the coffee-growing region there. To illustrate a similarity inference: if one knows a wallaby is like a kangaroo, only smaller, then one will infer that a wallaby probably has a pouch. We plan to add similarity information to SCHOLAR in the near future, because it will also be useful in making functional analogies which are discussed below. The recently added map facility (Warnock, Carbonell, and Carbonell[14]) which ties together visual and semantic representations, makes proximity inferences possible, but they are still a way off.

Figure 3

Tutor-Student Dialogue Excerpts

(T) There is some jungle in here (points to Venezuela) but this breaks into a savanna around the Orinoco.

(S) Oh right, that is where they grow the coffee up there?

(T) I don't think that the savanna is used for growing coffee. The trouble is the savanna has a rainy season and you can't count on rain in general. But I don't know. This area around Sao Paulo is coffee region, and it is sort of getting into the savanna region there.

(S) Are there any other areas where oil is found other than Venezuela?

(T) Not particularly. There is some oil offshore there but in general oil comes from Venezuela. Venezuela is the only one that's making any money in oil.

(S) Is the Chaco the cattle country? I know the cattle country is down there.

(T) I think it's more sheep country. It's like western Texas so in some sense I guess it's cattle country.

(T) And the northern part of Argentina has a large sort of semi-arid plain that extends into Paraguay. And that's a plains area that is relatively unpopulated.

(S) why?

(T) Because it's pretty dry.

Subordinate and subpart inferences follow a somewhat different pattern from the others discussed. If asked whether South America produces any oil, a person will answer "Yes" because Venezuela, which is part of South America, produces oil. But one does not want to conclude that South America is hot because the Amazon jungle is. We haven't worked out the details of the restrictions on these inferences as yet.

There are other transitive relations that are used to make deductive inferences but they are not as prevalent as the ones outlined here.

4.2 Negative Inferences

Negative information, such as the fact that men do not have wheels, is not usually stored but rather inferred. In a closed world this presents no problem; it is reasonable to assume that if something is not stored, then it is not true. In fact, SCHOLAR currently would say "No" if asked "Is oil a product of Brazil?" just because oil isn't stored for Brazil. But in the real world, the fact that something is not stored does not necessarily mean that it is not true. People seem to have complex strategies for deciding when to say "No" and when to say "I don't know". We are currently trying to develop these in SCHOLAR.

One kind of negative inference now in SCHOLAR is a simple contradiction procedure. It relies on contradictory values stored with various concepts: for example, barren contradicts fertile, and demoncracy contradicts dictatorship. Suppose SCHOLAR is asked "Is the Pampas barren?" It would find the soil of the Pampas is fertile, and since fertile contradicts barren, it would say "No, The soil of the Pampas is fertile."

There is an important class of contradictions that are not subsumed under the procedure above. For example, consider the question "Is Buenos Aires a city in Brazil?" The fact that Buenos Aires is not among the cities of Brazil is no reason to say "No", because there are cities in Brazil, such as Corumba, which are not stored. But there are three facts that together make a contradiction possible: (1) Buenos Aires is located in Argentina, (2) cities only have one location, and (3) Argentina and Brazil are mutually exclusive. We can illustrate the necessity for conditions (2) and (3): (2) even though Portuguese is the language of Portugal, it is also the language of Brazil (i.e., language can have more than one location); (3) even though Sao Paulo is in South America, it is also in Brazil (i.e., South America and Brazil are not mutually exclusive). Making an incorrect negative inference about cities with more than one location (e.g., Kansas City) or different cities with the same name (Rome, New York and Rome, Italy) is precluded by storing both locations specifically, just as with deductive inferences. The strategy we have worked out in flow chart form to find different contradictions of this kind is fairly complex.

Failure to find a contradiction leads to another kind of negative inference people use which we call the lack-of-knowledge inference (Collins, Carbonell and Warnock[5]). Example 2 of Figure 3 shows the tutor using this strategy. The basis of the tutor's inference is this: since he knows as much about other South American countries as he knows about Venezuela, it is a plausible but uncertain inference that if other countries produced oil, he would know about it. (His conclusion was at least somewhat wrong, because there are in fact several other countries in South America that produce oil, though for those countries oil is not nearly so important as it is for Venezuela.)

Such a strategy is currently being implemented in SCHOLAR in the following way: If asked a question like "Is oil a product of Uruguay?" where no oil is stored, SCHOLAR can look for oil under similar objects (e.q., Venezuela or Brazil) or objects with the same SUPERC and SUPERP. If SCHOLAR finds oil stored with Venezuela (say with an I-tag of 3) and if it has enough information stored about Uruguay (up to an I-tag of 8, say) to know about oil if it were at all important, then it can infer that Uruguay probably has no oil. The degree of certainty expressed in the answer should depend on the difference in I-tags between the depth of what it knows about Uruguay and the level at which oil is stored with similar objects. If SCHOLAR can find no similar objects that have the property in question, as with "Is sand a product of Uruguay?" the appropriate answer is something like "I don't know whether sand is a product of any country in South America". The lack-of-knowledge inference is based on the assumption that the extent of one's knowledge is fairly uniform for similar objects.

4.3 Functional Inferences

Functional inferences are common in the dialogues we collected (Collins, Carbonell and Warnock[6]). Examples 1, 3, and 4 in Figure 3 illustrate the three different ways we have seen people use functional knowledge: in quasi-calculations, in analogies, and in answer to "why" questions.

Functional knowledge, which includes knowledge about functional determinants and their interactions, is learned, just as is factual knowledge, and therefore is stored in SCHOLAR'S data base under concepts such as climate or agricultural products. We would argue that the representation of functional knowledge should be in a form that different procedures can use. One problem is to find a way to represent such knowledge in SCHOLAR so that it can be more or leas precise, and still be accessible to different subroutines that infer answers to questions or that describe the functional relation to students.

Functional calculations can be used in both a positive and negative way. One simple positive function now in SCHOLAR calculates the climate of a place if the information is not stored. Based on the functional determinants of climate, which are altitude, latitude, and distance from the sea, SCHOLAR will infer whether the climate is tropical, sub-tropical, temperate, or cold/polar. A negative use of calculation based on the agricultural products function is shown in the first part of the tutor's answer in Example 1. The functional determinants of agricultural products include the climate, soil, and rainfall. The tutor picked the lack of rain as a basis for a tentative "No". Negative calculations do not require as precise knowledge as positive calculations. They usually only require that one of the functional determinants have an <u>inappropriate</u> value.

Like functional calculations, functional analogies can be positive or negative. Example 3 shows the tutor making a positive functional analogy, again with the agricultural products function. There he thought of a region, western Texas, that matched the Chaco in terms of climate and rainfall, the functional determinants of cattle raising. Since he knew that western Texas was cattle country he inferred that the Chaco might be as well. A negative functional analogy might have occurred if the student had asked whether the Chaco produced rubber. Since the Amazon jungle and Indonesia produce rubber, the tutor could have said "No" on the basis of the mismatch between the Chaco and those regions, with respect to climate and rainfall.

A positive and negative analogy subroutine for SCHOLAR has recently been completed. It is a fallback strategy to be used if there is not enough information stored to calculate the functional relationship. For a functional analogy it is only necessary to know the functionally relevant attributes and their relative importance. Then SCHOLAR looks to see if it knows any similar objects where the property in question is in fact stored. It tries to find a match or a mismatch by comparing the given object and the similar object with respect to their values on the functionally relevant attributes. People frequently use such analogical reasoning, probably because of the ill-defined nature of their knowledge about functional relations.

The last example in Figure 3 shows the use of a functional relation to answer a "Why" question. The population density of a place depends on an indefinite set of functional determinants: climate, soil, and rainfall are major ones but distance from the sea, the particular continent, presence of valuable minerals, all contribute in different ways. The tutor picked one determinant that had a value inappropriate for a large population density and gave that as a reason. By contrast a geographer could probably write a whole treatise on why the Chaco has a low population density. What we aspire for SCHOLAR to do is what the tutor did, that is, to pick one or two of the major determinants with appropriate values and give those as a reason.

4.4 <u>Inductive Inferences</u>

We mention inductive inferences here only because they are a major class of human inference. We have not yet tried to program them in SCHOLAR since they occur mostly in storing rather than retrieving information. The generalization and discrimination processes underlying induction have been discussed in detail elsewhere (Becker[1]; Winston[16]; Collins and Quillian[7]).

4-5 <u>Combining Inferences and Accumulating; Uncertainty</u>

The inferential processes described can combine in a variety of ways. For instance, contradictions can combine with deductive inferences. SCHOLAR will answer a question like "Is the Atlantic orange?" with "No, it is blue", because it finds blue is stored with the SUPERC, ocean. Also one functional inference may call another. If the agricultural products function needs a value for the climate of some region, it could call the climate function to compute it.

A more important way that inferences combine shows up when different strategies reach independent conclusions about the same question. A good example is Example 1 in Figure 3. There a negative functional inference, with an implicit lack-of-knowledge inference, first led to a tentative "No" answer, but then a proximity inference produced a possible "Yes" answer, and so the tutor backed off his earlier "No". When several inferences combine to yield the same conclusion, they increase the certainty of the answer, and when they produce opposite conclusions, they decrease the certainty.

There are a number of sources of uncertainty in inferential procedures. Uncertainty can derive from the size of the difference between I-tags in the lack-of-knowladge inference, it can derive from the degree of match or mismatch in a functional analogy, it can derive from the degree of predictiveness of the functional determinants, and as we discussed earlier, it can derive from the degree of certainty about the information stored. These sources of uncertainty may be combined to produce an overall uncertainty (see for example Kling[9]). This overall uncertainty is important so that long.

tenuous chains of reasoning are not pursued to their pointless end, and so that the degree of uncertainty in the answer can be indicated to the student.

5. Conclusions

What we have tried to show in this paper is the fuzzy, ill-defined, uncertain nature of much of human knowledge and thinking. We want SCHOLAR to be just as fuzzy-thinking as we are.

6. Acknowledgements

This paper was started by Jaime R. Carbonell who died suddenly February 2, 1973. I have completed it as best I could following his outline. I want to thank Eleanor H. Warnock who helped me with the editing and Daniel G. Bobrow and Ross Quillian who have contributed many ideas to our work. The programming of various subroutines described in the paper was done by Jaime G. Carbonell, Mafk L. Miller, Joseph J. Passafiume, and Eleanor H. Warnock. Allan M. Collins.

This research was supported in part by the Office of Naval Research, Information Systems, under Contract No. N00014-70-C-0264, and also in part by the Office of Naval Research, Personnel and Training, under Contract No. N00014-71-C-0228, and by the Air Force Systems Command, Electronic Systems Division, under Contract Ho. F19628-72-C-0163.

7. References

1. J.D. Becker, "An Information Processing Model of Intermediate-Level Cognition," Bolt Beranek and Newman Inc., Technical Report No. 2335 (March 1972).

2. J.R. Carbonell, "Mixed-Initiative Man-Computer Instructional Dialogues," Ph.D. Thesis, M.I.T., Dept. of Electrical Engineering (June 1970).

3. J.R. Carbonell, "A.I. in C.A.I.: An Artificial-intelligence Approach to Computer-Assisted Instruction," IEEE Trans, on Man-Machine Systems, Vol. MMS-11, No. 4 (December 1970).

4. J.R. Carbonell, "Artificial Intelligence and Large Interactive Man-Computer Systems," Proceedings of 1971 IEEE Systems, Man, and Cybernetics Conference, Anaheim, California (October 1971).

5. A.M. Collins, J.R. Carbonell, and E.H. Warnock,"Semantic Inferential Processing by Computer," Proceedings of the International Congress of Cybernetics and Systems, Oxford, England (August 1972).

6. A.M. Collins, J.R. Carbonell, and E.H. Warnock, "Analysis and Synthesis of Tutorial Dialogues", Bolt Beranek and Newman Inc., Technical Report No. A.I.I (March 1973) .

7. A.M. Collins and M.R. Ouillian, "How to Make a Language User", in E. Tulving and W. Donaldson (eds.) Organization of Memory, Academic Press, New York (1972).

8. C. Fillmore, "The Case for Case" in Bach and Harms (eds.), Universale in Linguistic Theory. Holt, Rinehart, and Winston N.Y. (1968) .

9. R. Klinq, "FUZZY PLANNER; Computing Inexactness in a Procedural Problem Solving Language". Univ. of Wisconsin Technical Report No. 168 (February 1973).

10. J.Lyons, Introduction to Theoretical Lin?uistics, Cambridge University Press, ?ambridge, England (1968).

11. D.R, Olson, "Language and Thouqht: Aspects of a Cognitive Theory of Semantics", Psych. Review, Vol. 77, No. 4 (July 1970).

12. M.R. Quillian, "The Teachable Lanquage Comprehender: A Simulation Program and Theory of Language", CACM, Vol. 12, NO. 8 (August 1969).

13. B. Raphael, "SIR: A Computer Program for Semantic Information Retrieval", in M.L. Minsky (ed.) Semantic Information Processing, M.I.T. Press, Cambridge, Mass. (1968).

14. E.H. Warnock, J.R. Carbonell, J.G. Carbonell, "Discussing Maps with a Computer" (In preparation).

15. T. Winograd, "Procedures as a Representation for Data in a Computer Program for Understanding Natural Language", Project MAC TR-84, M.I.T. (February 1971).

16. P.H. Winston, "Learning Structural Descriptions from Examples", Project MAC TR-76, M.I.T, (September 1970).

17. W.A. Woods, "Semantics for a Ouestion-Answering System", Aiken Computation Laboratory, Report No. NSF-19, Harvard University (August 1967).

18. W.A. Woods, R.M. Kaplan, and E. Nash-Webber, "The Lunar Sciences Natural Language Information System", Bolt Beranek and Newman Inc., Report No. 2378 (June 1972).

19. L.A. Zadeh, "Outlining of a New Approach to the Analysis of Complex Systems and Decision Processes", IEEE Trans, on Systems, Man, and Cybernetics, Vol. SMC-3, No. 1 (January 1973).