Memory and Inference

Robert Wilensky
Division of Computer Science
Department of EECS
University of California, Berkeley
Berkeley, CA. 04720

ABSTRACT

A set of general knowledge structure manipulation primitives is proposed. These processing elements largely eliminate the need for a separate theory of inference. The elements are presented in terms of an abstract characterixation of the understanding task. The processing theory is given as a particular realisation of this characterixation.

1. Introduction

It has been generally recognized that world knowledge is important for making inferences, and that the structure and organization of knowledge in memory has important consequences in determining what inferences are made. For example, Schank and Abelson (1977), Charniak (1078), Colling ford (1078), Wilensky (1078), Lehnert (1082), and Dyer (1082) are all describe systems that use a particular kind of knowledge to make particular classes of inference.

However, the overall design the inference engine for text understanding remains obscure. Most of research has focused on the nature of the knowledge needed. The theory of processing per se has generally been relegated a secondary status, and usually developed only as far as was necessary to demonstrate the efficacy of one's knowledge structures.

The view taken here is that adequate theories of memory and the structure of knowledge largely eliminates the need for a theory of inference. In this formulation, what gets inferred is a function of what gets accessed in memory. This, in turn, is determined by the structure of knowledge, and by general memory access routines that are sensitive only to the structure and organization of knowledge, not to its content.

This view is parsimonious to the one espoused by Quillian (1068). Quillian's notion of representation was primitive by today's standards, however, and he was not concerned with the notion of inference per se. A more sophisticated view of inference is found in Charniak (1082). His approach is similar to the one pursued here, in that both theories endorse highly declarative representations and fairly genera] schemes for manipulating them. However, I am concerned less with the details of the implementation, and more with the overall architecture of the understander.

Is this paper, I first present an abstract characterization of the understanding task. Then a particular realization for this characterization is given.

In this discussion I refer to the representation produced by the story understander as a *construst* of the input. I call the process of deciding to include a particular fact in a construal *determinstion.* In addition, I shall use the term *assumption* to refer to the kinds of inferences used in story understanding, to distinguish these from the more general collection of processes that include logical deduction.

2. Principles of Text Comprehension

I first attempt to characterize the text understanding process by characterizing those facts whose determination constitute a satisfactory construal of the input. This characterization is given in terms of a set of principles. These are as follows.

> The Principle of Consistent Construal - View the inputs as instantiating a set of frames that are consistent with one's world knowledge.
>
> The Principle of Concretion (Overcommitment) - Determine as specific a fact as is possible from an input.
>
> The Principle of Least Energy - Make only the minimal assumptions necessary to determine a fact.
>
> The Principle of Exhaustion - Determine enough facts to account for all of the input.
>
> The Principle of Poignancy - Determine those frames that are pertinent to the intent of the discourse.

The Principle of Consistent Construal states that, as elements of a text are viewed as instantiating general frames, one should favor interpreting inputs as instantiating those frames that provide a consistent construal of the input.

For example, suppose we were given the sentence "The bottle fell off the table and it broke." One construal of this sentence assumes that (1) the bottle hit the floor, (2) it was the bottle (and not the table) thai broke, (3) the bottle was somewhat brittle, and (4) the bottle's hitting the floor, together with the bottle's brittleness, was the cause of the breakage. The Principle of Consistent Construal approves of this construal because it relates individuals in the story (i. e., the falling event, the breaking event, and the bottle) to knowledge in long term memory in a consistent manner.

However, the Principle of Consistent Construal does not inform us how to decide among alternative construals. For example, another construal of our example sentence assumes someone caught the falling bottle mentioned above and then broke it with a hammer. However, this construal seems more ad hoc. We require a way of distinguishing among such alternatives.

To address this problem, we introduce the Principle of Least Energy. Least Energy states that one should make those decisions that require the most tenable assumptions possible. Least Energy is meant be quantify the notion of the most tenable set of assumptions. Here we assume that facts are determined when they are supported by a sufficient amount of evidence. A fact that is supported by some evidence, but which is not yet determined, is said to be *activated.* In our processing model, when the activation of a fact reaches a certain threshold, then the fact will be determined. In addition, we assume that activation is spread to "nearby" facto in memory (cf. Anderson and Bower, 1073).

Our theory assumes that the introduction and determination of a frame or instance of a frame requires a certain amount of activation "energy". At any given point, various "activated frames" (i. e., those whose activation level is greater than zero) may have different levels of activation, and therefore some of these frames require more additional energy for determination. The Principle of Least Energy informs us that, when a choice is possible, we should determine those frames requiring the least amount of additional energy.

Least Energy allows for context sensitivity, as previous events may change the activation energy of activated frames. In addition, it allows inherent predispositions into the system. For example, if we mentioned that John ate a hamburger, we would probably assume that it was a normal-sized hamburger on a bun, etc. This assumption is made because the frame describing a prototypical hamburger is given a disproportionately large initial bias.

Note, however, that so far we have no reason to promote a specific concept (such as a prototypical hamburger) when only a general concept is mentioned. The Principle of Concretion is necessary to supply this direction. This principle instructs the system to find the most specific frame consistent with the input. This is derived from what I have termed "the First Law of Knowledge Application," which instructs us always to employ the most specific applicable knowledge. For example, if the understander learns that John loves Mary, it needs to realize that this is romantic love rather than sibling love in order to make accurate predictions about John's behavior. Thus a form of assumption b required in which semantic knowledge from the sentence (e. g., that some sort of love relation exists) is compared against pragmatic knowledge of relationship types (i. e., the various kinds of love), to find the interpretation that the speaker most likely intends.

The Principle of Exhaustion insures that an input is exploited fully in the understanding process. For example, consider computing an explanation for the sentence "John sneaked over to the door." One explanation for this action is that John wanted to be near the door. Indeed, if the understander knew that John wanted to leave the room, such an explanation is likely to be determined. However, it is important for the understander to realize that this explanation is not sufficient. The deficiency is due to the fact that the explanation explains only the change-of-location part of the action, but it does not explain the manner in which it was accomplished, that is, why John sneaked rather than just walked. The Principle of Exhaustion declares this an insufficient result as it leaves an unexplained residue.

The Principle of Poignancy distinguishes those aspects of the text that represent its essential content from those that are peripheral. Points are represented as rather abstract frames to which a certain affective importance is attached. As with any other invoked frame, such frames may be determined if enough evidence can be found. The resulting determined, elaborated frames constitutes the essense of a text.

3. Implementation

FAUSTUS (Frame Activated Unified Story Understanding System) is an frame manipulation program created by Peter Norvig that is designed to embody the principles suggested above. The idea behind FAUSTUS is that the program should make assumptions based not on specific rules of inference, but on general mechanisms for memory manipulation and on the structure of knowledge in memory.

FAUSTUS makes a number of kinds of assumptions, including the construction of explanations and the determination of story points in accordance with the principles specified above, by using the following set of fact manipulation primitives:

1) Invocation - Initially considering a fact. Facts are currently invoked because they have been indexed under a component that occurs as an input, or because they are explicitly associated with another invoked fact.
2) Determination - Deciding if enough evidence for an invoked fact exists to infer the fact.
3) Instantiation - Creating an instance of a determined frame to represent its occurrence in a story.
4) Elaboration - Filling in an empty slot of an instantiated frame. This process corresponds to the top-down processing performed extensively by systems like SAM and Ms. Malaprop.
5) Termination - Realizing that an instantiated frame is no longer relevant.
6) Attrition - Allowing an invoked but undetermined fact to fade from subsequence consideration. FAUSTUS measures time in terms of the number of fact invocations that have passed. Thus if enough new facts are invoked before a previously invoked fact is determined, FAUSTUS will drop that fact from subsequent consideration.

When FAUSTUS has an input, it first looks at instantiated frames to see if the input elaborates one of them, and then looks for frames that are indexed under the input. For example, frames suggesting the cause of an event are likely to be invoked by this process. FAUSTUS's determination process then looks for verification of these frames. Supporting evidence includes the previous mention of one of these frames or of their constituents. Thus FAUSTUS checks the story representation and other invoked frames for corroborating items.

FAUSTUS allows invocation to spread along the constituents of invoked frames. For example, if FAUSTUS knew John want to be at a particular restaurant, and stored in the permanent data base b the fact that being at a restaurant b a precondition for eating at one, then this stage would invoke eating at that restaurant as a candidate action. In effect, FAUSTUS will have hypothesized the explanation that John intended to eat at the particular restaurant he wanted to be at.

All the facto invoked in thb manner are now subject to confirmation. If one can be determined, the others will be either eliminated by exclusion or removed soon by attrition. If not, invocation will spread out again from the invoked facto. That is, if a reasonable explanation can be found within the story, it will be (a long, tenuous explanation may not be found due to attrition). If not, then intermediate hypothesises are *suggested* until a connection with the story can be found.

4. FAUSTUS Examples

Norvig's FAUSTUS implementation uses the PHRAN natural language analyzer, written by Yigal Arena, as its front end, and the PHRED generation system, written by Steve Upstill and Paul Jacobs, for output (see Wilensky and Arens, 1080). These stories demonstrates some general features of the understanding process.

INPUT TEXT:
Ann got into her car.
She went for a drive.
She arrived at the movie theatre.
She got a ticket.

INPUT: Why did Ann get a ticket?
OUTPUT: She wanted to see the movie.

INPUT TEXT:
Ann got into her car.
She went for a drive.
She got a ticket.
INPUT: Why did Ann get a ticket?
OUTPUT: She violated a traffic law.

In the first version of the story, buying a ticket is one of the steps of the "movie-going" frame, and thus the input is incorporated into this frame, which was previously activated by a reference to going to the movie theatre. That is, the input is considered to be an elaboration of the "ticket-buying" sub-frame of "movie-going." No other interpretation of the input is considered.

In the second version, getting a ticket is not found to be part of any activated frame. Indexing on this input returns a number of frames, including "traffic-violation" and "movie-going". Since traffic violation receives activation from the driving frame, it is chosen; the other unrelated alternatives are discarded. Thus concretion is performed on the input by finding the most precise frame consistent with the input that is supported by evidence from the story.

5. Conclusions

One feature of this theory is that it ties processing very closely to representation (but not entirely, since activation can presumably spread along non-epistemological associations). Since the theory is heavily dependent upon the nature of representation, a substantial commitment to a particular representational scheme is required before the details of the theory can be examined or evaluated. The details of such a scheme are beyond the scope of this paper.

The examples above are certainly not beyond the capabilities of previous story understanders. The claim made herein is that the proposed formulation will be conducive to the expansion of such systems. Presumably, if the theory of processing is correct, knowledge need only be represented correctly in order to be used. While it may be easier to build an inference engine tightly geared to a particular from of knowledge without solving these representation problems, I suggest that this only defers a problem that one will eventually have to confront.

6. Acknowledgments

The ideas in this paper emerged from several seminars at the Berkeley Artificial Intelligence Research Project and at the Berkeley Cognitive Science Program. Participants at this seminar included, in addition to this author, Yigal Arena, Margaret Butler, David Chin, Joe Faletti, Chuck Fillmore, Paul Jacobs, Paul Kay, Marc Luria, Jim Martin, Jim May field, Peter Norvig, and Lisa Rau. Independent work by Chuck Fillmore and Paul Kay and their students in the Reading Project was also responsible for some of the direction this work has taken. In addition, Peter Norvig has been working on implementing the processing scheme alluded to above; Yigal Areas has been working on the interaction of memory and linguistic processes. Many ideas described herein are due directly or indirectly to their research.

7. References

1) Anderson, J. R., ft Bower, G. H *Human Associative Memory* Washington: Winston ft Sons, 1073.

2) Charniak, E. On the use of framed knowledge in language comprehension. *Artificial Intelligence*. Vol. II. no. 3. 1078.

3) Cullingford, R. E. Script application: Computer understanding of newspaper stories. Yale University Computer Science Department Research Report #110, 1078.

4) Dyer, Michael G. In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension. Yale University Computer Science Department Research Report #210, 1082.

5) Kay, P. Three Properties of the Ideal Reader. Berkeley Cognitive Science Report #3, 1082.

6) Lehnert, W. G. Plot Units: A Narrative Summarization Strategy. In W. G. Lehnert and M. H. Ringle (eds.) *Strategics for Natural Language Processing.* Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1082.

7) Quillian, M. R. Semantic Memory. In M. L. Minsky (ed.) *Semantic Information Processing.* MIT Press, Cambridge Mass. 1068.

8) Schank, R. C. ft Abelson, R. P. *Scripts, Plana, Goals and Understanding.* Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1077.

0) Wilensky, R. *Understanding Goal-Baaed Stories.* Yale University Department of Computer Sciences Research Report #140, 1078.

10) Wilensky, R. What's the Point? In *Proceedings of the Third National Conference of the Canadian Society for the Computational Studies of Intelligence.* Victoria, British Columbia, May 1080.

11) Wilensky, R. *Planning and Understanding: A Computational Approach to Human Reasoning.* Addison-Wesley Publishing Company, Reading, Mass., 1083.

12) Wilensky, R. ft Arens, Y. A Knowledge-based Approach to Natural Language Processing. University of California, Berkeley Electronic Research Laboratory Memorandum No. UCB/ERL/M80/34, August 1080.