WITH A SPOON IN HAND THIS MUST BE THE EATING FRAME

Eugene Charniak
Department of Computer Science
Yale University

ABSTRACT

A language comprehension program using "frames" "scripts", etc. must be able to decide which frames are appropriate to the text. Often there will be explicit indication ("Fred was playing tennis" suggests the TENNIS frame) but it is not always so easy.("The woman waved while the man on the stage sawed her in half" suggests MAGICIAN but how?) This paper will examine how a program might go about determining the appropriate frame in such cases. At a sufficiently vague level the model presented here will resemble that of Minsky (1975) in it's assumption that one usually has available one or more context frames. Hence one only needs worry if information comes in which does not fit them. As opposed to Minsky however the suggestions for new context frames will not come from the old ones, but rather from the conflicting information. The problem them becomes how potential frames are indexed under the information which "suggests" them.

1 INTRODUCTION

Understanding every day discourse requires making inferences from a very large base of common sense knowledge. To avoid death by combinatorial explosion our computer must be able to access the knowledge it needs without irrelevant knowledge getting in its way. A plausible constraint on the knowledge we might use at a given point in a story or conversation (I shall henceforth simply assume we are dealing with a story) is to restrict consideration to that portion of our knowledge which is "about" things which have been mentioned in the discourse. So if we have a story which mentions trains and train stations, we will not use our knowledge of, say, circuses. This requires, of course, that given a topic, such as trains, or eating, we must be able to access its knowledge without going through everything we know. Hence we are lead in a natural way to something approaching a notion of "frame" (Minsky 1975): a collection of knowledge about a single stereotyped situation.

In the above discussion however I have made a rather important slight of hand. Given a story we only want to consider those frames "about" things in the story. How is it that we decide which frames qualify? I was able to gloss over this because in most situations the problem, at least at a surface level, does not appear all that difficult. If the story is about trains, it will surely mention trains. So we see the word "train", and we assume that trains are relevant. What could be easier.

Unfortunately, this ease is deceptive for the story may mention many topics of which only a few are truely important to the story. For example.

> The lawyer took a cab to the restaurant near the university.

Here we have "lawyer", "cab", "restaurant" and "university" all of which are calling for our attention. Somehow on the basis of later lines we must weed out those which our only incidental.

But a more immediate difficulty are those situations where a story deals with a well defined topic, yet never explicitly mentions it. So consider:

> The woman waved as the man on the stage sawed her in half.

Here we have no difficulty in guessing that this is a magic trick, although nothing of the sort has been mentioned. We are able to take "low level" facts concerning sawing, stages, etc and put them together in a higher level "magician" hypothesis. As such, the phenomena illustrated here is essentialy bottom up.

Of course, any time we try to infer relatively global properties from more local evidence we may make mistakes. That this creates problems in frame determination is shown by the nice example of Collins et. al. (forthcomming). (To get the full import of the example, try pausing briefly after each sentence.)

> He plunked down $5 at the window. She tried to give him $2.50 but he refused to take it. So when they got inside she bought him a large bag of popcorn.

The first line is uniformly interpreted as a buying act (most even going further and assuming something like a bet at a racetrack). The second line is then seen as a return of change, but the refusal is problematic. The third line resolves all of this by suggesting a date at the movies - a considerable revision of the initial hypothesis.

To sumarize the last few paragraphs, the problem of frame determination in language comprehension involves three sub-problems.

1) Stories will typically elude to many higher frames, any of which might serve as the *context for the incoming lines*. How do we choose between them?
2) The words used in a story may not directly indicate the proper higher frame. How do we do the bottom up processing to find it?
3) If we are lead astray in the course of (2), how do we correct ourselves on the basis of further evidence.

In the paper which follows I will be primarily concentrate on (2) with (3) being mentioned occasionally. In essence my position on (1) is that it will not be too much of a problem, provided that the cost of setting up a context like "restaurant" is small. If it is never used then as the story goes on it will receeded into the background. How this "receeding" takes place I shall not say, since for one thing it is a problem in many areas, and for another, I don't know.

Concerning (2) and (3), we will be lead to a position similar to that of Minsky (1975) and Collins et. al (forthcomming) in that a frame will be selected on the basis of local evidence, and corrections will be made if it proves necessary. We will see however, that there are still a lot of problems with this position which do not at first glance meet the eye.

2 THE CLUE INTERSECTION METHOD

Rather than immediately presenting my scheme, let me start by showing the problems with an alternative possibility, which I will call the "clue intersection" method. This alternative is by no means a straw man as one researcher has in fact explicitly suggested it (Fahlman 1977) and I for one find it a very natural way of thinking about the problem.

The idea behind this method is that we are given certain clues in the story about the nature of the correct frame, and to find the frame we simply intersect the possible frames associated with each clue. To see how this might work let us take a close look at the following example

> As Jack walked down the aisle he put a can of tunafish in his basket.

The clues here are things like "aisle", "tunafish' etc. Of course, I do not mean to say that it is the English words which are the clues, but rather the concepts which underlie the words. I will assume that we go from one to the other via an independent parsing algorithm. (However this assumes that there is no vicious interaction between frame determination and disambiguation. Given that disambiguation depends on prior frame determination (see (Hayes 1977) for numerous examples) this may be incorrect.) So the input to the frame determiner will be something like

```
ST-1 (WALK JACK-1 AISLE-1)
ST-2 (PERSON JACK-1)
ST-3 (EQUAL (NAME JACK-1) "JACK")
ST-4 (EQUAL (SEX JACK-1) MALE)
ST-5 (AISLE AISLE-1)

ST-6 (PUT JACK-1 TUNA-FISH-CAN-1 BASKET-1)
ST-7 (BASKET BASKET-1)
   . . .
```

The details of the representation do not figure in the paper, and those which do are fairly uncontroversial. An exception here is the use of specific predicates like BASKET or AISLE. We will return to this point in the conclusion.

Given this representation we can imagine one method of finding the appropriate frame. Our clues are the various predicates in the input, such as as AISLE, BASKET, etc. Index under each of them will be pointers to those places where it comes up. Under AISLE we might find CHURCH, THEATER, and SUPERMARKET, while BASKET will have LITTLE-RED-RIDING-HOOD, and SUPERMARKET. The point is that none of these clues will be unambiguous, but when we take the intersection the only thing which will be left is SUPERMARKET.

There are, however, problems with this view of things. For one thing it ignores what I will call the "clue selection" problem. Put in the plainest fashion the difficultiy here is deciding exactly what clues we will hand over to the clue resolution component, and in what order. So in the last example I selected some of the content of the sentence to hand over to the clue resolver, in particular AISLE, and BASKET. This seemed reasonable given that they do tend to suggest "supermarket", as desired. But there is more information in the sentence. It was Jack who did all of this. Why not intersect what we know about Jack with all of the rest, or WALK? Or again, suppose something ever so slightly odd happens, such as the basket hitting a screwdriver which is on the floor. SCREWDRIVER will have various things indexed under it, but more likely than not the intersection with the rest of the items mentioned above will give us the null set. For that matter, is there any reason to only intersect things in the same sentence? The answer here is clearly no, since there are many examples which require just the opposite.

> Jack was walking down an aisle. He was pushing his basket.

But if we do not stop a sentence boundries where do we stop? It is ridiculous to go through the entire story collecting clues and then do a grand intersection at the end.

A reasonably natural solution to the clue selection problem would start with the observation that usually we already have a general frame. When new clues come in we see if they are compatible with what we already believe. If so, fine. If not, we see if the clue suggests a different context frame. If not (as with, say, WALK which occures so often as to be unsuggestive) then nothing more need be done. If there are newly suggested context frames they should be investigated. This will be done for every predicate. Now the clue intersection method is compatible with this idea, but in its broad outline we are moving closer to what I have been characterizing as the Minsky proposal.

Furthermore, there are some problems with the clue intersection method which go beyond the mere suggestive. Consider the following example

> Jack took a can of tunafish from the shelf. Then he turned on a light.

After the first line the intersection method shoul leave us undecided between KITCHEN and SUPERMARKET. The next line should resolve the issue, but how is it that it does so? It must have something to do with the fact that normally a shopper at a store would not be the person to turn lights on or off, while it would be perfectly normal for Jack to do it in what presumably is his own kitchen. But this sort of reasoning is not easily modeled by clue intersection because it would seem to depend on making inferences which are themselves dependent on having the context frames available. That is to say, before we can rule out SUPERMARKET, we need some piece of information from the SUPERMARKET frame which will enable us to say that Jack should not be turning

on a light, given that he is cast in the role of SHOPPER in that frame.

Interestingly enough, Fahlman (who I earlier noted is a proponent of the clue intersection method) had a major role in the evolution of the Minsky proposal which I advocate. As such it behoves us to consider why he then rejected the idea in (Fahlman 1977). His primary reason is his observation that frequently in vision one does not have any single clue which could serve as the basis for the first guess at the appropriate frame. Rather it would seem that one has a multitute of very vague features, each one of which could belong to a wide variety of objects or scenes. To select one of them for a first guess would be quite arbitrary and would involve one in an incredible amount of backtrack. It would seem much more plausible to simply do an intersection on the clues and in this way weed out the obvious implausibilites.

While this analysis of the situation in vision is quite plausibile, I estimate that high level vision is still in a sufficiently rudimentary state that these conclusions need not be taken as anything near the final word. Furthermore, even if it were proved that vision does need an intersection type process, I can easily believe that the process which goes on in vision is not the same as that which goes on in language. For one thing in vision there is a natural cut-off for clue selection - the single scene. For another, within the scene there is a natural metric on the likelyness of two features belonging to the same frame - distance. Weither or not these in fact work in vision, they do suggest why someone primarily worried about the vision problem would not see clue selection as the problem it appears to be in language.

3 DIFFERENT KINDS OF INDICES

As I have already said, the scheme I believe can surmount the difficulities presented in the last section is a variant on one proposed by Minsky, and elaborated by Fahlman (1974) and Kuipers (1975). The basic idea is that one major feature or clue is used to select an initial frame. Other facts are then interpreted in light of this frame. If they fit, fine. If not then another frame must be found which either complements or replaces the original frame. In the previous propolsals the original frame contained information about alternate frames to be tried in case of certain types of incompatabilities. This may or may not work in vision (which was the primary concern of those mentioned earlier) however I shall drop this part of the theory. In discourse there are simply too many ways a frame can be inappropriate to make this feasible. For example, it stretches credibility to believe that SUPERMARKET would suggest looking at KITCHEN in the case the shopper turns on the lights.

So let us consider a very simple example.

Jack walked over to the phone. He had to talk to Bill.

It seems reasonable to assume that we guess even before the second sentence that Jack will make a call. To anticipate this we must have TELEPHONING indexed under TELEPHONE. When we see the first line we first try to integrate it into what we already know. Since there will be nothing there to integrate it into, we try to construct something. To do this we look to see what we have indexed under TELEPHONE, find TELEPHONING, and try that out. Indeed it will work quite well, since one of the things under TELEPHONING is that the AGENT must be in the proximity of the phone, and Jack just accomplished that. Hence we are able to integrate (AT JACK-1 TELEPHONE-1) into the TELEPHONING frame, and everything is fine.

Nothing is ever really this simple however, and even in this example, which has been selected for its comparative simplicity, there are complications. I suspect most people have assumed in the course of this example that Jack is in a room, and perhaps have even gone so far as to assume he is at home. Nothing in the story says so of course, and if the next line went on to say that Jack put a dime into the phone we would quickly revise our theory.

To account for our tendency to place Jack in room, we must have a second index under TELEPHONE which points to places where phones are typically found. (An possible alternative is to have this stated under TELEPHONING but this would make it difficult to use the information in cases where no call is actually being made, so TELEPHONING, even if hypothesized, would not stay around long.) So we will hypothesize two kinds of indices, an ACTION index and a LOCATION index. This distinction should mirror the intuitive difference between placing and object in a typical local and placing an action in a typical sequence. Other distinctions of this sort exist and may well lead to the introduction of other such index types locating objects and actions in time for example. However I would anticipate that the total number is small (under 10, say).

To illustrate how these index types might hook up to TELEPHONE I will use a slightly extended version of the frame representation introduced in (Charniak 1977) and (Charniak forthcomming). From the point of view of this paper nothing is dependent on this choice. It is simply to give us a sepecific notation with which to work.

```
(TELEPHONE (OBJECT) ;The frame describes an OBJECT
                    ;(and not, say, an event).
     VARS:(THING)   ,I only introduce one variable
          ...       ;THING which is bound to the
                    ;token in the story repre-
                    ,senting the phone

 LOCATION:((ROOM (HOME-PHONE . THING))
           (PUBLIC-LOC (PAY-PHONE . THING)))

      ;If we instantiate the ROOM frame then the
      ;HOME-PHONE variable in it should be bound
      ;to the token which is bound to THING.
      ;Similarly for PUBLIC-LOC and PAY-PHONE.

   ACTION:((TELEPHONING (PHONE . THING)))
          ...) ,Other portions of the frame would
               ;describe its appearance, etc.
```

We will not be able to integrate the first line of our story into any other frame, so we will hypothesize the TELEPHONING frame and either the room frame or the public place frame. Given my subject data on what people assume, the room frame is placed, and hence tried, first. This will cause the creation of two new statements which serve to specify the frames now active, and their bindings

```
(TELEPHONING (PHONE . TELEPHONE-1))
(ROOM (ROOM . ROOM-1)
      (HOME-PHONE . TELEPHONE-1))
```

The syntax here is the name of the frame followed by dotted pairs (VARIABLE . BINDING). Earlier I used a place notation for simplicity, e.g.,

```
(TELEPHONE TELEPHONE-1)
```

In fact this would be converted internally to the dotted pair format:

```
(TELEPHONE (THING . TELEPHONE-1))
```

I might note that my variables are what Minsky (1975) calles "slots". They are also equivalent (to a first approximation) to KRL slots such as HOME-PHONE in.

```
[ROOM-1 (UNIT)
        <SELF (a ROOM with
               HOME-PHONE = TELEPHONE-1)>]
```

So we are hypothesizing 1) an instance of telephoning, where the only thing we know about it is the telephone involved, and 2) a room (ROOM-1) which at the moment is only furnished with a telephone. Note that this assumes that in our room frame we have an explicit slot for a telephone. This is equivalent to assuming that rooms typically have phones in them.

We can now integrate the fact that Jack is at the phone into the telephoning frame, assuming that this state is explicitly mentioned there (i.e. we know that as part of telephoning the AGENT must be AT the TELEPHONE). With this added our TELEPHONING statement will now be:

```
(TELEPHONING (AGENT . JACK-1)
             (TELEPHONE . TELEPHONE-1))
```

When the second line comes in we must see how this fits into the TELEPHONING frame, but this is a problem of integration. The frame determination problem is over for this example.

CONSTRAINTS ON THE HYPOTHESIS OF NEW FRAMES

Early on we noted that it was only necessary to worry about a new frame if we received information which did not fit in the old ones. Then when we introduced the two kinds of indecies we noted that we wanted to place events in a sequence of events, and objects in their typical local. This immediately suggests that when we get an unintegratable action we use the ACTION index on the predicate, while for objects we would use the LOCATION index. However, this is not general enough in at least two ways.

For one thing, often we will have a non-integratable action where it is not the action frame, but rather the objects involved in the action which suggest the appropriate frame. Our example of someone going over to a phone is a case in point. Here GO tells us nothing, but TELEPHONE is quite suggestive. To handle this the search for ACTION indices must include those which are on OBJECT frames describing the tokens involved in the action. So since Jack is going to something which is a telephone, we look on the ACTION index of TELEPHONE.

We must also extend our analysis to handle states. If we are told that Jack is <u>in</u> a restaurant we must activate RESTAURANTING. In our current analysis (RESTAURANT (THING . RESTAURANT-1)) will not do this since it is an OBJECT frame and hence will only be looking for LOCATIONs in which the restaurant will fit. Hence in this case the IN frame must act like the GO frame in looking for ACTION indecies in which it might fit. More generally, any state which is typically modified by an action should cause us to look for ACTION indicies. So IN or STICKY-ON would do so, SOLID or AGE would not. (But if in the case at hand we are told that something did change the SOLID status then we would treat it like an action, as in "In the morning the water in the pond was solid".

Up to this point then the frame selection process looks like this:

1) When a statement comes in try to integrate it into the frames which are already active. In general this can require inference and a major open problem is how much inference one performs before giving up. If the integration is successful, then go on to the next statement.
2) If the statement is a description of an object (i.e. an OBJECT frame) then use the LOCATION index on the frame to find a frame which incorporates the statement. Keep track of yet untried suggested LOCATION frames.
3) If the statement is an action or changable state, then look for an ACTION frame into which the action (or state) can be integrated. First look on the frame for the

action (or state) and then on the object frames describing the arguments of the action (or state). Again, keep track of any remaining ones.

4) There must be a complicated process by which we test frames for consistancy with what we know about the story already. If it is not consistant we must involve an even more complicated process of deciding which is more believable, previous hypothesis about the story, or the current frame. I have nothing to say on this aspect of the problem.

There is however one type of example which raises some doubts about the above algorithm. These mention some object with associated ACTION frames, but only in connection with states which do not demand an ACTION frame for their integration. For example:

> The car was green. Jack had to be home by three.

In this example the above algorithm will not consider DRIVING because GREEN will not demand that we look at the action index assoicated with its arguments (the car). (Even if it did nothing would happen because the fact that the car is green would not integrate into DRIVING.) However, much to my surprise, when I gave this example to people they did not get the DRIVING frame either. However, with a modified example they do.

> The steering wheel was green. Jack had to be home by three.

This is most mysterious. One suggestion (Lehnert personal communication) is that to "see" the steering wheel the "viewer" must be in the car, which inturn suggests driving (since IN would demand action integration). This may indeed be correct, but we must then explain why in the first example the fact that the viewer must be NEAR the car does not cause the same thing. In any case however, these examples are sufficiently odd that it seems inadvisable to mold a theory around them.

5 MORE COMPLEX INDICES

There is one way in which the telephone example makes the problem look simpler than it is. In the case of TELEPHONE it seems reasonable to have a direct link between the object TELEPHONE and the context frame TELEPHONING. In other cases this is not so clear. For example, we earlier consider the example:

> The woman waved as the man on the stage sawed her in half.

Here it would seem that the notion of sawing a person in half is the crutial concept which leads us to magic, although the fact that the woman does not seem concerned, and the entire thing is happening on a stage certainly help re-enforce this idea. But presulably the output of our parser will simply state that we have here an incident of SAWING. Does this mean that we have under SAWING a pointer to MAGIC-PERFORMANCE? At first glance this seems odd at best. Some other examples where the same problem arise are:

> The ground shook.
> (EARTHQUAKE) (Example due to J. DeJong)

> There were tin cans and streamers tied to the car. (WEDDING)

> There were pieces of the fusilage scattered on the ground. (AIRPLANE ACCIDENT)

In the final analysis the real problem here is one of efficiency. If, for example we attach EARTHQUAKE to EARTH, then we will be looking at it in many circumstances when it is not applicable. (The alternative of attaching it to SHAKE is little better, and possibly worse since it would not handle "Jack felt the earth MOVE beneath him" assuming the average person gets EARTHQUAKE out of this also.)

One way to cut down the number of false suggestions is to complicate the indices we have on each frame. So far they have simply been lists of possibilities. Suppose we make them discrimination nets. So, under SAWING we would have various tests. On one branch would appear MAGIC-PERFORMANCE, but we would only get to it after many tests, one of which would see if the thing sawed was a person. In much the same way the discrimination net for EARTH could enquire about the action or state which caused us to access it. If it were a MOVE with the EARTH as the thing moved then EARTHQUAKE.

Note however that if there were few enough things attached to SAWING our net would not save significant time. Even if we were to access the MAGIC-PERFORMANCE frame the first thing we would do is check that the thing proposed for the SAWED-PERSON variable was indeed a person. The net only saves time when a single test in the net rules out a number of frames. At the present time I have not thought of enough frames associated with SAWING to make this worth while. But as I suspect this is primarily do to lack of work on my part, I will assume that discrimination nets will be required.

If we allow a discrimination net to ask arbitrary questions there will be the problem that it may ask questions which are not yet answered in the story. However a reasonable restriction which would prevent this would go as follows. Suppose statement A causes us to look at frames on an index of B. The discrimination net may only enquire about the predicate of A (EARTH looks to see if A was a MOVE), and what object frames describe the arguments of A or B (SAW looks to see if the thing sawed was a PERSON).

6 OTHER USES OF FRAME DETERMINATION

Earlier I noted that integrating a statement into a frame requires inference. Here I would like to point out that a modification of the above ideas would be helpful in this process as well. Consider the following:

Jack went to a restaurant. The menu was in Chinese. "What will I do now", thought Jack.

Our rules here will get us to RESTAURANTING after the first line. But if we are to understand the significance of the last line we must realize the import of line two; Jack can't read the menu. It would seem unlikely that RESTAURANTING would ask about the language of the menu, hence sentence two cannot be immediately integrated into RESTAURANTING. More reasonable would be to know that if something is in a foreign language it cannot be read, and one normally reads the menu so one can order. Only the second of these can plausibly be included in RESTAURANTING.

Given our algorithm the following will occur. The second line will become something like (IN-LANGUAGE MENU-1 CHINESE). Since the statement is not integrated we look to see if there is an ACTION pointer on IN-LANGUAGE. Indeed there is, and it will be to the following rule.

```
(READ (MOTIVATIONAL-ACTIVITY)
   VARS:  ...
  EVENT:
   (AND
      (SEE READER READING-MATERIAL)
      (IN-LANGUAGE READING-MATERIAL LANGUAGE)
      (KNOW READER LANGUAGE))
    ENABLES
   (KNOW-CONTENTS READER READING-MATERIAL))
```

In effect we are saying here that the typical signficance of something being in a certain language is whether a person can read it or not. This will cause us to activatve the READ frame. Initially there is little else we can do since at this point the we do not even know who is trying to read. However when we try to integrate READ we will be successful, and we will have further bound READER to JACK-1. At this point (and this is the modification required) we should return to READ and note that we can assume he does not know Chineese and hence will not be able to read the menu.

7 CONCLUSION

There is, of course, much I have not covered. The most glaring ommision is the lack of any discussion of how one detects a discrepency between a suggested frame and what we already know of the story. The problem is that a frame cannot afford to mention everything which is incompatable with it - there is simply too much. And the same is true for everything which is compatable. Furthermore, what would be enough to switch to a new frame under some circumstances would not be sufficient at other times. So "Jack walked down the isle and picked up a can of tunafish" takes us from CHURCH to SUPERMARKET. But if we added "from a pew" things are different. These are major problems and aside from (McDermott 72) and (Collins et. al. forthcomming) they have hardly been confronted, much less solved.

Early on I commented that the only controversial aspect of my representation was the use of very specific predicates (BASKET, AISLE, TELEPHONE, etc) rather than a break down into more primitive concepts. We might, for example, define AISEL as a path which is bounded on each side by things which are considered pieces of furniture (e.g., shelves or chairs). The problem with using a primitive representation here is that while it is somewhat plausible having SUPERMARKET and CHURCH indexd under AISLE, indexing them under PATH or some other component of the primitive definition is much less plausible. However, we can circumvent this problem by the use of discrimination nets, just as we did to get EARTHQUAKE from MOVE and EARTH. But, it should be noted that by using this method we are eliminating one of the benefits of a primitive analysis - we can no longer assume that we can get our information in a piecemeal fashion and come out with the same analysis. In particular we must get "aisle", or else we must get all of its components at the same time. If we do not then the discrimination net will fail to notice that we do not have any old path, we have an AISLE. Given this restriction the primitive and non primitive analyses come out pretty much the same. A primitive decomposition just becomes a long name for a higher level concept. Or to turn this around, the use of high level discriptions is not so controversial after all - it is simply a short name for a primitive decomposition.

ACKNOWLEDGEMENTS

I have benefited from conversations with J. Carbonelle, J. DeJong W. Lehnert, D. McDermott, and R. Wilensky, all of whom have been thinking about these problems for a long time. Many of their ideas have gone into this paper. This research was done at the Yale A.I. Project which is funded in part by the Advanced Research Projects Agency of the Department of Defense and monitored under the Office of Naval Research under contract N00014-75-C-1111.

REFERENCES

Charniak, E., A framed PAINTING on the representation of a common sense knowledge fragment. Journal of Cognitive Science, 1, 4, August 1977.

Charniak, E., On the use of framed knowledge in language comprehension, forthcomming.

Collins, A, Brown, J. S., and Larkin, K. M., Inference in text understanding, in: R. J. Spiro, B. C. Bruce, and W. F. Brewer (Eds.) Theoretical issues in reading comprehension. Hillsdale, N. J., Lawrence Erlbaum Associates, forthcomming.

Fahlman, S. E., A hypothesis-frame system for recognition problems, Working Paper 57, M.I.T. Artificial Intelligence Lab, 1974.

Fahlman, S. E., A system for representing and using real-world knowledge. Unpublished Ph.D. thesis, M.I.T., September 1977.

Hayes, P. J., Some association-based techniques for lexical disambiguation by machine.

TR25, University of Rochester Computer Science Department, June 1977.

Kuipers, B., A frame for frames, In D. Bobrow and A. Collins (Eds.) Representation and understanding, New York, Academic Press, 1975

McDermott, D., Assimilation of new information by a natural language understanding system, TR 291, M.I.T Artificial Intelligence Lab, 1972.

Minsky, M., A framework for representing knowledge. In P.H. Winston (Ed.), The psychology of compher vision, New York, McGraw-Hill, 1975, pp. 211-277.