

CHAPTER 1

The Meaning of In-Depth Understanding

1.1 BORIS -- A Computer Program

The theory of narrative comprehension presented here is implemented within BORIS, a computer program. BORIS is able to answer questions about a narrative in order to demonstrate that it has read and understood it. Thus BORIS also incorporates a theory of memory search and retrieval for question answering.

What follows is an actual understanding and question answering session with the BORIS program running on a narrative concerning a divorce (called DIVORCE-1). Input to the program is in lower case; output in uppercase. BORIS receives the English shown here as verbatim input and produces verbatim English below as output:

I/O Behavior on DIVORCE-1

*(BORIS 'DIVORCE-1)

PROCESSING PARAGRAPH 1:

Richard hadn't heard from his college roommate Paul for years. Richard had borrowed money from Paul which was never paid back. but now he had no idea where to find his old friend. When a letter finally arrived from San Francisco, Richard was anxious to find out how Paul was.

READY TO ACCEPT Q/A:

What happened to Richard at home?
RICHARD GOT A LETTER FROM PAUL.

Who is Paul?
RICHARD 's FRIEND.

Did Richard want to see Paul?

YES, RICHARD WANTED TO KNOW HOW PAUL WAS.

Had Paul helped Richard?

YES, PAUL LENT MONEY TO RICHARD.

Why didn't Richard pay Paul back?

RICHARD DID NOT KNOW WHERE PAUL WAS.

How did Richard feel when the letter appeared?

RICHARD FELT GLAD BECAUSE PAUL AND HE WERE FRIENDS.

PROCESSING PARAGRAPH 2:

Unfortunately, the news was not good. Paul's wife Sarah wanted a divorce. She also wanted the car, the house, the children, and alimony. Paul wanted the divorce, but he didn't want to see Sarah walk off with everything he had. His salary from the state school system was very small. Not knowing who to turn to, he was hoping for a favor from the only lawyer he knew. Paul gave his home phone number in case Richard felt he could help.

READY TO ACCEPT Q/A:

What was the letter about?

PAUL AND SARAH WERE GETTING A DIVORCE.

Did Paul write Sarah?

NO, IT WASN'T SARAH, IT WAS RICHARD WHO GOT THE LETTER.

Why was Paul upset about the divorce?

PAUL AND SARAH WERE FIGHTING OVER THE FAMILY POSSESSIONS.

What did Paul do for a living?

PAUL WAS A TEACHER.

What did Sarah want?

SARAH WANTED TO HAVE THE CAR AND THE HOUSE AND
THE KIDS AND THE ALIMONY.

Why did Paul write to Richard?

PAUL WANTED RICHARD TO BE HIS LAWYER.

PROCESSING PARAGRAPH 3:

Richard eagerly picked up the phone and dialed.
After a brief conversation, Paul agreed to have
lunch with him the next day. He sounded extremely
relieved and grateful.

READY TO ACCEPT Q/A:

Why did Richard call Paul?

RICHARD WANTED TO ARRANGE A MEETING WITH PAUL.

Did Richard contact Paul?

YES. RICHARD CALLED PAUL.

How did Richard contact Paul?

BY PHONE.

How did Paul feel when Richard called?

PAUL WAS HAPPY BECAUSE RICHARD AGREED TO BE PAUL'S
LAWYER.

PROCESSING PARAGRAPH 4:

The next day, as Richard was driving to the restaurant, he barely avoided hitting an old man on the street. He felt extremely upset by the incident, and had three drinks at the restaurant. When Paul arrived Richard was fairly drunk. After the food came, Richard spilled a cup of coffee on Paul. Paul seemed very annoyed by this so Richard offered to drive him home for a change of clothes.

READY TO ACCEPT Q/A:

Why did Richard eat out with Paul?
RICHARD WANTED TO DISCUSS THE DIVORCE CASE WITH PAUL.

What happened to Richard on the way
to the restaurant?
RICHARD ALMOST RAN OVER AN OLD MAN.

Why did Richard get drunk?
RICHARD WAS UPSET ABOUT ALMOST RUNNING OVER
THE OLD MAN.

What happened at the restaurant?
RICHARD SPILLED COFFEE ON PAUL.

How did Richard feel when the coffee spilled?
RICHARD WAS UNHAPPY.

Why did Richard spill the coffee?
RICHARD WAS DRUNK.

When did Richard almost hit the old man?
WHILE RICHARD WAS DRIVING TO THE RESTAURANT.

Where did Richard have lunch with PAUL?
AT A RESTAURANT.

PROCESSING PARAGRAPH 5:

When Paul walked into the bedroom and found Sarah with another man he nearly had a heart attack. Then he realized what a blessing it was. With Richard there as a witness, Sarah's divorce case was shot. Richard congratulated Paul and suggested that they celebrate at dinner. Paul was eager to comply.

READY TO ACCEPT Q/A:

Why did Richard drive Paul home?
SO PAUL COULD CHANGE HIS CLOTHES.

What happened to Paul at home?
PAUL CAUGHT SARAH COMMITTING ADULTERY.

How did Paul feel?
PAUL WAS SURPRISED.

Why did Sarah lose her divorce case?
SARAH COMMITTED ADULTERY.

Why did Richard congratulate Paul?
PAUL WON THE DIVORCE CASE.

TEXT PROCESSING COMPLETED ON: DIVORCE-1

1.2 What BORIS Is Up Against

This section is included to give the reader some idea of the kinds of problems that arise in almost any fragment of narrative text. In this case the examples which appear below are taken from DIVORCE-1. Each problem raised in this section will be addressed at some point in the thesis, once a sufficient theoretical foundation has been developed.

(a) Richard hadn't heard from his college roommate Paul in years.

Here BORIS is told a non-event. What should BORIS do with this? Simply storing an event as not having occurred is inadequate. For instance, when people read: "John walked into the room and Mary was not there." they infer that John wanted to see Mary and had been expecting her there. They do not simply store in memory the fact: FALSE [LOC [MARY,OFFICE]].

(b) ... which was never paid back.

This phrase also refers to a non-event, but what is its effect on processing? Many readers said that Richard was willing to help Paul because Richard felt obligated to Paul. Therefore, this non-event must be understood in terms of the relationship between Paul and Richard. In addition, it should be noted that this phrase never explicitly mentions WHO failed to pay WHAT to WHOM, yet people never notice this, and automatically fill on these roles from knowledge of loans and their repayment.

(c) ... but now he had no idea where to find his old friend.

How do we go about parsing an expression like (c) and once parsed, how do we represent its semantics? That is, we must determine what representation the parser should produce. The word "had" does not refer to physical possession. The word "old" does not refer to the age of Richard's friend. How do we represent the meaning of "finding" someone? Also, we must understand the connection between (b) and (c), that finding Paul is an enabling condition for paying him back.

(d) When a letter finally arrived from San Francisco ...

People assume that the letter is from Paul even though this is never explicitly stated.

(e) Unfortunately, the news was not good.

What effect does "unfortunately" have on processing and how is

it represented? Furthermore, people assume that the entire second paragraph in DIVORCE-1 refers to information contained in the letter and that Richard is reading this letter. But we are never explicitly told: "Richard began to read and the letter stated the following: ..."

(f) Paul's wife Sarah wanted a divorce.

How do we represent a divorce in memory? At the very least, it must refer both to marriage and to the legalities required in terminating a marriage. Otherwise the subsequent mention of a lawyer would make no sense.

(g) ... but he didn't want to see Sarah walk off with everything he had.

Parsing and representation here is complex. Representing "walk off" as physical movement is inadequate. Similarly, "see" does not refer here to literal vision. Nor can "everything" be taken too literally. Metaphors and suggestive references must be handled with care.

(h) His salary from the state school system was very small.

What does (h) have to do with the story? People immediately realize the connection between a small salary and lawyers' fees, alimony, etc. But BORIS must explicitly establish this connection.

(i) Not knowing who to turn to, he was hoping for a favor from the only lawyer he knew.

How are "who to turn to", "hoping", "favor", "only" to be parsed and represented? Here BORIS must know about lawyers in order to understand why "lawyer" has been mentioned. Notice, also, that the story never explicitly states that Richard is the lawyer being referred to, yet people automatically make this inference.

(j) Richard eagerly picked up the phone and dialed.

The story never explicitly states who Richard dialed or engaged

in conversation. Furthermore, protocols of individuals indicate that many people infer at this point that Richard will agree to be Paul's lawyer.

(k) He sounded extremely relieved and grateful.

What is to be done with affects such as "relieved" and "grateful"? People use their knowledge of these affects to assign an appropriate referent to "he". Most people immediately assume that "he" is Paul, not Richard.

(l) ... he barely avoided hitting an old man on the street.

BORIS must realize that "hitting" here refers to a vehicle accident rather than a fist fight.

(m) When Paul arrived Richard was fairly drunk.

The story never explicitly states where it is that Paul has arrived. People nevertheless interpret Paul's arrival in terms of their arrangement to meet. But how?

(n) ... so Richard offered to drive him home for a change of clothes.

Why did Richard make this offer? Why does Paul have to change clothes? Connections must be made between these events to answer such questions.

(o) When Paul walked into the bedroom ...

What is Paul doing in the bedroom and how did he get there? Most scene changes are made only implicitly in narratives. Nor are we explicitly told that Richard drove Paul home, only that he offered to.

(p) ... and found Sarah with another man he nearly had a heart attack.

The story never explicitly states that Sarah was having an affair. This must be inferred, along with Paul's probable reaction. Only then does it make sense for "heart attack" to denote surprise, and not a cardiac arrest.

(q) Then he realized what a blessing it was.

What is "blessing"? What is the effect of this sentence on processing, and what is constructed in memory?

(r) With Richard there as a witness, Sarah's divorce case was shot.

BORIS needs to know something about witnesses to successfully understand this sentence. Also, the word "case" here is not a container or unit of measurement, as in "case of beer", and "was shot" is metaphorical.

(s) Richard congratulated Paul and suggested they celebrate at dinner.

How do we represent "congratulated" and why does Richard make his suggestion? BORIS must understand why congratulations are in order. But the story never explicitly states that Paul has won, simply that Sarah has lose. The implication is clear, but must be inferred nonetheless.

1.3 Knowledge and Memory for Comprehension

As the excerpts above suggest, understanding narrative text is a complicated task, requiring the interaction of many disparate sources of knowledge. The story DIVORCE-1 requires the following abstract constructs to represent its conceptual content: object primitives [Lehnert and Burstein, 1979] [Lehnert, 1979], scripts [Cullingford, 1981], [DeJong, 1979a] settings [Dyer and Lehnert, 1980], goals [Wilensky, 1978a], plans [Schank and Abelson, 1977], affects [Roseman, 1979], themes [Wilensky, 1978b], interpersonal relationships [Schank and Abelson, 1977], physical states [Schank, 1975], events [Dyer, 1981a], social acts [Schank and Carbonell, 1979],

MOPs [Schank, 1982a], and TAU's [Dyer, 1981b]. This thesis deals with issues of knowledge representation, knowledge application, memory instantiation, memory search and processing interactions within multiple knowledge source domains.

To understand narratives, one must have knowledge about the actions and situations which arise within them. A story involving a divorce, for instance, would be incomprehensible to one who did not already know something about divorces. But what does knowledge of divorce look like and how is it organized? Any theory of narrative understanding must include a theory of knowledge representation for the kinds of things narratives deal with.

But a representational system alone is not enough. A theory of narrative comprehension must also include processing knowledge -- i.e. knowledge which determines when and how representations are to be used. For instance, when we read:

Paul wanted to change his clothes. When Paul
walked into the bedroom...

we use our knowledge about clothing in order to understand why Paul went into the bedroom. To do this, we must know: a) to change clothes one must obtain new clothes, b) clothes are commonly kept in one's bedroom, c) to obtain an object, one must be near that object, and d) this may require walking into the room containing the desired object. There are other things we also know about clothes -- e.g. clothes are bought at stores, and they cost money. But this particular knowledge happens to be irrelevant in the example above. So a theory of narrative processing must explain how knowledge is organized during processing so that only knowledge relevant to the situation at hand will be applied.

These processes of comprehension must construct in episodic memory [Tulving, 1972] a conceptual representation of the narrative. Memory for a narrative will be composed of instantiations of many different knowledge structures that must be causally connected. For instance, when people read about a husband who catches his wife being unfaithful, they do not immediately think of lawyers. However, if a lawyer is mentioned next:

After George caught his wife committing adultery
George decided that he needed a lawyer.

they are not surprised. In contrast, consider the following fragment:

After George caught his wife committing adultery
George decided that he needed a basketball player.

In this case people become very confused because they can not find any relationship between adultery and basketball players.

In addition, access links must also be built for subsequent use by search and retrieval processes, particularly during question answering. If narrative episodes can not be appropriately accessed, then the result of the narrative comprehension process will be of little use.

Finally, narrative episodes do not reside in isolation in memory. When we read a narrative, often a related narrative will come to mind [Schank, 1982b]. This indicates that both narratives are being processed and indexed in a like manner at some level in episodic memory. Often, the level of indexing is at a very abstract level, dealing with very general themes. When this occurs, a story may remind a reader of another story which shares little at the content level, and yet is typified or characterized at a more abstract level by the same themes.

1.4 Natural Language Processing: Some Background

Research in text processing has grown explosively in the past ten years. Consequently, no effort is being made here to review all potentially relevant literature. Instead, the following section is intended to give the reader a perspective from which the scope and aims of the BORIS project can be examined.

1.4.1 Early Semantics-Based Approaches

Some of the earliest semantics-based approaches to natural language processing were represented by the SHIRDLU [Winograd, 1972] and MARGIE [Schank, 1975] systems:

SHIRDLU engaged in an interactive dialog concerning a micro-

world of blocks. It could answer questions about blocks and carry out simple commands, such as "Pick up the red pyramid on a blue block." SHRDLU represented the meaning of a sentence in terms of a procedure to carry out a set of actions within the blocks micro-world. For instance, "the red pyramid on a blue block" would translate into a program to examine each block until a blue one was found, and then check to see if a red pyramid was on top of it, and if not, continue searching for another blue block, etc. Unfortunately, this approach of procedural representation made it difficult to represent the meaning of a sentence outside the context of a pre-specified micro-world.

In contrast, MARGIE was based on a representational system intended to be independent of any particular micro-world. This representational system was based on a fixed set of primitive semantic ACTS which were related by a number of Conceptual Dependencies (CD)¹ [Schank and Abelson, 1977] [Schank, 1973]. MARGIE was composed of three modules: ELI [Riesbeck, 1975], MEMORY [Rieger, 1975], and BABEL [Goldman, 1975]. ELI parsed English sentences, producing CD representations. MEMORY then generated all of the inferences which arose from the conceptual representations produced by ELI. These inferences were themselves represented in CD. Finally, BABEL [Goldman, 1975] generated paraphrases by expressing each CD in English.

There were several novel aspects to MARGIE. First, ELI was driven by semantic rather than syntactic concerns. Second, ELI used the notion of "expectations" to aid in its conceptual analysis. That is, lexical items would access CD structures. Associated with each structure would be one of more expectations for what might follow. For instance, the primitive ACT of INGEST represented knowledge about putting substances inside one's body (such as eating, smoking, etc.). One expectation associated with INGEST looked for the occurrence of an edible object. When an expectation was satisfied, the expected entity would be connected to the CD structure associated with the expectation and an instantiation of an INGEST conceptualization would be built in memory.

Expectations are one way of representing an active context. Among other things, this active context can be used to aid in

¹For a brief discussion of CD theory, see appendix II.

disambiguating word senses. For instance, an expectation associated with INGEST arising from “ate” in the following sentence:

John ate a hot dog.

will select the food meaning of “hot dog” (rather than a canine with a high fever). However, MARGIE had great difficulty handling more than one sentence at a time. This occurred because MEMORY generated every possible inference it could. Since each inference would generate a CD, and since that CD had many potential inferences associated with it, very quickly MEMORY would be overcome by a combinatorial explosion of inferences. There seemed to be no way of constraining inferences to just those which were relevant to the text at hand.

1.4.2 Script-Based Understanding

The combinatoric problem was partially solved by the notion of a script [Schank and Abelson, 1977]. Scripts² contain pre-arranged causal chains which represented stereotypic actions associated with a given setting. For instance, a restaurant script includes the knowledge that people sit down, order, eat, tip, pay, and then leave. Once a script has been activated, the relevant inferences, concerning actions to follow, are acquired for free. A computer program, SAM [Cullingford, 1978], used scripts to understand simple stories. For instance, given the story:

John went to a restaurant. He ordered lobster from
the waitress. He left a big tip and went home.

SAM could use its script to fill in the relevant information not explicitly stated within the story. For example:

Q: What did John eat?

A: Lobster.

²Similar work has been conducted with the more general notion of a frame [Minsky, 1975] [Minsky, 1977] [Charniak, 1978].

Q: Who brought the lobster to John?

A: The waitress.

The restaurant script contains information the the diner usually eats what he orders and that the food is usually brought by a waitress or waiter. Thus SAM knew that John ate lobster even though the story had never explicitly mentioned this fact.

1.4.3 Goal-Based Understanding

The strictly script-based approach of SAM had serious limitations. These limitations were a natural consequence of what scripts had been originally designed to accomplish. Since scripts were intended to capture only very stereotypic knowledge, they lacked intentionality. For instance, although SAM knew that the waitress brought food to the diner, SAM did not know *why* she did this. As a result SAM was incapable of handling stories in which the characters' mental states lead them to deviate from the actions predicted by the script.

In order to understand novel situations when they occurred, therefore, it was necessary to track the goals and plans of the characters in the story. This approach was taken by PAM [Wilensky, 1978b], a program which explained the action of a character by using the character's goals to infer a plan which the character was using to achieve his goal. Consider the following text handled by PAM:

Willia was hungry. She picked up the Michelin Guide and got into her car.

A script-based program like SAM would be unable to connect up these three events unless it already had some HUNGER-READ-MICHELIN-GUIDE-DRIVE script. However, readers understand that Willia is using the Michelin guide to locate a restaurant to drive to even though it is the first time they have encountered this text. Instead of relying on a script, PAM used general knowledge about goals and plans to infer causal connections between physical actions. Regarding the story above, this knowledge included:

The goal of hunger is satisfied by eating.
One plan for eating is the restaurant script.
Michelin Guides tell the location of restaurants.
One plan for getting somewhere is to drive there.

PAM encoded general relationships between goals and plans. Furthermore, most scripts in PAM were represented as plans. Any action by a character was explained once PAM found a plan (compatible with the character's goals) of which that action was a part. By concentrating on inferences which connected up only goals and plans, PAM avoided some of the combinatoric problems that had arisen in MARGIE's mode of inference.

1.4.4 Text Skimmers

Both SAM and PAM tried to understand everything they read. They examined every word and attempted to build up as complete a set of causal and intentional connections as possible. Each program was capable of answering questions about stories read in order to demonstrate their comprehension.

In contrast, the text skimmers FRUMP [DeJong, 1979b] and IPP [Lebowitz, 1980] represent a different processing philosophy. Instead of answering questions about the text, each program displays its understanding by constructing a summary which represents the "gist" of a story. Unlike SAM and PAM, both FRUMP and IPP ignore words in the text. The resulting memory of a story often missed events and situations which occurred within the story.

Both these programs read numerous stories on the UPI (United Press International) news wire. FRUMP contains over 50 "sketchy" scripts, concerning such things as: earthquakes, kidnappings, visiting dignitaries, labor strikes, and breaking diplomatic relations.

These programs operate in a "top-down" manner. Once a knowledge structure is referred to in the text, this structure directs subsequent parsing strategies. For instance, when FRUMP finds a story which involves one of its scripts, it uses that single script to guide its analysis of the text. Given an earthquake "sketchy" script, FRUMP can extract from a UPI story the time, place, magnitude and number of casualties caused by the quake. Everything else, however, will be ignored.

Instead of knowing a little about many things, IPP knows a lot about one domain -- i.e. terrorism. IPP contains specific knowledge structures on extortion, kidnappings, bombings, skyjackings, shootings, etc. and their relations to one another. Unlike FRUMP, IPP maintains a long-term episodic memory of the stories it has read, and uses this memory as a basis for making generalizations about terrorism. For instance, if IPP reads several stories about attacks against the British by the IRA, then IPP will make the generalization that terrorist attacks in Britain are normally caused by the IRA.

The top-down approach of fulfilling prior expectations and selectively ignoring information not conforming to these expectations gives both systems a good deal of robustness. The main negative consequence is that unusual or unexpected information is often missed and not incorporated into a final memory representation of the story. This approach is adequate, however, for the task of extracting a sparse summary for stories.

1.5 In-Depth Understanding

In contrast to the text skimmers, BORIS attempts to understand narratives as deeply as possible. For BORIS, understanding a narrative "in-depth" means the following: 1) reading in a careful mode rather than skimming, 2) handling narratives which involve multiple interacting knowledge sources, 3) parsing text in an integrated fashion, where memory search and construction processes are evoked on a word-by-word basis, and 4) recognizing the key thematic patterns which characterize a narrative at very abstract levels.

Careful Reading -- BORIS reads each sentence (or phrase) in a left-to-right manner, without backing up over the text. Unlike skimmers, BORIS attempts to construct a complete representation of a narrative, including all physical events and mental states, along with the causal connections between them. As a result, BORIS operates in a very bottom-up manner. Processing is directed more from information arising in the input than from pre-determined expectations. Expectations in BORIS are encoded in the episodic memory of the narrative read thus far and are activated by search processes. This bottom-up approach gives BORIS the capability of

noticing unusual events, which are often missed by exclusively top-down processing approaches. It is the unusual and unexpected events, including the mistakes and failures of the characters, which often make a story memorable. By their very nature, such events can not be predicted in a top-down manner.

Multiple Knowledge Domains -- Earlier systems at Yale were largely "dedicated" to handling one or two knowledge structures. MARGIE dealt with Conceptual Dependency structures, SAM with scripts, PAM with goals and plans, FRUMP with sketchy scripts, OPUS [Lehnert and Burstein, 1979] [Lehnert, 1979] with knowledge about object primitives, etc. However, no single program existed which was capable of handling stories involving the interactions of all these different sources of knowledge, including novel sources, such as the affective reactions of narrative characters.

An initial attempt to handle complicated narratives had been made by simply applying ELI, PAM and SAM in a "round robin" fashion. This approach failed because ELI, PAM, and SAM each operated in isolation from one another and each lacked knowledge about how various knowledge structures interacted.

Parsing Integration and Unification -- ELI parsed sentences in isolation and then passed them to SAM or PAM. As a result, any information in SAM or PAM which could have helped ELI in its parsing tasks remained unavailable. It does not appear that people process text in separate phases [Marslen-Wilson et al., 1978] [Tyler and Marslen-Wilson, 1977]. Instead, they perform memory manipulations and inference tasks on a word-by-word basis.

With the advent of FRUMP, parsing was integrated with other processes to the extent that sketchy scripts in FRUMP directed the parsing task. Since the creation of FRUMP other projects at Yale have stressed the notion of integrated processing.

BORIS is a highly integrated system. All memory search, instantiation, and inference tasks occur as side-effects of a single, unified parsing process which occurs on a word-by-word basis. Furthermore, narrative questions are parsed by the same processes which handle the narratives themselves. One natural consequence of this integration is that BORIS often knows the answer to a question before it has completely understood the question. Another natural consequence is a "Loftus Effect" for narratives [Loftus, 1979] [Loftus,

1975]. That is, asking a question about a narrative may cause the memory of the narrative to be altered.

Thematic Level of Analysis -- Understanding a narrative “in depth” involves recognizing the moral or point of a narrative. This is analogous to being able to characterize the theme of a narrative in some appropriate way, as in selecting an apt title or adage for it.

In BORIS, abstract thematic patterns, such as hypocrisy, are handled by memory structures called TAUs (Thematic Abstraction Units). These structures arise when expectation failures occur, causing episodes to be organized around errors in planning. As such, they contain an abstracted intentional structure, which represents situation-outcome patterns in terms of: the plan used, its intended effect, why it failed, and how to avoid (or recover) from that type of failure in the future. This information is often expressed in terms of an adage, such as “the pot calling the kettle black”, and is abstracted from the specific content making up the episodes that each TAU organizes. This abstraction allows each TAU to organize episodes (which share the same failures in planning) across widely differing contexts. TAUs account for the phenomena of cross-contextual remindings, and serve as a way of sharing planning information in new domains.

1.6 Methodology, Scope, and Aims

As natural language programs tackle more complicated narratives, it becomes harder to characterize their domain of applicability. A simple metric, such as “this program correctly reads all stories 100 words or less in length” is unattainable. Therefore, it is important to briefly discuss the approach to natural language processing being taken in the BORIS project.

Psychological Validity -- Currently, people are the only examples of systems capable of understanding complicated narrative text. Therefore, BORIS tries to mimic the behavior exhibited by people who have performed the task of reading various narratives such as DIVORCE-1.

Specifically, subjects were asked to read a given narrative. After a short amount of time or an intervening task subjects were asked to answer questions about the narrative. Two weeks later they were

asked to recall the narrative in as much detail as possible. Others subjects were given the opposite task of first recalling the narrative and then two weeks later answering questions about the narrative. This data was then used to guide program design decisions.

Experience with BORIS has also lead to a number of psychological experiments made concerning various knowledge structures used in BORIS [Lehnert and Robertson, 1981], [Lehnert, Black, Robertson, 1982], [Seifert, 1981]. These experiments lend support to approaches taken in BORIS and will be discussed later.

Generality -- Since BORIS is not a skimmer, but rather an in-depth understander, it is not practical to hook BORIS up to the UPI wire, as was done for FRUMP. It might be quite some time before a story appears on the UPI wire which contains interactions of the same knowledge structures utilized in BORIS. However, a program which only reads one story might be open to the criticism of being ad hoc.

To demonstrate that the principles embodied in BORIS are general, BORIS is able to understand stories which are different than DIVORCE-1, but yet involve the same kinds of knowledge structures. For example, BORIS was able to read DIVORCE-2 (below) without any additional knowledge than was needed for DIVORCE-1.

DIVORCE-2

George was having lunch with another teacher and grading homework assignments when the waitress accidentally knocked a glass of coke on him. George was very annoyed and left refusing to pay the check. He decided to drive home to get out of his wet clothes.

When he got there, he found his wife Ann and another man in bed. George became extremely upset and felt like going out and getting plastered.

At the bar he ran into an old college roommate David, who he hadn't seen in years. David offered to buy him a few drinks and soon they were both pretty drunk. When George found out that David was a lawyer, he told him all about his troubles and asked David to represent him in court. Since David

owed George money he had never returned, he felt obligated to help out.

Later, David wrote to Ann, informing her that George wanted a divorce. Her lawyer called back and told David that she intended to get the house, the children and a lot of alimony. When George heard this, he was very worried. He didn't earn much at the junior high school. David told him not to worry, since the judge would award the case to George once he learned that Ann had been cheating on him.

When they got to court, David presented George's case, but without a witness they had no proof and Ann won. George almost had a fit. David could only offer George his condolences.

If one compares DIVORCE-1 and DIVORCE-2, it should be clear that they share many of the same knowledge structures. Both DIVORCE-1 and DIVORCE-2 make use of knowledge about divorce, lawyers, bedrooms, drinking, letters, phones, driving, friendship, etc. However, these two narratives also differ at many different levels: lexical, syntactic, semantic, and episodic.

Lexical/Syntactic Level

Different lexical phrases and syntactic constructs are used to describe similar events:

1. Spilling liquid:

DIVORCE-1: Richard spilled a cup of coffee on Paul.

DIVORCE-2: ...the waitress
accidentally knocked a glass of coke on
him.

2. Discovering adultery:

DIVORCE-1: When Richard walked
into the bedroom and found Sarah with

another man...

DIVORCE-2: He found his wife and
another man in bed.

Semantic Level

1. In DIVORCE-1, Richard learns that Paul needs a lawyer after having read a letter from Paul. In DIVORCE-2, David learns that George needs a lawyer after bumping into George at a bar. In neither narrative is the reader told explicitly that the lawyer has agreed to take the case. This fact must be inferred. In each story this inference is evoked by a different event:

DIVORCE-1: Richard eagerly picked up the phone
and dialed.

DIVORCE-2: ...he felt obligated to help out.

2. In both stories one character spills liquid on another. However, the consequences are very different:

In DIVORCE-1 it is Richard who spills something
on Paul, thus causing Richard to drive Paul home.

In DIVORCE-2 it is the waitress who spills
something on George in DIVORCE-2, causing
George to refuse to pay the check and then drive
himself home.

Notice that it would make no sense in DIVORCE-1 for Paul to refuse to pay the check. This is a result of the different social relationships in each story between the spiller to the one being spilled on.

Episodic Level

Both stories differ greatly in terms of the ordering and outcome of events. For instance:

1. Discovering adultery:

In DIVORCE-1 Richard catches his

wife in bed at the end of the narrative.

In DIVORCE-2 George catches his wife toward the beginning of the narrative.

2. Emotional Responses:

In DIVORCE-1 Paul “almost has a heart attack” when he discovers his wife in bed.

In DIVORCE-2 George’s “near fit” is the result, instead, of losing the divorce case.

3. Final Outcomes:

In DIVORCE-1 the husband wins although the characters never actually get to court.

In DIVORCE-2 the characters all appear in court and the husband loses.

It can be rather difficult to make up narratives which share exactly the same knowledge structures contained in DIVORCE-1, yet involve interestingly different interactions. Thus DIVORCE-2 both lacks some structures contained in DIVORCE-1 while containing some new knowledge not included in DIVORCE-1. Because of this, it was necessary to add some knowledge and processing rules to BORIS before DIVORCE-2 could be handled. Where the same knowledge structures are used, however, no changes were necessary.³ Given the dramatic differences between these two narratives, the success of BORIS in handling both stories seems to indicate that BORIS has achieved a level of generality.

³For examples of BORIS answering questions concerning DIVORCE-2, see appendix sections I.5 and I.6.

1.7 A Guide to the Reader

The organization of this thesis is based on the assumption that a reader's attention diminishes through time. I've tried to construct this thesis so that the major points and larger brush strokes are made toward the beginning of the thesis, while the more detailed, minor brush strokes occur toward the end. In addition, each chapter contains a section discussing related work. This organization is appropriate in a broad-based thesis dealing with interactions among many sources of knowledge.

The thesis is divided into four major parts. Part I deals with the role of abstract themes in narratives: i.e. Chapter 2 presents TAUs, which both account for one class of cross-contextual reminders during narrative comprehension, and supply an abstract level at which narratives can be characterized in terms of planning situations involving expectation failures. Chapter 3 discusses a categorization scheme of planning metrics. This scheme provides the foundation for the recognition of a large class of TAUs. Finally, chapter 4 presents a system of AFFECT representation and processing, including a discussion of how TAUs and AFFECTs are related.

Part II examines issues of process integration and their consequences for memory retrieval during question answering: Chapter 5 compares the way BORIS parses text at narrative comprehension time with parsing at question understanding time, and the resulting effects on episodic memory. Chapter 6 contains an overview of how different knowledge sources in BORIS interact during comprehension along with an overview of processing control.

Part III deals with a number of specific knowledge structures used in BORIS: Chapter 7 presents a system of intentional links which are used to build up larger knowledge structures and discusses their role in answer retrieval. This chapter also introduces MOPs (Memory Organization Packets) [Schank, 1982a] and their role in knowledge representation. Chapter 8 compares MOPs with scripts and examines a number of related MOP structures, including CONTRACTs and LEGAL-DISPUTE_s. Chapter 9 discusses the representation of space and time through the use of scenario mapping, while Chapter 10 deals with the role of the interpersonal level in narratives.

Part IV consists of an overview of what has been covered: Chapter 11 contains a detailed, annotated trace of BORIS reading DIVORCE-2, along with some sample question-answering and English generation traces. Finally, chapter 12 contains a summary of major claims, along with current BORIS limitations and a number of directions for future research.

The reader who wants an overview should read just Parts I, II, and the conclusions chapter. The reader who is interested in how specific knowledge structures were represented and applied in BORIS should also read chapters 7,8,9, and 10. Finally, those who want to know some of the nitty gritty concerning how everything came together can examine the annotated traces in chapter 11 of part IV. At the end of chapter 11 are included some program 'specs' concerning the current BORIS implementation.

Appendix I contains copies of various stories read by BORIS for easy reference, along with sample questions and answers generated by BORIS. The last section in appendix I lists the major memory structures currently implemented in BORIS.

Appendix II contains a brief discussion of Conceptual Dependency theory, including scripts, plans and goals. Readers totally unfamiliar with the Yale school of natural language processing can refer to this appendix in order to gain an initial introduction to the theoretical material upon which this dissertation was built.

Appendix III contains both a description and implementation of McDYPAR, a micro version of DYPAR (Dyer's Parser). McDYPAR is a demon-based conceptual analyzer written in TLISP, a dialect of LISP at Yale. In addition to implementation code, Appendix III includes a sample lexicon, sample execution trace, sample demon definitions, and a set of exercises for expanding the capabilities of McDYPAR.

PART I. RECOGNIZING NARRATIVE THEMES

There is more to understanding narratives in depth than being able to later answer questions of fact and causality. In-depth understanding entails recognizing what is interesting in a narrative; grasping the moral, point, or significance of a narrative; recognizing what makes it memorable -- including why, in the first place, it was a narrative worth telling. This deep level of analysis depends upon the recognition of thematic patterns in narratives. In this part of the thesis we present a class of thematic structures which are based upon abstract planning situations, and examine how such structures are recognized during narrative processing. We also explore the role of affect in narratives and its relationship to these thematic structures.

CHAPTER 2

Thematic Abstraction Units

2.1 Introduction

People often rely on common sayings or adages, when asked to summarize or title stories. Why do adages often serve as an effective way of characterizing a story? In this chapter we will present a memory construct, called a TAU, which serves as a basis for story characterization by means of cross contextual reminders and the recall of adages.

When asked to characterize the following story:

MINISTER'S COMPLAINT

In a lengthy interview, Reverend R severely criticized President Carter for having "denigrated the office of president" and "legitimized pornography" by agreeing to be interviewed in Playboy magazine. The interview with Reverend R appeared in Penthouse magazine.

readers often responded⁴ with adages such as:

Adg-1: The pot calling the kettle black.

Adg-2: Throwing stones when you live in a glass house.

Clearly, these adages are an effective characterization of MINISTER'S COMPLAINT. But how do we recognize this fact? By what process does an 'appropriate' adage come to mind, and to what purpose?

⁴Some of these responses were elicited in protocols. Other adages were generated spontaneously by subjects during summarization experiments.