Massachusetts Institute of Technology
Artificial Intelligence Laboratory

AIM 431

November 1977

Frame-based Text Processing

by

Steven T. Rosenberg

This paper presents an overview of a theory of discourse structure, and discusses a model for assimilating text into a frame-based data structure. The model has been applied to the analysis of news articles. The theory assumes sentences contain links to the database which are relatively easy to compute. These links point to prior themes which contain expectations and procedural knowledge. This knowledge is used to assimilate new sentences to these themes. At any given time, only procedural knowledge from the indicated theme is active in processing new sentences.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the Laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

Frames-based Text Processing

Introduction

If understanding connected discourse can be viewed as combining the meanings of individual sentences, then the placing of new information from successively read sentences into some conceptual structure becomes the fundamental operation. The essential tasks of a discourse processor capable of doing this can be divided into two operations: 1) Linking by locating a prior context in the story database in which to place new knowledge; and 2) Mapping the new information in a sentence into that context. From this perspective, ease of understanding depends on reducing the amount of processing spent on mapping operations which fail due to an incorrect link. In other words, a goal of well written text is to assure that only allowable mappings are attempted. Such text is sprinkled with various cues to guide the reference processes.

The rules for linking will take advantage of the context-specifying information available in the text itself. To do so, the model assumes that new sentences are linked to the story database derived from the text through a restricted set of contexts. Each new sentence is viewed as containing a linking cue to one of a relatively small set of themes.

The more explicit the linking cue, the easier it is to focus the understanding process on mapping meanings into the appropriate part of the data base. The usefulness of theme structure and thematic expectations will be determined by the ease with which various cues in the sentences can be used to indicate intersentence links. These cues provide the initial focus by directing attention to the particular mappings which should be attempted.

Knowledge Representation

A parsing scheme which can function as input for this model must provide case information as well as a syntactic description. The discourse understanding program currently being designed

takes deep case frames as input. It does not deal with the problem of mapping from the surface case structure into these deep case frames. However, other researchers in the Intelligent Support Systems group at MIT are working on this issue (see Bullwinkle, 1977). The partitioning of the processing stream from natural language into surface case and syntactic structure, then into deep case frames is a common approach of natural language systems.

Goldstein and Roberts (1977) have developed a frame representation language which forms the basis for this semantics. Surface case frames are ordinarily defined in terms of the selectional restrictions on the cases a verb may take. Deep case frames may also contain slots for such things as further cases, paraphrase labels, links to episodically related frames, pointers to generic information, and procedural knowledge. Additional processing on noun phrases or clauses in the surface case structure which specify particular cases allows substitution of computed values for the phrase or clause. (e.g. substituting the current time for the time case value specified by the phrase "at this time").

Within the frame system, a slot can be further specified through the use of associated "Facets". Each slot in the frame may have an arbitrary number of user and system defined "Facets". Useful system defined facets are the $Value facet, which contains the value of that slot; the $Default facet, which specifies a default value; $Require, which specifies requirements which any value for that slot must meet; $To-Add, which specifies procedures which must be run to compute a value for that slot; $If-Added, which specifies actions to be taken if a value is added to that slot, and $If-removed. Procedural knowledge is contained in the facets associated with a slot.

The frame system provides a slot, which each frame must have, called an AKO (A Kind Of) slot. The value of this slot will point to the more generic frame which the current frame is a modification of. Thus the frame system is organized as a tree structure. Generic information is "bumped" up the tree, with particular frames specifying new distinguishing knowledge. The

generic knowledge is inherited automaticly, including computational procedures. The inheritance can be restricted, if desired.

Input to our discourse understanding model consists of deep case tokens representing a sentence meaning. Tokens distinguish particular instances of a frame from the generic type. Since there are no primitives in the frame system, any frame may have an instance which is a token. The tokens cannot be implicitly distinguished by finding terminal nodes in the frame tree, since such nodes may or may not be tokens. All frames can inherit a type marker unless they are specifically marked as tokens.

Discourse Links

Let us consider the various types of cues available for linking related sentences to a prior theme. We wish to link each new sentence to an appropriate part of a Story Data Base (SDB). To do so we examine the sentences looking for linking cues which will indicate which theme is to be the current one. To find these cues, we assume first that there are only a small number of types of cues in sentences which indicate the link, and second that every sentence contains such a link. (Of course, in less than very well written text this is not always true. Such text is more difficult to understand precisely because more work must be done to determine how what is said in one sentence relates to prior themes.)

Some of the forms intersentence links can take are: anaphoric reference (B.L. Nash-Webber, 1976; Hankamer and Sag, 1976), cleft and pseudo-cleft sentences (Carpenter and Just, 1977), noun phrase reference (C.Bullwinkle, 1977), conversational postulates (Grice, 1967), functional sentence perspective (Kuno, 1975), the topic-comment distinction (Moore, 1967), and compiled phrases (Becker, 1975). The particular links this paper focuses on form a small but useful subset of this larger group. Anaphoric references are cleaned up in the transition from surface to deep case frames, and I do not deal with them. Bullwinkle (1977), however, is tackling this particular reference

problem, among others, within the context of frame based semantics. The set of linking cues I use are: (A) Paraphrase, (B) Contextual Reference, (C) Frame reference and (D) Contiguity.

A) Paraphrase (Frame reference by name) Unlike anaphoric references, which presuppose the existence in the database of an object which that pronoun refers to, paraphrase allows the referenced object to be directly evoked by using a label. This has several advantages. If the refered-to objects already exist in long term memory, skimming becomes a much easier job, since each sentence can stand alone.

Consider how a concept (represented as a Frame) could be referenced. It might have a name. For example, the frame for the Oil Cartel might have the name "Oil Cartel". Whenever this name occurs in text, the frame is directly referenced, if it already exists in the story data base, or else is added to the S.D.B. Suppose after reading a sentence with "Oil Cartel" as its topic, we read a sentence containing "O.P.E.C." Presumably this also is a name for the same frame. There are other names which can be used as unique designants of the frame. For instance, "the oil exporting nations". We can think of these different PARAPHRASEs as all being permissable labels for the frame. Their use results in a reference to the "Oil Cartel" frame if a previous use of a label has added it to the story database. However, the use of a paraphrase, unlike a pronoun, does not presuppose that the frame already exists in the story database. The sentence is understandable even if this is not the case, with the paraphrase serving to evoke the frame. Consequently, successive uses of PARAPHRASEs function as discourse links, while the first use serves to place the frame in the story data base. Each frame will have an associated list of allowable names which directly access it. (This holds both for frames which are used to represent objects or things (Oil Cartel) and for frames which are used to represent events. For instance, we have "the Oil Crisis", "Oil Embargo", etc.) New paraphrases are built up through repeated use of contextual referents or the creation of new idioms.

B) Contextual Reference (Frame reference by description) is closely related to paraphrase. Contextual reference is the use of a descriptive term in a context where it functions as a reference or paraphrase. In the following sentences

(S1) The report discussed the recent border clash.
(S2) The incident was not considered important.

the use of the phrase "The incident" in (S2) is an unambiguous reference to the "border clash" of the first sentence. In the context of the discourse, only the "border clash" is an incident. Other contexts might result in the phrase having a different referent. Consequently contextual reference involves the use of a general term in a context where its referent can be uniquely disambiguated. Repeated use of a particular contextual referent in a text causes it to become a local paraphrase. After its initial introduction, "incident" would refer to the "border clash" throughout the text. If "incident" itself had paraphrases, these would also function as local referents to "border clash". Such local contextual referents can be saved to reduce later processing. A sub-class of this category is the use of more or less generic names as references (e.g. "gun" vs. "forty-five caliber automatic pistol").

C) Frame Reference (Reference using Slots) refers to a form of reference which depends on the empty slots of a frame for effect. It operates either within or between sentences, and illustrates that the processing of intersentence links is not separate from intrasentence processes. Consider the following two sentences.

(S3) John shot his wife.
(S4) The gun was a forty-five caliber automatic.

In this case the sentences can be linked through the realization that the action of the first sentence ("shot") has an unspecified instrument. The second sentence specifies this instrument. Suppose a sentence like

(S3a) His motive was jealousy.

is interposed between sentences (S3) and (S4). We still have no difficulty in linking them. Consequently, the scope of the case expectations can extend over more than one sentence, although my hypothesis is that when a new theme has been introduced, they become more difficult to disambiguate.

D) Contiguity I propose as a rule of discourse that if there is no explicit link to a previous theme (i.e. frame), the current sentence will be linked to the same theme as the preceeding sentence. This is refered to as the CONTIGUITY option. This allows us to treat the cues as a complete closed set.

To determine how well these categories covered the field of discourse links, and their relative frequency of occurence in well written text, an examination was made of several newspaper articles. Newspaper articles provide a good domain, since they have evolved under very practical constraints into prose which provides a maximum of information in a minimum of space, in a format which is one of the most easily understandable.

The articles were chosen from the New York Times. The procedure was to take one front page article a day for several days, and count the frequency of occurence as a thematic link of each of our cues. The category of anaphoric links was included, to determine the proportion of sentences for which this link is used. A total of eight articles were examined in this way. (Appendix 1 gives a sample article.) Out of a total of 259 sentences, only 68 were linked through anaphoric references. This is not surprising, since Rosenberg and Lambert (1974) have found that sentences with many intersentence anaphoric references were more susceptible to processing difficulties than those with fewer such links. Paraphrase links were used in 97 sentences, while 54 sentences were linked by contextual references. Contiguity occured in 40 sentences. No instances of frame reference as a thematic link occurred in these particular articles.

To further explore the use of these cues, three articles were examined intensively, and their

thematic structure charted. Consider the nature of thematic organization. The two possible dimensions are depth and breadth. A depth first thematic organization would first introduce a theme, discuss it exhaustively, and then pop back up the stack and introduce a new theme, without ever returning to the previous theme. A breadth first theme structure would first introduce all sub-themes, saying a few things about each. It would then return to the first, and say a few more things about it, then go on to the next, etc. This process would repeat itself to some arbitrary level of depth. Actual prose is a combination of these two strategies. In either case, we find local groups of sentences discussing some sub-theme. At some point it becomes necessary to "pop up" and switch to a new sub-theme. An interesting question is whether links between sentences with different themes differ from those used locally within a group of sentences which discuss a single theme.

To answer this, a numerical value was assigned to each link in the three articles. Those links which were "local" were given a value of 0, while those used in "poping up" or otherwise linking a theme with other themes were given a value of 1. A "local" theme link can be defined as a link to the current theme. A non-local link refers to a different theme. The mean value associated with each cue indicates the relative uses to which that cue is put. (CONTIGUITY cues are not included in this statistic since they are by their nature local cues). Anaphoric reference was never used to link across themes. The mean values for paraphrase links in the three articles were .75, .36, and .43. The corresponding values for contextual referents in those articles were .17, .45, and .50.

These results indicate that in our sample anaphoric reference is never used to link across themes. However, "poping" and "pushing" themes will be a crucial part of understanding discourse. Consequently, although the sample is quite small, and we should be cautious in drawing conclusions from it, it seems that our discourse cues will be useful in handling discourse structure in news articles.

Discourse mapping

Once a surface link is found, the new information must be mapped into the indicated theme. A sentence may be related to a prior theme in many ways. The most important of these ways are:

A. Instantiating the frame description of that theme.

B. Augmenting expectations for attributes of the frame.

C. Providing reasoning knowledge involving that theme.

D. Introducing a new theme as the value of some attribute of the current frame.

These are not exclusive. More than one choice can occur in a sentence. Lets consider these four alternatives.

(A) Instantiating a frame description of a theme. The theme can either be an object ("Oil Cartel") or an event ("Oil Crisis"). In either case we use a FRAME to represent the theme. In many cases there are no values specified for some slots. Such defaulted frame slots are UNINSTANTIATED. INSTANTIATING a frame involves substituting actual values given in sentences for default values. The discourse processor expects that in a piece of discourse all uninstantiated slots for a theme will be instantiated somewhere in the course of the article. New knowledge in a current sentence related to that theme may instantiate part of it.

Thus, uninstantiated slots create EXPECTATIONS. The program EXPECTS these slots to be instantiated by information in succeeding sentences. Once a link has been found between the current sentence and a previous theme any expectations associated with that theme become candidates for instantiation. This mapping is effected by attempting to instantiate the new knowledge from a current sentence onto the expectations of the linked frame. Of course, in any particular article, not all expectations will be instantiated, but this is not important. We wish to restrict the alternatives considered at any one time. Hence the assumption that any new knowledge in a sentence is related to one of a set of known themes, and that by examining cues in the

sentence this theme can be discovered. For each new sentence, only a small number of alternatives (relative to the total number of frames in the story data base) are ever actively considered. Although there may be many expectations, only the small set associated with a particular theme are ever active at any one time. The function of discourse structure is precisely to limit the number of alternatives which have to be considered.

(B) Augmenting expectations The process of instantiating expectations may also result in the following alternative. Although a new sentence may not actually instantiate an expectation, it might help us by providing further information about the form of that instantiation. Consider a frame for murder evoked by the sentence

(S5) John murdered his wife.

An expectation for an instrumental case exists. Suppose the next sentence had said:

(S6) He used a blunt instrument.

Although the case has not been instantiated, in the sense that a particular object has not been specified as the value, we now know that a semantic feature of the instrument is that it is blunt. This can be encoded in the $req Key for the instrument slot by the addition of a new requirement which any instrument must meet. Our expectation has been AUGMENTED by the knowledge in the new sentence. If the third sentence said:

(S7) A frozen leg of lamb and a vial of poison were found next to the corpse.

The expectation for the instrument slot can now be instantiated with "leg of lamb", since this fulfills the requirements of the slot. Discourse often gives requirements for the semantic features of an instantiation, before an actual slot value which fills these requirements occurs. As more and more requirements are added to the theme, the context becomes more restricted. This produces, over the course of a text, more and more specific expectations. Once a sufficient level of domain specific knowedge is encoded, expectations can be extended to include the use of prior context

knowledge (scripts) and plans. These can be used to generate expectations when the text indicates the appropriate theme, although the knowedge in the text itself does not support such an elaborate set of expectations.

(C) Reasoning knowledge Some sentences do not instantiate a theme, but instead provide reasons, causal links to previously described events, etc. We will call such information REASONING KNOWLEDGE. Such knowledge often expresses a simple causal relation in a non-quantitative fashion. Articles often seek to explain theme events by expressing reasoning knowledge about the event.

(D) Sub-themes The term theme may have been used in a confusing manner. Every sentence designates at least one topic, but not all topics are themes. The topic of a sentence corresponds to the intuitive notion of "what the sentence is about", although linguists have not been able to agree on a single formal definition. We represent a sentence meaning as a set of deep case frames. One of these frames will represent the action or thing the sentence is "about". This frame is the topic, and is linked to the other frames in the sentence representation. The topic may only instantiate a previous theme. (i.e. it is mapped into an existing deep case frame.) If it also has forward consequences across sentence boundaries it is considered a theme. It is not always possible on first encountering a new topic to decide if it is also a new theme. Bullwinkle has suggested that if the next sentence does not expand on some expectation in the current topic, it is no longer available as a default theme, unless an explicit reference to it is put in (Bullwinkle, 1977).

Popping a theme occurs whenever the discourse reverts back to discussing a previous theme. Pushing a theme occurs when a slot value in the current theme becomes a new theme. All slots in a given frame are, of course, frames themselves. Not all these slots are of equal importance. The values of some slots provide links to frames which are part of the same theme as the current frame. These are slots (such as the cause slot), which link events, or slots (such as the domain slot),

which link objects, where the links are episodic in nature. For instance if two events are linked in a sentence through a causal connection (i.e. the cause slot in one contains the other as value), we can say that both share the same Context-Horizon (C-H). In examining the first frame as a theme, we also wish to consider the caused event as part of that theme, by getting the value of the "cause" slot, and consider that value's slots' expectations. However we wish to avoid the infinite regression of trying to instantiate a slot value's slot value's slot value's . . . . . Thus the C-H defines the boundaries of the theme. Slot values of frames in the C-H are the terminal depth boundary of the theme. Some slots, such as "cause" are known to the system as potential paths to other parts of the theme.

Linking Given a token, the linking component tries to find links of the following sort (contextual reference and paraphrase have had sub-categories broken out): (1) Literal reference; (2) Paraphrase; (3) generic reference; (4) Frame Reference; (5) Contextual Reference; (6) Contiguity.

Lets consider a general model for matching two frames. Any two frames will share part of a heritage path in the frame tree descending from the top most node, which will diverge at some point. The common portion defines the semantic match between the two frames. Frames in the branching portions define semantic characteristics which are not shared. Two frames may match if the heritage path of one entirely contains the heritage path of the other. In this case one frame is a semantic subset of the other.

Two frames may share semantic features in common, by having identical slots, but contain different values for those slots. Thus, the linked token is mapped onto the corresponding token of the indicated theme. This mapping compares all slots which have values in either token against the corresponding slot's values, augmentations, and requirements. All slot values in common should either match, or the new value should be in a slot which takes multiple values. (It is also possible

for one value to be part-of the other value, as we shall see later.) New values for uninstantiated slots may trigger procedural knowledge associated with that slot, which evaluates the potential new value. The semantics for the particular slots will determine the outcome of the mapping. For instance, some values may change over time, hence a difference may not be important. This type of mapping procedure corresponds to determining whether a literal or paraphrase reference is correct.

If one frame matches another frame's heritage path, but descends deeper into the frame tree, we have an instance of possible generic reference. Generic reference is the use of more or less generic terms (e.g. dog vs. animal) as referents. There exists a direct path in the frame tree between these two terms. The path may involve going up one AKO link from the more generic term, as when both "rifle" and "flintlock" inherit from "gun", but not directly. However, if the path is more indirect, the two terms will not be co-referential. Both "gun" and "knife" are weapons, but not references. Generic links can be found by determining if such a heritage path exists between a potential referent and a previous token. Once again, the mapping determines the appropriateness of the reference.

Not all frames are linked to all possible subordinates. Contextual referents are general terms which do not usually fit into the direct inheritance tree of the terms they can refer to, except in the trivial sense that something is an event, or object. For instance, consider a "forecast". Almost anything could be a forecast, if it had the right slot values (i.e. had not yet occured.) However, we do not wish to make everything inherit from the forecast frame. Slot requirements and augmentations of the frame are used to determine whether any particular recent frames are actually referents.

A contextual reference provides a description of its referent through its slots and their associated values, requirements, and augmentations. An "international incident" for example, is any

event whose duration slot value is small; whose participants slot values are countries; and whose action slot value can be described as hostile. This description can be matched against frames in the current strory database.

Frame reference can be implemented by using the augmentations on a slot to define the lowest possible node in the frame tree from which a value must inherit. A demon on the instance slot of this frame will examine each new token inheriting from this frame, filtering it through the requirements and augmentations, to determine if it is an appropriate value for the slot. Such demons are not computationally expensive since they do not need to examine each new input, but are instead automaticly triggered by likely candidates.

MAPPING Once a link has been found, it is necessary to map the sentence tokens into the theme. This is a straight-forward process consisting of successive maps. If the link is a frame in the context-horizon of the sentence, it is mapped onto the corresponding frame in the theme, perhaps filling some of the expectations. A mapping step then occurs by determining if the sentence context frame is linked to any other frames in its context horizon. If so, a check is made for a corresponding frame in the theme. If it exists, the map step is repeated; if it doesn't exist, the token for the new frame is linked to the appropriate frame in the story data base by instantiating the link value. If the sentence frame is a value of a slot of a frame in the context-horizon of that sentence, it is necessary to determine if the corresponding dominant context-horizon frame exists in the story data base, and map these to determine if the referent is correct. If so, processing proceeds as above.

An Example

Consider the problem of understanding the following first paragraph from a Commodities Report article. This article is used in a program designed to allow the user to explore the use of expectation and instantiation of new topics onto themes. The program takes the deep case frame

representations of the headlines and sentences as input. From these it builds a tree whose nodes and links represent the thematic structure. Each node points to a frame representing a topic. All new topics are added to the tree. Some topics are also themes. The nodes for those topics indicate this. Topic nodes are linked to theme nodes, but not directly to each other. The nodes are pointers to the clusters of frames organized around topics which represent the sentence meanings. The program trys to add each new topic to the tree. On finding a link between a new topic and an existing theme, it tries to instantiate the topic onto that theme.

If no theme exists, the first topic encountered becomes the theme. If more than one theme exists in the tree, the program will move about the current sentence representation generating possible candidates for a theme link. The problem of choosing the right link is currently finessed by using crude heuristics and allowing the user to choose from among the potential links which are found. (See, however, Rosenberg and Roberts, forthcoming, for a more detailed description of the matching procedures necessary for finding theme links in a system implemented in FRL.) When the user chooses a potential link, the program generates the expectations of that frame, presents these to the user, and asks whether the user wishes to attempt to instantiate the new sentence onto these expectations. In the event he does not, or the instantiation fails, the program makes use of its knowledge of how one frame is linked to another frame in a theme to suggest alternatives. Nodes on the tree point to individual frames for topics and themes. These frames are actually part of a cluster of frames representing sentence meanings and knowledge which has been instantiated onto a particular theme. Thus the tree functions as an index to the themes and topics. The frames pointed to are entry points into individual clusters of frames, and are the most important frames (thematically speaking) in those clusters. However, the actual expectations which a new topic fills may not be in this initial frame but in one of its associated frames. Similarly, a topic is linked to the other frames representing the sentence meaning. These frames, rather than the topics, may be

the ones which instantiate thematic expectations. Thus in order to instantiate a new topic onto a theme it is necessary to find the appropriate points of correspondence between the theme frames' expectations and the frames associated with the topics. The program can "walk" through clusters of frames. In whatever frame it is in, it can generate its expectations, and try and map the candidate topic onto these expectations. It can also examine the links between this theme frame and other frames, and if the instantiation fails, suggest a likely candidate. For example, it may suggest that the current frame is the domain of some larger action, which it would be appropriate to consider. In trying to instantiate a topic onto a theme it will consider other frames linked to the new topic as possible candidates. If all attempts at instantiation fail, the user can return to the current topic for more candidate theme links. The program is interactive and allows the user to "walk" around the various theme structures and frames, choosing which matchings and expectations he wishes to try. The following scenario describes how a future version of the program would find thematic links as well as assimilate the new information.

WHEAT

AN ANALYSIS OF THE LATEST U.S.D.A. SUPPLY PROJECTIONS

This week the U.S.D.A. issued a new 1975-76 forecast for the supply of wheat. The figures reflect the department's opinion of the latest announced indications of Russian interest in U.S. grains, therefore the change seems important. Previously, the U.S.D.A. had expected a wheat carryover next june 30th of 497 million bushels. The forecast this week adjusts that total downward from 497 to 395 million.

The frames which must be specified for this task fall into several areas. Basic domain-independent knowedge about one of these domains, supply and demand, is contained in the Supply frame.

SUPPLY		
AKO	$VALUE	THING
DOMAIN		
SUPPLY	$DEFAULT	CARRYOVER, PRODUCTION, IMPORTS
OLD-CARRYOVER	$IF-ADDED	TOTAL-SUPPLY, CARRYOVER
	$IF-REMOVED	TOTAL-SUPPLY, CARRYOVER
	$DEFAULT	0
PRODUCTION	$IF-ADDED	TOTAL-SUPPLY, CARRYOVER
	$IF-REMOVED	TOTAL-SUPPLY, CARRYOVER
	$DEFAULT	0
IMPORTS	$IF-ADDED	TOTAL-SUPPLY, CARRYOVER
	$IF-REMOVED	TOTAL-SUPPLY, CARRYOVER
	$DEFAULT	0
TOTAL-SUPPLY	$DEFAULT	TOTAL-SUPPLY
USE	$DEFAULT	DOMESTIC, EXPORTS
DOMESTIC	$IF-ADDED	CARRYOVER, TOTAL-USE
	$IF-REMOVED	CARRYOVER, TOTAL-USE
	$DEFAULT	0
EXPORTS	$IF-ADDED	CARRYOVER, TOTAL-USE
	$IF-REMOVED	CARRYOVER, TOTAL-USE
	$DEFAULT	0
TOTAL-USE	$DEFAULT	TOTAL-USE
CARRYOVER	$DEFAULT	CARRYOVER
DATE	$DEFAULT	NOW

The Wheat-supply frames are daughters of the Supply frame which specify that the domain is wheat, and the units are million bushels. The slot values for the current crop are stored in instances of the Wheat-Supply Frame.

WHEAT-SUPPLY		
AKO	$VALUE	SUPPLY
DOMAIN	$VALUE	WHEAT
UNIT	$VALUE	MILLION BUSHELS

WHEAT-SUPPLY-76		
AKO	$VALUE	WHEAT-SUPPLY
YEAR	$VALUE	1976
OLD-CARRYOVER	$VALUE	320
PRODUCTION	$VALUE	2138
IMPORTS	$VALUE	1
DOMESTIC	$VALUE	694
EXPORTS	$VALUE	1400

The bottom-most frame (Wheat-supply-76) only contains values. The more general Supply frame has several $If-addeds whose purpose is to maintain relations among the changing slot

values. For instance, total-supply is the sum of domestic production and imports. If the estimate of domestic production is changed, the total supply will be updated. These procedures are inherited by Wheat-supply-76.

Since the task is to learn the current crop values by understanding the article about the current report, we need to represent several properties of reports in a frame:

(1) There are different KINDs of reports, among which is Supply-and-Demand.

(2) Reports have a SOURCE, which in the current case is the U.S.D.A.

(3) Reports have a FORM, such as Figures, Values, Tables, etc.

(4) They also have a DOMAIN, such as Wheat-supply.

(5) They have a STATUS (Final, Current, Projected, etc.)

(6) They have a TIME.

Reports are represented in a Report frame. We will also need a frame for the event of reporting:

```
REPORT
AKO          $VALUE          THING
DOMAIN
SOURCE       $REQUIREMENT    (MUST-BE ROLE ACTOR)
FORM         $DEFAULT        TEXT
STATUS
PARAPHRASE   $VALUE          INDICATION
DATE         $DEFAULT        NOWC
             $IF-ADDED       ADJUST-STATUS


REPORTING
AKO          $VALUE          EVENT
PARAPHRASE   $VALUE          ISSUE
SOURCE       $REQUIREMENT    (MUST-BE ROLE ACTOR)
RESULTS-IN   $REQUIREMENT    (KIND-OF REPORT)
ACTION
DOMAIN       $IF-ADDED       RESULTS-IN
DATE
```

The most recent report is always "current", while the previous one has the status "previous".

When a new report on a particular domain occurs, we want the status of other reports to be relative to this new one. The $If-added "adjust-status" achieves this by comparing dates for reports which have the same domain and changing statuses appropriately.

The $If-added on the source slot ensures that sources are either actors (animate objects), or corporate objects which can function as actors, such as the government.

Reports can be about almost anything. Consequently, there are no requirements on the domain slot. Different kinds of reports specify restrictions on their domains through requirements on this slot. Forecasts are reports which discuss things which have not yet taken place. Since forecasts may be useful, they are given their own frame:

FORECAST		
AKO	$VALUE	REPORT
DOMAIN	$REQUIREMENT	FUTURE-DATE
PARAPHRASE	$VALUE	PROJECTION
STATUS	$DEFAULT	CURRENT

Analysis of the Headlines

At the beginning of the article there will be no story database. In this case, the topic of the first input becomes the theme. Headlines are often phrases or sentence fragments, rather than full sentences. Since these can just as easily be represented in frames as full sentences, this does not pose a representation problem. However discovering the theme of a fragment is more difficult. Sentences describe actions and can result in changes to the database. Either the expected change or the event itself can be treated as the theme. However, sentence fragments often do not specify a specific theme. Consider the first headline in the commodities article:

(S7) WHEAT

Wheat is known to be used in many different ways. Consequently we have no way of choosing one of these as the theme, unless we have some prior expectations. The headline itself does not

describe a theme which allows us to derive specific expectations. Since it is the major headline, we can assume it is central to the theme explored in the article. We treat it as constraining the future theme. More formally, we expect the headline to instantiate this theme. We will place a demon on our "try" list (which is a short term buffer of strategies), which says that when a theme is found, try to instantiate "wheat" onto it .

The second headline is:

(S8) AN ANALYSIS OF THE LATEST U.S.D.A. SUPPLY PROJECTIONS.

"Analysis" is a nominalization of the action "analyze". An analysis is then the result of analyzing some domain. Many things bear this "result-of" relation to actions. For instance, a report is the result-of a reporting action. Because we know the analysis is the result of analyzing something, the value of the domain slot for an analysis frame (which is AKO report) is required to have the characteristics of something which has been analyzed. This means that we expect its slot values to have been determined. We also wish to express the relation between the action and the nominalization through complimentary "results-of" and "results-in" slots. e.g.

```
ANALYZE
AKO          $VALUE          EVENT
RESULTS-IN   $REQUIREMENT    (MUST-BE ANALYSIS)
DATE
DOMAIN       $IF-ADDED       (EXPECT DOMAIN INSTANTIATION)

ANALYSIS
AKO          $VALUE          REPORT
RESULT-OF    $REQUIREMENT    (MUST-BE ANALYZE)
DOMAIN       $IF-ADDED       (EXPECT DOMAIN INSTANTIATION)
```

Both frames take domain slots. Nominalization relations noticed when mapping one frame onto another will be taken account of in determining coreference.

The rest of the sentence fragment forms the domain of the analysis: "the latest U.S.D.A. supply

projections". This is represented by a frame for "projection" with the modifiers filling slot values. The headline is then represented by:

PROJECTION1			ANALYSIS1		
AKO	$VALUE	FORECAST	AKO	$VALUE	ANALYSIS
STATUS	$VALUE	CURRENT	DOMAIN	$VALUE	PROJECTION1
DATE	$VALUE	LATEST			
DOMAIN	$VALUE	WHEAT-SUPPLY1			
	$REQUIREMENT	MUST-BE SUPPLY			
		(COULD-BE WHEAT-SUPPLY IF-INSTANT DOMAIN WHEAT)			
SOURCE	$VALUE	U.S.D.A.			

"Projection" is a paraphrase for "forecast". Thus Projection1 is a token of the "forecast" frame. (I shall not indicate tokens on the example frames, but all frames representing sentences in the story are tokens, as well as previous wheat-supply frames which contain crop values.) Of course, a "Supply" forecast is one which has its domain specified to be a kind of supply frame. This is an augmentation which is put in as a new requirement, since the supply frame itself does not accept values. Generally the frame being instantiated will indicate the requirements of a value, as opposed to an augmentation; Wheat-supply values, for instance, must be quantities.

"U.S.D.A." instantiates the source slot inherited from the report frame, meeting the requirement that it be an animate or corporate actor. The frame for "U.S.D.A." specifies that the supply of wheat is among the values reported on by the U.S.D.A. Sources of reports are always checked so that the kinds of things the sources report on can be used as expectations. This makes use of the relation that a source for a nominalization (e.g. report) must also be the source for the related action (e.g. reporting). This can be specified on the source slot of the Report frame with an $If-added. When a source is established, the $If-added will specify that the values for the report slot of that source are expectations on the domain slot of the Report frame. Once one of the expectations is satisfied, the rest are eliminated. However, if the domain slot took multiple values,

we would not treat these expectations as a disjunctive set. We indicate this with a $Require on the domain slot which is a list of "could-be's" (only one of which is included here).

The commodities world involves many indicators of relative and absolute time: "most recent", "current", "latest", "past", "earlier", etc. The appearance of a new report can change the status of earlier reports relative to other information. Indefinite temporal terms can also function as expectations. If no absolute date for an event is given, phrases such as "the latest" are sent to the time specialist which attempts to compute them. If it cannot do so, they are considered as augmentations of expectations (i.e. further restrictions on) on a time slot. Such relative phrases will require the services of a temporal specialist (Kahn, 1975) to automatically update such judgements, and keep them consistent with the "present time" state of the world. The status slot is kept current by the specialist through the inherited $If-added "adjust-status" on the date slot.

The sentence context consists of Analysis1 and Projection1. These are substantially instantiated, except for a date and domain slot. The try list is now examined, and the first headline, "wheat" is mapped against expectations in the sentence context. This headline satisfies the requirements on the "domain" slot by specifying the expected type of Supply frame to be a Wheat-supply frame, since this is the only "could-be" satisfied by wheat. A token of the Wheat-supply frame becomes the domain, augmenting the context-horizon. (Domain values, like cause slots, are used to link frames in the same context horizon).

The body of the article

The first sentence is:

(S9) THIS WEEK THE USDA ISSUED A NEW 1975-76 FORECAST FOR THE SUPPLY OF WHEAT.

This sentence recapitulates the headlines, adding only two new bits of information: the date of issue of the report, "this week", which confirms the computed value for "latest" (this is calculated by

asking the database how often these reports are issued, and if the interval since the last one is within this range, using the current date to some approximation, such as a week); and the crop year, 1975-76, which is the default value, being the current crop year at the time the article was written. The action of "issuing" can be considered a paraphrase for "reporting". (In treating these and other terms as paraphrases, I am discarding nuances of meaning which a more subtle system should capture by expanding the frame hierarchy):

The sentence is therefore:

ISSUE1

AKO	$VALUE	REPORTING
SOURCE	$VALUE	U.S.D.A.
DOMAIN	$VALUE	FORECAST2
DATE	$VALUE	THIS WEEK

FORECAST2

AKO	$VALUE	FORECAST
STATUS	$VALUE	NEW
SOURCE	$VALUE	U.S.D.A.
DOMAIN	$VALUE	WHEAT-SUPPLY2

WHEAT-SUPPLY2

AKO	$VALUE	WHEAT-SUPPLY
DATE	$VALUE	1975-76

Sources are passed down by the $If-added "results-in" on the domain slot of reporting. This illustrates another use of the nominalization relation. Since the action and resulting state share many aspects in common, such as a source, if a domain is instantiated, the $If-added checks the domain to determine if it is a nominalization of the event. If so it passes down all common values such as the source.

The three frames Issuel, Forecast2, and wheat-supply2, form the sentence context. I will introduce the term context pointer to indicate which frame in the sentence or theme context is being examined. A Candidate-generator sets the context-pointer at Issuel, and sends this reference candidate to the linking component. No reference will be found since no other reporting event has occured. Since this frame has slots with values, these are considered before switching the pointer. The date slot is not an object or action, and the domain slot is used to change the context-pointer.

Thus the source slot value is examined next. U.S.D.A. is a proper name, and a literal reference. A previous token is recognized as having occured in the theme in the Projection1 frame. The mapping component now tries to match the two frames this value occurs in. This fails since they are not tokens of the same frame. The fact that issue1 can have a nominalization was noted when the sentence frames were examined for reference candidates, and put on the try list. When the mapping fails, it is triggered and tries to find either an actual nominalization value in the domain slot of Issue1, or an action of which Projection1 is a nominalization. The former case succeeds. The mapper next maps the domain value of Issue1 against the Projection1 frame, since both of these are nominalizations of the same kind-of reporting event. Forecast2 is now mapped against Projection1. This map succeeds, since all comparable slots with values match. Issue1 is now added to the database as the action which results in Projection1. The context-pointer is shifted to the third frame, Wheat-supply2, and this is mapped against the corresponding Wheat-supply1 which is the value of the domain slot of the Projection1 frame. These also match, with the result that a few minor values are instantiated: the date of the wheat-supply crop as 1975-76, and "this week" for "latest".

The next sentence is:

(S10) THE FIGURES REFLECT THE DEPARTMENT'S OPINION OF THE LATEST ANNOUNCED INDICATIONS OF RUSSIAN INTEREST IN U.S. GRAINS, THEREFORE THE CHANGE SEEMS IMPORTANT.

The sentence consists of two clauses linked by "therefore". This sentence is one where we shall wish to interact with the reasoning component. The reasoning component knows about particular kinds of assertions, and particular relations among assertions. Both "reflect" and "therefore" trigger it to pay special attention to this sentence. "Reflect" in this usage is understood to mean that "the figures" represent "the department's opinion". To determine this involves reasoning on these figures. "Opinion" is understood to be a kind of "analysis" in which reasoning knowledge is

associated with slot values. "Announced" is understood in terms of being a paraphrase like "issuing". Similarly, "indication" is considered to be a paraphrase for " report". "Russian interest in U.S. grains" forms the domain of this report. "interest" is understood only in terms of supply relations. It means a "plan to buy". Thus "Russian interest in U.S. grains" is understood as Plans Russia has to purchase our grains. In the database, the Use slot label of the Supply frame is itself a frame which contains knowledge of specific buyers. Each buyer is assumed to have a buying plan which is modified in light of actual purchases. The U.S.D.A. is the source of knowledge about these plans. These plans are hypothetical, and must be reevaluated periodicly in light of new evidence. The reasoning component compares new buying by a buyer against the plan, to re-evaluate it. Changes in the Wheat-supply summary statistics are treated as being evidence for either the implementation or change of plans. Thus if export demand increases, it must be because either a planned buy has occured, or some buyer has changed his purchase plan. Since the present report is a prediction for the current crop, changes will contain both confirmed purchases plus expected purchases. Therefore, whenever the change frame is used with certain domain slot values, the reasoning component is triggered into re-evaluating (in the present case, Russian buying interest). The second clause is meant to tell us that we should pay attention to the change. e.g.

```
FIGURES1
AKO                $VALUE      DATA
CONSEQUENCE-OF     $VALUE      OPINION1
                   $IF-ADDED   REASON

OPINION1
AKO        $VALUE  ANALYSIS
SOURCE     $VALUE  U.S.D.A.
DOMAIN     $VALUE  INDICATION1

INDICATION1
AKO        $VALUE  REPORT
DOMAIN     $VALUE  BUYING-PLAN1

BUYING-PLAN1
AKO            $VALUE  BUYING-PLAN
BUYER          $VALUE  RUSSIA
COMMODITY      $VALUE  WHEAT
PLAN-BODY

THEREFORE1
AKO        $VALUE      ASSERTION-RELATION
ASSERT1    $VALUE      OPINION1
           $IF-ADDED   REASON
ASSERT2    $VALUE      CHANGE1
           $IF-ADDED   REASON
```

"U.S. grains" is a generic reference to wheat, which is used as the actual value (i.e. wheat is AKO grain. "U.S." which modifies grain, must then map onto the Wheat Frame. It does, since we are talking about american supply statistics.)

There is no instanitated Change frame yet. We can create a frame for the event class "change" with the form:

```
CHANGE1
AKO          $VALUE          CHANGE
SOURCE
DOMAIN
DIRECTION
FROM         $REQUIREMENT    (MUST-BE QUANTITY)
TO           $REQUIREMENT    (MUST-BE QUANTITY)
TIME-FROM
TIME-TO
```

Since the Change frame requires a quantity as certain slot values, the "figures" is a likely caandidate. Understanding this reference will also provide the domain, since this will be what the

figures are of. It is apparent to a reader that "figures" is used to refer to the report. Figures are one way of characterizing the form slot value. The form slot has not been instantiated yet. A frame reference demon, triggered by "figures" will instantiate the slot. This also provides wheat-supply1 as the domain for the change frame. All new frames in the sentence are examined to assign all references, as these may be impossible to decipher later. "Department" is a contextual reference to the "U.S.D.A."

The reasoning component will use this new sentence to set up reasoning expectations which will be added to all slots accepting quantitative values in the wheat-supply frame. As these values are added, the change frame will be evoked, resulting in reasoning. Each successful piece of reasoning shouts on completion, and all unsuccessful pieces give an explanation, when the article is finished. This explanation will consist of describing which values and requirements were not met. This could then be used to actively inquire for the missing data. While the change frame token pushes the theme, it is an unusual case. First, its instantiation depends on the Wheat-supply1 frame being instantiated. Thus, expectations are put on the slots of this frame. These will "wake up" the change frame should they be instantiated, and add the prior values as well as the current values. The change frame can then cause reasoning to occur. The other unusual aspect of this frame is that since its domain value is a frame which itself contains many values, and any of these may change, we actually expect many instances of the change frame to occur.

The reasoning knowledge which must be drawn on to set up expectations in this case is the following:

(a) Russian interest in U.S. grains affects supply.

(B) The crop report reflects the government's estimate of (A)

(C) The change in the crop report reflects the change in the government's estimate of (A). Although these cannot yet be used, since no values are given, once these are instantiated with

particular values from the current report, a prediction can be made. e.g. The amount of change in the crop report reflects a change in Russian plans to import U.S. grains.

The next sentence is:

(S11) PREVIOUSLY, THE U.S.D.A. HAD EXPECTED A WHEAT CARRYOVER NEXT JUNE 30TH OF 497-MILLION BUSHELS

To create the sentence context, we add the following semantics for wheat-carryover.

WHEAT-CARRYOVER		
AKO	$VALUE	CROP-STATISTIC
PART-OF	$VALUE	WHEAT-SUPPLY
VALUE		
DATE		

The sentence can now be represented as:

EXPECTED1			WHEAT-CARRYOVER1		
AKO	$VALUE	FORECASTING	AKO	$VALUE	WHEAT-CARRYOVER
DATE	$VALUE	PREVIOUSLY	VALUE	$VALUE	497
DOMAIN	$VALUE	WHEAT-CARRYOVER	DATE	$VALUE	JUNE 30, 1976

Expected1 is the first reference candidate. This is a possible reference to Issue1. The date slot value "previously" is a relative time term. When evaluated by the time specialist, it indicates that the previous token for Forecasting which involves the same Wheat-carryover domain is the approriate token. In other words, this sentence is seen as referencing a token for Forecasting in the S.D.B. which was placed there by some previous article. This evoked token must now be linked to the current theme structure. Issue1 is the possible link. This mapping will succeed. The context-pointer now moves on to the domain values. Wheat-carryover is not a reference to Forecast1. However, a report and a report value cannot be directly compared. Thus the domain of Forecast1, which is Wheat-supply1, is used. (It is implicit that information is conveyed by a medium such as a report, conversation, etc. We do not always want to be creating dummy report frames to contain

knowledge, but neither do we wish matches to fail because in one case the report is explicitly given, while in the other it is not. Thus, if one slot value in a match is a report we will use its domain in matching against a non-report slot value.) The part-of relation is checked. A domain value in a subsequent reference may be a sub-set of the original. This succeeds. The value in that slot matches the new value. Since there are no expectations on the old token for Forecasting, the current theme is maintained, but a hint is placed on the try list stating that if the next reference to this theme gives trouble, consider this previous state.

(S12) THE FORECAST THIS WEEK ADJUSTS THAT TOTAL DOWNWARD FROM 497 TO 395 MILLION.

Note in this sentence the definite reference "The forecast", the contrasting "this week" to the earlier "previously", and the use of the definite reference "that total". In addition the old total of 497 million is restated.

ADJUST1		
AKO	$VALUE	CHANGE
SOURCE	$VALUE	FORECAST3
DOMAIN	$VALUE	THAT TOTAL
DIRECTION	$VALUE	DOWN
FROM	$VALUE	497
TO	$VALUE	395

FORECAST3		
AKO	$VALUE	FORECAST
DATE	$VALUE	THIS WEEK

Adjust1 is the first candidate. It is a paraphrase for the Change frame. Consequently Change1 is retrieved as the theme. Mapping the frames, Forecast3 instantiates the source slot. It is a literal reference to Forecast1, which becomes the source value. Mapping these, "this week" matches the date value, resulting in success. The domain of Change1 is the Wheat-supply frame; of Adjust1 "that total". We first try and disambiguate "that total". One way to do this is to find all tokens which are "totals" (i.e. quantities) which have been mentioned in the current or previous theme. The old value for "Wheat carryover" is recognized as the most recent referent of "that total". The

sentence tells us that this total has been changed, which allows the "current" expectation for "Wheat carryover" to instantiate. In other words, we require complementary expectations on the change frame to those we placed on the Wheat-supply frame. If the Change frame is instantiated, we wish to instantiate the correct slot in its domain (the Wheat-supply frame) before creating a new change theme which focuses on that particular change.

To conclude, using a formal frame-based knowledge representation scheme, it is possible to specify processes for determining certain classes of intersentential links. Together with a model of discourse processing whose components can also be computationally specified, they form a first specification of a text processing module.

Appendix 1

ISRAEL-P.L.O. CLASH MARKS U.N. DEBATE ON WEST BANK ISSUE

Representatives of Israel and the Palestinian Liberation Organization clashed in debate before the Security Council today as the council began discussing anti-Israeli unrest in the occupied West Bank of the Jordan river.

The clash came after the council had voted, over United States objections, to allow the P.L.O. to participate in the session with the same rights as any United Nations member country whose interests are affected.

The 15-nation council overruled the objection by a vote of 11 to 1 with 3 abstensions. The lone dissenting vote - not a veto because it came on a procedural rather than a substantive matter - was cast by the new delegate of the United States, William W. Scranton, in his first appearance since his appointment.

As the debate began, the acting observer of the P.L.O. at the United Nations, Zehdi Labib Terzi, accused Israel of "Hitlerite atrocities" in the occupied territory and likened the wave of demonstrations to "the glorious Warsaw ghetto uprising" against the Nazis in World War II. Delegates of a number of Arab countries backed him.

Israel's chief delegate, Chaim Herzog, dismissed the arab charges as lies and said the demonstrations were undertaken by youths incited by what he described as false propaganda. The bulk of the West Bank population, he said has not been involved.

Mr. Herzog called the attention of the council to the current "tragedy of horrifying proportions" in Lebanon.

Mr. Terzi then raised a point of order, demanding that Mr. Herzog stick to the situation in his own country.

Britain's delegate, Ivor Richard, also on a point of order, pointed out that the council had already heard a series of speakers hostile to Israel. "The Israeli delegation is entitled to have its say," he declared.

The council president, Thomas S. Boya of Benin, ended the clash by asking the Israeli delegate to continue his statement.

This was the first time that Israel and the Palestinian Liberation Organization had faced each other in the Security Council.

RULING BROUGHT UNREST

Israel had boycotted the Council sessions in December and January on the ground that the P.L.O., which it denounced as an organization bent on Israel's destruction, was admitted. The P.L.O. attended its first session in December.

The Israeli government decided last Friday to present its case before the Council during a debate on the situation in Jerusalem and the West Bank that had been requested by the two Moslem members - Libya and Pakistan.

The anti-Israel demonstrations in the occupied territories of the last few weeks were touched of by a ruling of an Israeli magistrate that it was not unlawful for Jews to pray on Temple Mount

in the Old City of Jerusalem, the site of two Islamic shrines and the ruins of King Solomon's temple.

Israeli policemen had banned Jewish prayers there to avoid protests from Moslems, and the ban was upheld by the Israeli Supreme Court yesterday.

For the current debate, the Security Council admitted as nonvoting participants Egypt, Jordan, Saudi Arabia, Syria and Yugoslavia, as well as Israel and the P.L.O.

The three countries abstaining in the 11 to 1 vote on admitting the P.L.O. were Britain, France and Italy. Another European member of the council, Sweden, voted with the majority, composed of Communist and third-world countries.

Before the vote, Mr. Scranton declared that the United States would not have opposed participation of the P.L.O. under Rule 39 of the Council's procedural rules. This provision says that the Security Council may invite members of the United Nations Secretariat "or other persons whom it considers competent for the purpose," to supply it with information on matters under consideration.

However, the american delegate said, the United States rejected the P.L.O. attendance at the meeting under Rule 37, which provides that any member of the United Nations that is not a member of the Security Council may be invited to participate, without vote, in the discussion of an issue affecting its interests.

The American representative, a former Governor of Pennsylvania, remarked in today's statement that the United States stand on P.L.O. participation in council debates was based on principle, "a principle that cannot be eroded either by its continuing violation no matter how many times, or by time itself." Referring to his own debut in the Council debates, Mr. Scranton said he hoped "to play a part in preserving the Security Council for future generations."

Mr. Scranton spoke calmly and in measured tones. Several diplomats noted afterward that his style was in marked contrast with that of his predecessor, Danial P. Moynihan, who often seemed emotional in debate.

A dispute over seating arrangements delayed the morning meeting 90 minutes. After the procedural vote was taken, Mr. Herzog was assigned a place at one end of the horseshoe table, and Mr. Terzi was seated at the opposite end.

The observers of the other participating countries were given seats on the side of the Council table.

Israeli and P.L.O. reresentatives have on many earlier occasions been present together, though in widely separate sections, in meetings of the 144-country General Assembly. They also pass one another almost daily in United Nations lobbies, but have always studiously ignored the presence of the adversary. Today's was the first meeting eyeball-to-eyeball.

References

Becker, Joseph D. 1975. The phrasal lexicon, B.B.N. Report 3081, Bolt Beranek and Newman, Cambridge, Mass.

Bullwinkle, C. 1977. Levels of complexity in discourse for reference disambiguation and speech act interpretation. MIT AI laboratory memo 413.

Carpenter, Patricia A, & Just, Marcel Adam. 1977. Integrative processes in comprehension, in D. Laberge & S. J. Samuels (eds.) Perception and comprehension, Lawrence Erlbaum Associates, Hillsdale, N.J.

Charniak, Eugene. 1972. Toward a model of children's story comprehension Ph.D. Thesis, Artificial Intelligence Lab, M.I.T.

Goldstein, Ira & Roberts, B. 1977. Nudge, a knowledge-based scheduling program. MIT AI Laboratory Memo 405.

Goldstein, Ira & Papert, Seymour. 1975. Artificial intelligence and the study of knowledge. MIT AI Laboratory Memo 337.

Grice, H. P. 1967. William James Lectures, Harvard University.

Hankamer, Jorge & Sag, Ivan. 1976. Deep and surface anaphora, Linguistic Inquiry 3, 391-426.

Kahn, K. 1975. Mechanization of temporal knowledge. M.Sc. Thesis, M.I.T. Laboratory for Computer Science. Document MAC-TR-155.

Kuno, S. 1975. in Grossman, San & Vance (Eds.) Papers from the parasession on functionalism Chicago Linguistic Society.

Minsky, Marvin. 1975. A framework for representing knowledge, in Patrick Winston (Ed.), The psychology of computer vision, Mcgraw Hill.

Moore, T. 1967. The topic-comment function: a performance constraint on a competence model. University of California at Los Angeles dissertation.

Nash-Webber, B. L. 1976. Semantic interpretation revisited BBN report 3335, Bolt Beranek & Newman, Cambridge Mass.

Rosenberg, S. & Lambert, W. E. 1974. Contextual constraints and the perception of speech J.E.P. 102, 178-180. 1974