[Proceedings of the *Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Colgate University, Hamilton, New York, August 14-16, 1985]

The Level Hypothesis in Discourse Analysis

James Pustejovsky
Department of Computer and Information Sciences
University of Massachusetts at Amherst, 01003
CSNET Address: JAMESP@UMASS

Abstract

In this paper I would like to explore some difficult questions related to topics in discourse analysis (henceforth DA) and offer a partial solution to some of them. In particular, I will address the issue of *levels* in DA and how the various approaches taken within the field can be classified according to a leveled model. I then want to consider an approach I have been pursuing for representing the semantics of discourse, and consider how it fits in to the proposed model for DA.

1. Approaches to Discourse Analysis

There has been a great deal of renewed interest generated lately in the area of DA, motivated in part by the influence of researchers in Artificial Intelligence (AI), attempting to design "natural language conversation systems." As with many branches of AI, it at times appears as though we are reinventing the wheel, failing to take stock of past work done in related disciplines such as linguistics, philosophy, and psychology. However, much of the work has added new and complex dimensions to the study of DA (including speech act theory). I am thinking in particular of the works of Allen, Cohen and Perrault and the role of planning in speech acts; Wilks and Bien and the *Point of View* principle; and the recent work done on conversational moves and clue words, by Webber, Grosz, Sidner, Reichman, and others. The immediate uniformity between these approaches is that they are concerned with process oriented models of discourse understanding rather than claims to being competence models.

In what follows I will attempt to classify the different factors influencing the "understanding" of a discourse, and how these have been analyzed and dealt with in the field. I will assume a traditional classification of the communicative content of an utterance, *U:*

(1)

1. Truth-conditional semantics for *U.*
2. Entailments from *U.*
3. Presuppositions from *U.*
4. Conventional implicatures from *U.*
5. Conversational implicatures from *U.*
6. Felicity conditions associated with *U.*

Along with the these semantics aspects of an utterance, we must include the deeper coherence relations in a discourse, such as causal, temporal, spatial, and definitional considerations.

It is difficult to address one of the areas above without getting involved in at least one other. Therefore no clearly delineated classification is possible for "who works on which topic" and just what is meant by "semantics." Nevertheless, I would like to compare the work done on these topics by establishing what feeding relationship exists between them.

Let us begin by identifying what appear to be the three major approaches to DA:

(2) a. Structural Analysis
b. Goal Recognition
c. Model Theory

1.1 Structural Analysis

This approach is primarily concerned with how the structure of a discourse influences the interpretation as well as the linguistic realizations of a text. Chief proponents of this view are Grosz, Webber, Reichman, and Sidner.

Early work by Webber (1979) and Grosz (197) was aimed at identifying the contexts within which discourse anaphora was licensed. The notion of *focus* and *topic* was adopted to delimit the space within which anaphoric binding is possible. That is, only if something is labeled with such a discourse marker can certain pronominal references be licensed.

As Reichman (1984) puts it, the purpose of DA is to identify "a conversation's deep structure in terms of the structural relations between the discourse elements."[1] In this view discourse structure is defined by the conversational moves (CM) taken by the participants in the discourse. Each CM takes the discourse into a new stage; that is, each CM has associated effects. Also central to this model for DA is the notion of *context space,* which is an "abstract structure" taking into account the following components:

(3)

1. The propositional representation of the discourse utterance.
2. The conversational move (CM).
3. The Preconditions for the CM
4. Links to previous discourse spaces.
5. Focus level assignments for various elements in the context space.

According to Reichman's view, all discourse utterances obey certain rules, regardless of the type of discourse. A few of the more important ones are given below.

(4)

1. Conversation is a series of CMs linked by functional relations.
2. Utterances in a single context space serve the same CM.
3. A CM has preconditions and effects associated with the underlying discourse structure.
4. While in a subspace, the containing context space retains control.
5. Inter-sentential anaphoric binding is possible only with high focus items.

Central to this model of DA is the belief that conversational moves (CMs) are recoverable from the specific linguistic structure of the text. Thus, we have a taxonomy of possible CMs and the clue words most frequently associated with them:

---

1. Cf. Reichman (1984) for a full discussion.

(5)

MOVE	CLUE WORD
1. support	Because; Like
2. restatement and/or conclusion of point supported	So
3. Interruption	By the way
4. Return to interrupted space	Anyway
5. Indirect challenge	Yes, but
6. Direct challenge	(No) but
7. Subargument concession	All right
8. Prior logical abstraction	But look
9. Further development	Now

The "deep structure" of a discourse consists of a sequence of the above CMs, through which a conventional interpretation (the understanding of the discourse) is accomplished. This essentially involves recovering the *mutual knowledge* between the participants in the discourse.

1.2 Goal Recognition

A very different approach to DA is that which I will call *Goal Recognition.* This differs significantly from the structural analysis school in one important respect: what is being recovered from an utterance and what is being represented as the understanding of the discourse (or text) is something much deeper than the structural form of the text. Within this approach we can single out two major schools of thought: those concerned with narrative form, coherence, and story understanding (Schank, Abelson, Hobbs, and Wilensky); and those concerned with the recognition of speech acts and intentions (Cohen, Allen, and Perrault).

For Schank and Abelson (1977), and much of the Yale school, understanding a text is a problem of inference generation and control. That is, a reader attempts to find the implicit connections between the sentences in the text. As a solution to the infinite search space problem of inferences, they proposed that there are script-like knowledge structures which we can access in order to understand stories. Thus we recover these prototypical event-sequences, the scripts, and form a coherent understanding of the text.

Wilensky (1982) points out a number of problems with this approach, chief among them the fact that not all stories or texts can be characterized as stereotypical sequences of events. He proposes a theory of text coherence that incorporates the goals and plans that actors in a text may have. Thus, we try to recognize what the intention of the actor is and piece together the text on the basis of this goal.

Whereas Wilensky is concerned more with the underlying intentions and goals of the agents in a text, Hobbs (1978, 1982) attempts a general classification of coherence relations that may exist in a text. The two that he examines in detail (Hobbs 1982) are *elaboration* and *occasion.* These relations are formal constraints on an inference mechanism which constructs a tree-like structure for a discourse containing all the asserted and presupposed propositions (cf. Hobbs (1980)).

Lehnert (1978, 1982) is also of the purely script-based and story grammar approach to understanding as being too top-down oriented. She proposes a system of text analysis and memory organization which has the features of bottom-up processing as well.

In this theory the underlying notion of coherence is based on *affect states* and *plot units.* Affect states are a set of primitive predicates over states and events, with values *positive, negative,* or *neutral.* That is, an event is positive, etc. with respect to an object. These states are bound to objects.

In addition to these primitive predicates are links between event/state pairs that describe causal coherence relations. These are: motivation, activation, termination, and equivalence. From these notions Lehnert then defines the notion of plot unit: a *plot unit* is a directed labeled link from one affect state value to another. The underlying coherence of a narrative, then, is captured in terms of these units.

It is important to note that for these approaches, the inference processes are spawned as a result of the knowledge structures associated with propositions (and the plans they fall into) rather than linguistic or surface structural clues.

---

1. This is not completely true, of course. Some researchers in this school make use of clue words just as Reichman and Polanyi and Scha.

Alterman (1985) proposes an interesting theory of text coherence based on the notion of *event concept coherence*. This property is part of the dictionary entry for an event/state description, and provides a way to group text into structured bundles, based on their relative coherence. Alterman makes three claims for this theory: (1) text is composed of structured chunks of conceptual event/state descriptions; (2) events can be bundled together without stating their complete causal connections; (3) the initial grouping and structuring of text can be done with simple augmentation of case relationships by inter-event relations.

The concept coherence relations assumed by Alterman are characterized as follows:

(6) a. Taxonomic-class/subclass
b. Partonomic
i. sequence/subsequence
ii. coordinate
Co Temporal
i. before
ii. after

Thus in an example such as (7), it is the relative proximity of the concepts *chop* and *drop* via the concept *hold* that establishes the coherence between the two sentences.

(7) a. The peasant was chopping a tree in the woods.
b. He dropped his axe...

Another approach that addresses questions of goal recognition is taken by Cohen, Allen, and Perrault. 1 These researchers have as their primary concern the recognition and modeling of the speaker's plans in a dialogue. According to this view, speakers' intentions can be thought of as plans, and speech acts are no different from any other actions. Hence, they can be planned and recognized with algorithms and heuristics already employed in

1. The work of Grosz (1977) deals with tracking a dialogue topic in a task-oriented domain. She employed plan-tracking heuristics to this end, but did not embed speech acts into a general planning environment.

AI for planning systems (e.g. STRIPS).

Following Cohen and Perrault (1979), this approach treats actions as operators defined in terms of *preconditions* (applicability conditions), *effects,* and *bodies,* which explicate how to achieve the effects. These are evaluated relative to the speaker's models of their listeners. Thus discourse processing in this view has nothing to do with the structure of the discourse per se but rather with the intentions of the speakers.[1]

The model that a speaker has of his listeners involves representing the beliefs and goals of those people. Belief is interpreted for Cohen and Perrault as a modal operator, A-BELIEVE, taking propositions as its argument. This formal treatment (cf. Hintikka (1969)) allows for infinite embeddings of belief contexts, with the advantages and problems of such an approach.[2]

Recently Litman and Allen (1984) have extended the planning paradigm to allow plans about the planning process itself. This allows for tracking clarification subdialogues while still keeping track of the plans associated with the speech act being performed.

Finally, another important approach to belief (and goal) recognition is that taken by Wilks and Bien (1983). This "least-effort" approach to language understanding and belief representation is to be contrasted with that just mentioned, such as Allen and Perrault (1980). Wilks and Bien argue that deep nestings of beliefs could not possibly be efficient from a psychological or computational perspective. They propose as an alternative a theory of belief *percolation,* whereby temporary frames (pseudo-texts) indicating belief states can be pushed down into another such frame, if necessary the understanding of a discourse.

1.3 Model Theory

---

1. Recently Litman and Allen (1984) have proposed a model of plan recognition that does incorporate some of the strategies found in the structural analysis school. We will return to this theory below.
2. For discussion of this topic, cf. Cohen and Perrault (1979).

Recently there has been much work done on discourse within formal approaches to linguistics and semantics. I am thinking in particular of the Discourse Representation Theories of Kamp (1981) and Heim (1982) and the recent work on Situation Semantics by Barwise and Perry (1983). These approaches take (at least in spirit) as their point of departure the formal framework proposed by Montague (1974) and Kaplan's work on indexicals and demonstratives (Kaplan 1977). There isn't room here to examine these works in detail, but I will review the major points of their theories.

Kamp's (1981) main concern is the correct interpretation and representation of discourse referents. Essentially, Kamp argues that deictic and anaphoric occurrences of pronouns are identical, and that identifying their antecedents involves selection from specified sets of previously available entities. Associated with an utterance is a discourse representation structure (DRS) containing the appropriate quantification over the entities in the proposition, as well as the propositional content. To illustrate, consider the DRS for (8a), shown in (9):

(8) a. Pedro owns Chiquita.
b. He beats it.

(9)

u v
Pedro owns Chiquita
u = Pedro
v = Chiquita
u owns v

Now, the novel aspect of Kamp's proposal comes with the DRS for (8b). Because there are no possible referents within (8b) for the two pronouns, it does not license a separate DRS but must rather be embedded within (or bound by) another, satisfying structure; in this case, (9). Hence we have the DRS for the discourse pair, shown in (10).

(10)

u	v
Pedro owns Chiquita	
u = Pedro	
v = Chiquita	
u owns v	
He beats her	
u beats her	
u beats v	

The proper linking is now possible between the pronominals and their antecedents, since their is a common scope delimiter, viz. the DRS, which contains both binder and variable.

Heim's (1982) approach is similar to Kamp's in many respects, but her concerns is how to represent the presuppositions carried by utterances Crucial to this theory is the notion of a *file,* a record on which descriptions of entities can be kept, and which is evaluated with respect to rules of familiarity and *file-change.*

According to Heim, every sentence has "file change potential." That is every utterance has the potential to change the context set of the utterances following it. The common ground, in Stalnaker's (1979) terms, between the speaker and the hearer, is the set of presuppositions common to both. This is what is contained in the file of a context in Heim's theory.

Barwise and Perry (1983) provide the groundwork for a theory of situations and attitudes that allows for partial models rather than being tied to the exhaustive models of Montague semantics. I will have nothing further to say here about this approach.

2. Levels of Discourse Analysis

It is clear from our discussion above that what counts as a representation of the discourse or as the "understanding" of the text differs wildly. In this section I would like to explore how these different representations interact and propose a model for DA incorporating these component parts.

2.1 Conversational Moves versus Coherence

Let us begin by examining the logical distinction between possible conversational moves in a discourse and possible types of coherence that tie a text together. Reichman and others, following Grice (1971), classify utterances according to the roles they play in the discourse; e.g. supporting, elaborating, interrupting, etc. Others working in goal recognition have classified the types of coherence relations that exist between sentences in a text or discourse. These include causation, temporal ordering, but also notions such as elaboration and occasion. The problem here is that what some are calling *moves* in a discourse others term *coherence relations.*

Hobbs (1982), for example, describes the two coherence relations, *elaboration* and *occasion.* In the dialogue shown in (9), (b) is said to elaborate (a).

(9) a. John can open Bill's safe,
b. He knows the combination.

Similarly (10) is said to be an instance of an elaboration.

(10) a. Go down Washington St.
b. Just follow Washington St three blocks to Adams St.

Although the (b) examples above clearly elaborate the (a) sentences, there is much more that can be said about the coherence relations between them than this. The notion of elaboration Hobbs is using here is *structural* coherence and is not significantly different from a conversational move for the structural analysts. In this sense I agree that both (9) and (10) are structural elaborations.

A deeper description, however, of the connectedness between the two sentences in (9) would involve something like a because-of relation; that is, the real coherence link here is *enablement* and not elaboration. The connectedness between (l0a) and (l0b), on the other hand, is one of identity. Although structurally an elaboration, (l0b) reflects a changed performative strategy by the speaker, due to his/her model of the hearer's beliefs.

The other coherence relation Hobbs mentions is *occasion,* which can be defined simply as follows: A occasions B if A creates a state so that B can occur. An example of this is a text involving direction giving:

(11) a. Turn left.
b. Go to the corner.

By performing the action denoted in (11a) a change of location is effected

that allows the action in (11b) to occur. The structural relationship between (a) and (b) is simply a continuation or further development, and I agree with Hobbs that the coherence link here is one of occasioning.

While Hobbs and others fail to make a careful distinction between conversational moves and deeper coherence relations, still others ignore the role of discourse moves entirely. Alterman (1985), for example, develops a taxonomy of concept coherence terms with which his system creates a complete representation of a narrative text without recourse to textual moves or CMs. The obvious problem with this approach, in my opinion, is that without the structural clues provided by a discourse or text (such as topic and focus) it is impossible to adequately recover the interpretation of pronouns and deictic terms. For example, in the partial text mentioned in section 1 (cf. (10)), *he* is bound by the NP mentioned in the previous sentence, *the peasant.* But it is not the underlying coherence relation that licenses this as much as the structural positioning of the antecedent relative to the pronoun.

Determining such structural environments for discourse anaphora has been the concerns of researchers such as Sidner, Grosz, Webber, and Reichman. One such licensing context is the *domain of focus,* which accounts for the anaphoric behavior of the pronoun discussed in the previous paragraph. These theories suffer, however, from the lack of any coherent representation of the deeper semantic relations between the discourse entities.

As discussed above, Reichman proposes a theory of discourse structure based on conversational moves. Clue words act to signal when a shift in context is being made. This model takes a surface representation (call it SS) and maps it into a discourse representation (DR) using these clue words as triggers for interpretation. Thus, an utterance such as (12b) is construed as a *support* for (12a).

(12) a. I don't like John, (b) because he's rich.

Let the interpretation of (12a) be represented by -P, and (12b) by Q. The derived DR for this pair is then,

(13) -P *because* Q → supports(Q,-P)

Interestingly, however, there is another interpretation of (12) with the because connective (operator) inside the scope of the negative in (12a). The reading here can be paraphrased as, "It is not the case that (P because of Q), but (P because of Q')." The function of *because* under this reading is not direct support, but rather it triggers an entirely different set of presuppositions. Namely, the fact that there is some other support to -P that is not explicitly mentioned, and that Q does *not* support -P.

This points to the problem of what to take as the input to DA. Reichman assumes that surface structure is the natural choice, as do most structural analysts. This example, however, seems to indicate that Logical Form (LF) may have a feeding role into DA. Any presuppositions or discourse moves associated with the second interpretation would have to be derived from the LF, where the appropriate scope assignments are represented (cf. (14)).

(14) -[P *because* Q] → supports(Q',P)

Although this is an isolated example, I think it is important to study such interactions in order to establish the feeding relations between the various interpretation levels.

Another criticism that can be leveled at Reichman concerns her misunderstanding of the Toronto school's (Allen, Cohen, Perrault) meaning of "understanding." She points out that one must distinguish between a person's *intention for an utterance* and the *communicative effect* of the utterance in context: "[While] a speaker's intent may well be reflected by a communication, grasping that intent cannot be a necessary precondition for understanding."(Reichman (1984)). The confusion here is this: Reichman states that a hearer's interpretation is dependent on the communicative effect of the utterance in context, and this may or may not be identical to the speaker's intent. I agree with this, but I would not call this *understanding* the speaker. This is in fact the basis for misunderstanding in a communicative act. In order to fully understand the speaker, it is not a sufficient condition, but at least a necessary condition to recover the intent.

Finally the question arises as to where the model theoreticians fit into the discussion above. First, it is obvious that the major concerns are different for these researchers. Although questions of anaphora and reference are dealt with, Kamp's theory doesn't address the problems of inferencing or goal recognition and planning. Nor does he look at the structure or semantics of meta-sentential text and ask questions pertaining to coherence. Yet these are not his immediate interest. Heim addresses many topics related to DA as well, the emphasis being on the presuppositions from utterances, and the proper characterization of the *common ground,* the mutual belief space. Although this work highlights the importance of LF for later interpretation strategies, her concerns do not extend to the deep coherence relations addressed by Hobbs and others.

2.2 The Level Hypothesis

In this section I would like to outline a model for discourse analysis based on fairly strict levels of interpretation and I will try to establish what the feeding relations are among the conflating factors. We will address the following questions:

1. What are the levels of analysis for DA?
2. What is the unit of analysis for DA?
3. How does Discourse Representation (DR) affect interpretation?
4. Since DR is not the real semantics, what is?

This is a proposal for a process-oriented model rather than a competence model, but I do not discuss this distinction at any length.

Let us begin by separating the structural or syntactic aspects of a discourse from the coherence relations, what I will call the deeper semantics. The conversational moves discussed in Section 1.1 are structural descriptions for the constituency of the discourse itself. Also of a structural nature are domain notions such as focus and topic. These too, however, are of a meta-descriptive nature, and can be isolated from the coherence relations of the utterances. Now, in the previous section we suggested that perhaps LF is the appropriate input to DA rather than SS. We will continue with that assumption here. Let us then propose as the first link in the model the following hypothesis:

(15) LF → DR

The Logical Form of an utterance is seen as feeding the discourse representation.

Establishing such a model, however, is useless without examining what the unit of analysis for DA is. We will assume that the utterance, as defined by linguists, is the unit for analysis. But one utterance may have several communicative effects, in terms of conversational moves (CMs), and the speech acts effected. Thus, the mapping from LF to DR is not one-to-one, but rather one-to-many. For example, any nonrestrictive relative clause can be thought of as (at least) an elaboration or further development of the NP it modifies. Yet for purposes of intra-sentential anaphora and binding, we must treat it as one sentence. Similarly, adjunct clauses containing temporal adverbials and other connectives may very often signal a CM on the part of the speaker.

In order to capture this mapping let us say that one of the primitives of DR is the *clause,* i.e. a simple proposition. The syntax of DR establishes the connectedness of these clauses in terms of the CMs taken by the speaker (or inherent in the text). We express this as follows, where CF is Clausal Form:

(16) $LF \rightarrow \{_{DR}CF_i\}$

This level creates the structure from which we interpret,

a. The antecedents of discourse anaphors and deictic terms, and
b. The deep coherent relations between the Clausal Forms.

This interpretation derives a level that I will call Intentional Form (IF), and add to the model below:

(17) $LF \rightarrow \{_{DR}CF_i\} \rightarrow IF$

The primitive notions of IF include the coherence relations described in Section 12 above. Two CFs can be connected by the following deep coherence relations:

1. Causal
2. Spatial

3. Temporal
4. Definition

Causality can be thought of as a covering term to include *occasioning, enablement,* and stronger senses of causation. For now, let us think of causation as a operator that limits or prunes the possible state space following an event. Thus, where b is *temporally subsequent* to a, we determine the strength of a causing b by examining b relative to the rest of the state space generated by a.

IF will represent the goals and plans associated with the utterance as well. This will be the representation of the speaker's intention. Still, the most elusive aspect of this level is the representation of mutual belief, the common ground. Speaking in terms of what is presupposed by a listener, we will distinguish between those clauses that are asserted, those that are presupposed by the lexical structure of an item, those clauses presupposed on the basis of structural configuration, and those as a result of convention. That is, presuppositions are triggered by different elements in different environments. Now we ask, at what levels are the various presuppositions derived or computed? Lexical presuppositions, we claim, accompany the LF structure into DR; they are already computed. Structural presuppositions, on the other hand, are computed from LF and feed into DR. Conversational implicatures will be read off of DR itself, and the presuppositions associated with beliefs and common ground will be computed at IF. IF, notice, feeds into itself, indicating that inferencing is spawned as a result of these conventional inferences.

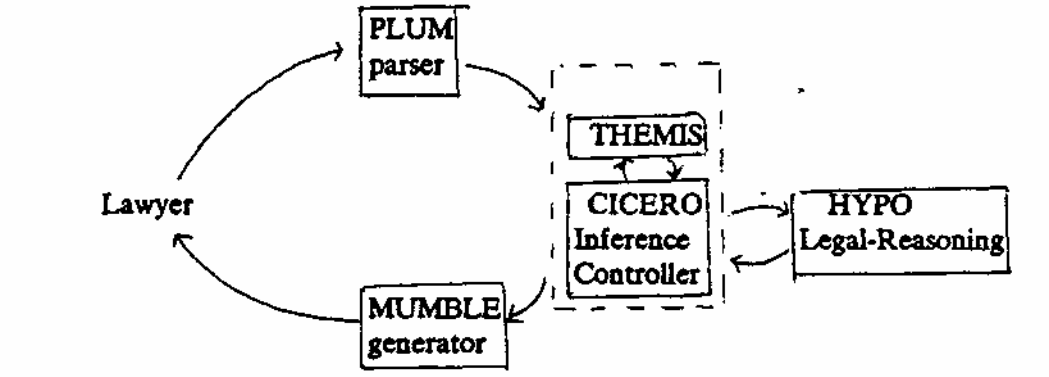
3. CICERO: Inference Controlling for Discourse Analysis

In this section I would like to describe the current and proposed capabilities of a system being designed at the University of Massachusetts. This project is part of a large natural language understanding system, COUNSELOR, currently under development in our department. I will first describe the scope of the research involved and how the various components interact. I will then give a detailed description of the discourse interpreter, CICERO, as well as the knowledge representations used by the system. At all

times I hope to make it clear bow this system's functioning relates to the model proposed in the previous section.

3.1 A Natural Language Interface for a Case-based Legal Reasoning System

COUNSELOR is the combined efforts of four separate projects to develop a case-based legal reasoning system with full natural language capabilities. The projected capabilities will allow a lawyer to interactively input the facts of a case, let the system analyze them, and propose the strongest arguments and counterarguments based on the given facts. The expert system that actually does the legal reasoning (HYPO) is essentially the intentional agent for the natural language front end, which consists of a parser, a generator, and a discourse interpreter.[1] The interaction of the systems is illustrated in (/) below.



As an example of the text and discourse encountered by the system, consider the fragment below from an interactive session between an attorney, P, and the system, S.

P: I represent a client named RCA VICTIM who wants to sue SWIPEINC and Leroy Soleil for misappropriating trade secrets in connection with software developed by my client. RCA VICTIM markets the software,

---

1. We will not be concerned with the actual reasoning capabilities of HYPO. Cf. Ashley and Rissland (1984) for details on the argumentation process involved in the system.

known as AUTOTELL, a program to automate some of a bank teller's functions, to the banking industry. In 1982, Leroy Soleil, one of RCAVICTIM's personnel, left RCAVICTIM and began working for SWIPEINC on a competing product, TELLERMATIC, also an automated teller program. SWIPEINC has begun marketing TELLERMATIC in competition to AUTOTELL.

S: Did Soleil enter into a noncompetition agreement or a nondisclosure agreement with RCAVICTIM?
P: Assume there are no agreements.
S: Did Soleil work on the AUTOTELL project?
P: Yes, he was a key employee on the project.
S: Was Soleil the only source through which SWIPEINC could get RCAVICTIM's confidential information about AUTOTELL?

This example illustrates two aspects of the understanding process: (1) fact and plan recognition (the opening paragraph); and (2) a question-answer interaction soliciting facts for the express purpose of forming an argument.

3.2 Managing the Discourse

There are two programs which contribute to the interpretation of the discourse, THEMIS and CICERO. The former can be thought of as an expert system for tracking and predicting the structure of a discourse based on conversational moves, interpreted through keywords. In its current implementation it is essentially a modification of Reichman's (1984) ATN model.

The other system, CICERO, manages and controls the representation of deeper semantic relations between discourse entities and predicates. The basic components of the system are: (1) A Knowledge Base defined in terms of *clustered objects;* and (2) A best-first control strategy generating and recognizing the plans of the speaker and hearer, respectively.

A *cluster* is a particular way to represent both the objects in the world as well as mental objects such as plans and goals that operate over them.

The ontology consists of the following types:[1]

1. objects: frames representing real-world objects with associated role-goal pairs.
2. states: predicates over the objects.
3. events: functions from one state to another state.
4. scripts: prototypical event sequences.

Using examples from the dialogue above, let us examine what structure these clustered objects have, and what role they play in the interpretation of the discourse.

Under the current implementation, when the system begins to interpret the input from the user, the discourse tracking program (THEMIS) has already set the system-mode to expect a case-facts summary from either a layman or an attorney. That is, CICERO is expecting a particular kind of speech act; namely an *inform*. This top-down expectation is represented in the current discourse frame under the slot :discourse-mode, along with the contextual parameters, :participants, :speaker-goal, and :hearer-goal.

After the parse of the initial sentence, CICERO's task is to confirm any expectations it has concerning the speaker-goal, as well as to form a coherence representation of the semantic content of the proposition. The parse output for this sentence is a legal-representation frame, and percolates the knowledge to CICERO that the speaker is an attorney. This in turn satisfies the precondition for the discourse-script shown in (18)—the coherence representation—and confirms the system's expectation for what the speaker's goal is; viz. to inform about a case.

---

1. We are assuming a standard temporal logic, such as Allen (1984) for interpreting the tense-based objects above.
2. The clusters including scripts have been implemented as flavors in Zetalisp on the Symbolics. For implementation details cf. Pustejovsky and Gallagher (in preparation).

The script illustrated in (18) clusters together the rhetorical moves associated with presenting information about a case for this particular situation. Each speech act of *inform* is represented as a separate action in the :events field, and this defines part of the larger textual structure of this preamble in the dialogue.

(18)

```
(define-cluster accept-information-about-case script
    :participants ((hearer)
                   (speaker))
    :props ((lawsuit))
    :preconditions ( (speaker '(:type attorney)))
    :events ((t0 '((:optional)
                   (:code (establish-relationship-of-lawyer-to-party))))
             (t1 '((:head)
                   (:code (action-taken-by-the-plaintiff))))
             (t2 '((:head)
                   (:code (elaboration-of-case-perspective))))))
```

In addition to the instantiation of the discourse script above, the semantic representation of the "desire to sue", the lawsuit frame from the parser, is bound as the value of the conceptual-frame for this discourse space, and in particular, it is of type misappropriation. The state of the discourse at this point (after the first sentence) is represented by the following discourse-frame and bindings:

(19)

```
(define-cluster legal-discourse-frame discourse-frame
     participants ((hearer 'COUNSELOR)
                   (speaker '((:type attorney)
                              (:infer from legal-rep attorney))))
     :hearer-goal ()
     :speaker-goal ((inform 'legal-rep))
     :discourse-mode ((mode 'expect-inform))
     :discourse-script ((script 'accept-information-about-case script))
     :conceptual-frame ((lawsuit '(:type $misappropriation))) )
```

At this point the system operates in a top-down expectation-driven mode, triggered by the value for the conceptual-frame slot. That is, $misappropriation is itself a script, and the best-first control strategy used by CICERO chooses to instantiate the script as part of its inferencing about the coherence relations in the (upcoming) text.

(20)

```
(define-cluster $misappropriate script
    "legal concept"
    :participants
            ((plaintiff-corporation '((:type corporation)
                                                    (:inherit thru parent lawsuit *)))
            (defendant-corporation '((:type corporation)
                                                    (:inherit thru parent lawsuit *)))
                                )
    props
            ((plaintiff-product '((:type product)
                                            (:infer from plaintiff-corporation product)))
              (defendant-product '((:type product)
                                            (:infer from defendant-corporation product)))
             (misappropriated-knowledge '((:type knowledge-about-a-product))))
    :preconditions ((t0 '((:code (produces plaintiff-corporation
                                                            plaintiff-product))))
                                (t1 '((:code '(used-in plaintiff-product
                                                            misappropriated-knowledge)))))
    :events ((t2 '((:code $legitimate-access-to-knowledge)))
                    (t3 '((:code (equal misappropriated-knowledge
                                                    (get-value defendant-product :knowledge-used))

                     (t4 '((:code $competitive-advantage))))
    )
```

This representation provides us with the logical arguments to a relation (the entailments), as well as a large set of presuppositions that will direct the inferencing—to establish the deep coherence— in later processing.

Notice that the discourse frame in (20) keeps a dual representation of the information streaming in from the parser. For structural bookkeeping purposes, the $misappropriate frame is bound to action-taken-by-the-plaintiff, in that it satisfies a particular structural property of such preamble paragraphs. For deeper semantic coherence, however, the same frame is bound to type

of a lawsuit, and carries the complex of information shown above in (/21.

There are two interesting aspects to the representation shown in (/):

1. Any inferences possible due to the presupposition-set of an utterance are computed by CICERO rather than the expert system.
2. The exact same representation is used for understanding text as for generating text.

4.0 Conclusion

I have sketched in this paper a very rough model of DA based on a level hypothesis, wherein the conflating factors of discourse interpretation have been teased apart. In the previous section I attempted to demonstrate a working system, CICERO, which is aware of these levels at the stages of analysis outlined above. The system, however, is still incomplete at this point, in that it fails to adequately simulate and model the speaker's belief space. Furthermore, the role of goal recognition as recovering the speaker's intention was minimal, due to the nature of the interaction in the domain. These topics are being addressed currently in our ongoing research.

Acknowledgements I would like to thank Kevin Gallagher, John Brolio, and Sabine Bergler for fruitful discussion on this topic. This work was supported in part by a grant from the Defense Advanced Research Planning A... contract no. N00014-85-K-0017.

REFERENCES

Allen, J. (1984) "Towards a General Theory of Action and Time", *Artificial Intelligence,* **23.2**

Allen, J. and C.R. Perrault (1980) "Analyzing Intention in Utterances", *Artificial Intelligence* **15.3**.

Ashley, K. and E. Rissland (1984) "Toward Modelling Legal Argument".

Barwise, J. and Perry, J. (1983) *Situations and Attitudes.* MIT Press.

Cohen, P. and C.R. Perrault, (1979) "Elements of a Plan-Based Theory of

Speech Acts", *Cognitive Science* 33. Ablex Publishing Corp.

Grice, H.P. (1971) "Meaning", in Steinberg and Jacobovits (Eds.) *Semantics: An Interdisciplinary Reader in Philosophy.*

Heim, I. (1982) *The Semantics of Definite and Indefinite Noun Phrases,* Unpublished PhD. Dissertation, UMASS.

Hintikka, J. (1971) "Semantics for Propositional Attitudes".

Hobbs, J. (1982) "Towards an Understanding of Coherence in Discourse", in Lehnert and Ringle (eds), *Strategies for Natural Language Processing.*

Karttunen, L. (1976) "Discourse Referents," in J. McCawley (ed.), *Syntax and Semantics 7,* New York, Academic Press.

Kamp, H. (1981) "A Theory of Truth and Semantic Representation", in Groenendijk, Janssen, and Stokhoff (eds.) *Formal Methods in the Study of Language.*

Kaplan, D. (1977) "Demonstrative: An Essay on the Semantics, Logic, and Metaphysics and Epistemology of Demonstratives and other Indexicals", Draft 2., ms. UCLA.

Lehnert, W. (1978) *The Process of Question Answering,* Lawrence Erlbaum Ass. Hillsdale, NJ.

Lehnert, W. and M. Ringle (1982) *Strategies for Natural Language Processing* (eds.) Lawrence Erlbaum Ass. Hillsdale, NJ.

Litman, D. and J. Allen (1984) "A Plan Recognition Model for Clarification Subdialogues", Proceedings of COLING-1984.

Pustejovsky, J. and K. Q. Gallagher (in preparation) "Clustered Objects as Knowledge Representation for Natural Language".

Schank, R. and Abelson, R. (1977) *Scripts, Plans, Goals, and Understanding.* Lawrence Erlbaum Associates.

Stalnaker, R. (1979) "Assertion," in J. McCawley (ed.), *Syntax and Semantics 7,* New York, Academic Press.

Webber, B. (1978) *A Formal Approach to Discourse Anaphora,* Ph.D. Dissertation, BBN Report No. 3761.

Wilensky, R. (1982) "Points: A Theory of the Structure of Stories in Memory" in Lehnert and Ringle (eds), *Strategies for Natural Language Processing.*

Wilks, Y. and Bien J. (1983) "Beliefs, Points of View, and Multiple Environments" *Cognitive Science* **7.2**