

TEXT UNDERSTANDING:

A SURVEY

ROBERT YOUNG

Department of Computer Sciences
University of Texas
Austin, Texas 78712

Copyright © 1978

Association for Computational Linguistics

TABLE OF CONTENTS

Introduction	3
1. The Content of Connected Discourse	5
1.1 The Nature of Discourse Content	5
1.1.1 Barclay, Bransford and Franks	6
1.1.2 Frederiksen	8
1.1.3 Kintsch	10
1.1.4 Thorndyke	11
1.2 The Structure of Discourse Content	12
1.2.1 Bartlett	13
1.2.2 Crothers	14
1.2.3 Grimes and Meyer	15
1.2.4 Kintsch and van Dijk	19
1.3 Some Considerations Regarding Memory Experiments	25
1.4 Discussion and Conclusions	26
2. The Form of Connected Discourse	31
2.1 Structural Analysis of Text	31
2.2 Kintsch and van Dijk	34
2.3 Rumelhart	36
2.4 Mandler and Johnson	38
2.5 Thorndyke	40
2.6 Text Grammar	42
2.7 Discussion and Conclusions	44
3. Computational Models of Text Understanding	47
3.1 The Necessity of World Knowledge	49
3.1.1 Charniak	49
3.1.2 Rieger	56
3.1.3 Schank	60
3.2 The Organization of World Knowledge	67
3.2.1 Schank and Abelson	68
3.2.2 Phillips	76
3.3 Other Work	84
3.4 Discussion and Conclusions	87
4. Final Observations	91
References	92

INTRODUCTION

The goal of this study is to examine work that has something to offer toward the construction of a computable model of text understanding, and therefore toward cognitive models in general. The reason for the criterion of computability is that this rather strict requirement makes vague generalizations impossible (or at least more difficult), and forces exact specification of processes being hypothesized. (However, this criterion is frequently difficult to apply, and in some instances the author's decision was undoubtedly a subjective one.)

It should be clearly pointed out that no attempt is made here to deal with the semantics of individual sentences. A familiarity with recent work in semantic representation in cognitive psychology, linguistics and artificial intelligence is assumed (see Norman and Rumelhart [75], Schank and Colby [73] and Steinberg and Jakobovits [71]). Although many problems in representing the meaning of individual sentences remain unsolved, this study focuses on those aspects of meaning that are conveyed only by groups of connected sentences - texts. Additionally, only work that attempts to deal with the semantics or understanding of texts, as opposed to statistical or syntactic analysis, is considered. This focus on text has also led to the omission of studies of non-textual memory and of computational systems that understand and solve problems stated in English or carry on dialogue. This omission is not meant to imply that work in these areas has no relevance to text understanding, but the different focus and additional constraints of these studies make such implications difficult to isolate.

The study is divided into three parts, which occasionally overlap. The first section deals with the content of connected text - that is, what exactly does text communicate. The work in this section is primarily that of experimental psychologists interested in memory and recall. The second section deals with the structure of text, apart from its specific content. This work has been done mainly by cultural anthropologists and those interested in the theory of literature, but is also being investigated by experimental psychologists. The third section discusses computational models proposed by computational linguists, which include attempts to implement in computer programs some of the processes discussed in the other sections. Each section is concluded with a discussion of the emerging model of text understanding. Finally, some directions for needed research are suggested. It is hoped that this kind of interdisciplinary survey will call attention of workers in one area to work in other areas addressing the same problems. This type of communication is valuable in two distinct ways. When similar conclusions are reached in different disciplines frequently using different methods and with somewhat different goals, these conclusions must be given special credence. On the other hand, it is sometimes the case that one discipline completely overlooks a problem or some aspect of it due to their own interests and biases. Such omissions need to be brought into sharp focus.

1.0 THE CONTENT OF CONNECTED DISCOURSE

This section will present a brief survey of the recent work of some experimental psychologists concerned with the process of understanding connected discourse. The details of their experiments will not be presented, but the types of experiments and general conclusions will be discussed, as well as any models proposed by the investigators. These results seem to converge toward a single model, in spite of apparent contradictions. Psychological work on text understanding is relevant to computational models because humans are still the only good text understanders. Also, although some aspects of psychological research, such as forgetting, are not currently included in any serious way in computational models, it would appear that such models have a very useful role to play in pointing out where significant details of the process of text understanding are being handled intuitively by the theorist, and have not really been defined in the theory.

1.1 The Nature Of Discourse Content

The nature (general characteristics and features) of the meaning of connected text will be examined first. The principle focus of these studies is to compare the understanding of a text to the actual text and attempt to characterize the types of differences that are found.

1.1.1 Barclay, Bransford And Franks -

Barclay, Bransford and Franks in a large body of work (Bransford, Barclay and Franks [72], Barclay [73], Bransford and Franks [73], Bransford and McCartell [74] and Franks and Bransford [74]) argue for the existence of semantic representations independent of the surface input, and propose active processes operating on the input. They describe sentences as information which is used by the understander to construct a description of a situation. Furthermore, they claim that the information used to construct the semantic description is not wholly contained in the sentences, but that an understander uses his previous knowledge to a great extent. They do not make any definite proposals regarding the form of the semantic representations. Typical experiments performed to test these hypotheses consist of recognition or recall of compound sentences like:

1. Three turtles rested beside a floating log, and a fish swam beneath them.
2. Three turtles rested on a floating log, and a fish swam beneath them.

The physical situation described by the second sentence is essentially the same if the pronoun "them" is replaced with "it". This situation differs from that described in the first sentence, however. When subjects were presented with a sentence in which the pronoun substitution had been made, those who had heard the first sentence were able to make the distinction, while those who had heard the second made no distinction. The explanation offered is that subjects used their spatial knowledge to create situation descriptions, and

when two different sentences produce the same description, it is difficult to distinguish them.

Similar investigations used a sequence of sentences, each comparing two objects from a set of five according to some dimension (e.g. position, height, speed, weight). Subjects who knew that a single situation was being described had a better memory for the situation, but were less able to distinguish sentences they had heard from those implied by the described situation. Once again, the conclusion is that a non-linguistic representation is created when a meaningful text is understood.

A final group of studies involved descriptive texts that were ambiguous or difficult to comprehend without an appropriate context. The context might be a picture of the described situation or a meaningful title for the text. In every case the results indicated the understander attempted to build a description of the whole situation, and when this was very difficult or impossible, comprehension was low. (Related studies by Anderson, Reynolds, Schallert and Goetz [76] and Scallert [76] utilize texts, each with two totally different meanings. Evidence is obtained that one or the other alternative is almost always chosen for the entire text. The effect of subjects backgrounds on interpretations of these ambiguous texts is also studied, and a correlation between background and chosen interpretation is claimed.)

The principal conclusion of these studies is that the process of understanding text involves constructing a consistent unified meaning, which is not simply the set of individual sentence meanings. It differs in that

additional knowledge and organization is introduced by the understander.

1.1.2 Frederiksen -

Frederiksen [75a, 75b] continues this line of thought, treating the understanding of a discourse as the semantic knowledge that is acquired in listening to the discourse. This knowledge is acquired by processes that utilize prior knowledge, context, etc. and includes information which is inferred as well as that explicitly presented. Frederiksen also argues that the processes involved in understanding a discourse are directly related to the process limitations of the processor. The two processes that he discusses are overgeneralization and inference. Overgeneralization is the discarding of detail information resulting in a more general concept. He suggests that overgeneralization reduces the amount of information to be understood, thus reducing the processing load on the understander. Inferred information is information which is assumed to be true even if it is not explicitly present. This process is claimed to reduce the processing load by eliminating the necessity of completely understanding every sentence. Frederiksen performed neutral recall experiments and recall experiments in which the context - the task assigned to the subject prior to presentation of the story - was changed. His conclusions are that overgeneralization and inferences are incorporated into the semantic representation of the story as it is understood, and that the context can influence the amount of inferencing that is done during comprehension.

Frederiksen uses a simple, informally defined semantic representation in the above studies. It consists of set inclusion, identity, and logical implication relationships that hold among the concepts of the text. He indicates the need for a detailed model of the semantic representation, and in other work, Frederiksen [75c] begins to define such a representational scheme. He proposes the use of two networks: a semantic network and a logical network. The semantic network contains the representation of individual propositions while the logical network is composed of relations that hold between propositions that exist in the semantic network. Frederiksen's system is probably the most elaborate yet proposed using the basic ideas of case grammar represented by semantic networks. He begins with the fundamental ideas of object and action hierarchies, although his are quite large and contain numerous distinctions. He proposes a system that distinguishes sixteen verb cases, and a large number of relations specifying states, quantification location, manner, time, order, proximity, tense and aspect. His logical network consists of relations drawn from propositional logic plus causality, and he proposes several modal operators. The complexity of this proposal gives the impression of representational power, but it seems that Frederiksen is still primarily oriented toward the representation of individual propositions. Although he shows the representation of simple time and causally ordered actions sequences, he has not yet demonstrated the adequacy of his system for text, in general. And, he has utilized a number of elements of quantified modal logic without demonstrating their usefulness for modelling human understanding of discourse. His discussions do raise a number of interesting questions about representation, and certainly deserve

consideration in the development of any set of semantic types and relations.

1.1.3 Kintsch -

Kintsch [74] (reviewed by van Dijk [75b]) attempts to set forth a fairly complete theory of language understanding. Kintsch's basic representational unit is the proposition, and he includes discussion of the usual difficult problems of definiteness (including generic versus specific distinctions) of noun phrases, quantification, modality, implication and presupposition, location, time and tense. He proposes a text base which underlies discourses, but his revised and elaborated thought on this is discussed in Section 1.2. He describes a model of discourse processing in which he makes a distinction between episodic and semantic memory, argues that the processing operations must be well defined (he suggests pattern matching and completion, abstraction and generation) and argues that recall is essentially a different process from recognition in that recall requires input organization while recognition does not.

Kintsch reports a number of experiments undertaken to test aspects of his theory. One conclusion that he reaches is that the semantic representation is partially independent of the actual input sentences. One kind of experiment performed to test this was the performance of a common task by subjects who had read substantially the same material but in forms of varying complexity. In other experiments he notes that sentences containing multiple propositions are much less likely to be recalled (completely) than are single proposition sentences. This also supports the idea of an underlying representation. A

related conclusion is that the agent of a proposition is the most likely case to be recalled. However, Kintsch did find that in passive sentences, the subject, rather than the agent, was the most likely to be recalled. A final conclusion concerns the presence of inferred information in memory. Kintsch found that immediate recall showed some difference between implicit and explicit propositions, but that after twenty minutes or more the difference had disappeared and the two were indistinguishable to the subject. An interesting side result was the distinct difference in reading and recognition times for argumentative versus descriptive discourse. The argumentative discourse required significantly more time in both tasks.. Kintsch also includes reaction time experiments to study proposition retrieval, determination of the truth or falsity of general propositions and processing of complex lexical items but these will not be discussed.

1.1.4 Thorndyke -

Additional investigations on the role of inference in understanding text have been carried out by Thorndyke [76]. He uses compound sentences asserting a causal connection which is not familiar or obvious. For example:

The hamburger chain owner was afraid that his love for french fries would ruin his marriage.

Sentences like these are imbedded in a meaningful text, and are followed later in the text by a continuation sentence which references the previously mentioned relationship. This continuation sentence might be neutral like

He decided to see a marriage counselor in order to save his marriage.

or it might encourage one particular explanatory inference like

He decided to join weight watchers in order to save his marriage.

which strongly suggests the inference

He is fat from eating french fries.

Thorndyke uses recognition tests after the presentation of the story to compare inferences that have been reinforced by a continuation sentence with neutral and inappropriate inferences. He found the reinforced inferences much more likely to be recognized as part of the text than the neutral inferences, while recognition of inappropriate inferences was very unlikely.

Thorndyke suggests that although this evidence indicates that inferences are made and stored as part of the understanding of a text, a more important implication exists. This is that the role of inferencing is to aid in the integration of new information into the larger framework of the understanding of a text when no appropriate understanding could be obtained from only the explicitly stated information.

1.2 The Structure Of Discourse Content

Attempts to characterize the content of text necessarily lead to a structuring of that content. Simple one-dimensional representations in which propositions are connected only by co-reference or by time or causal ordering

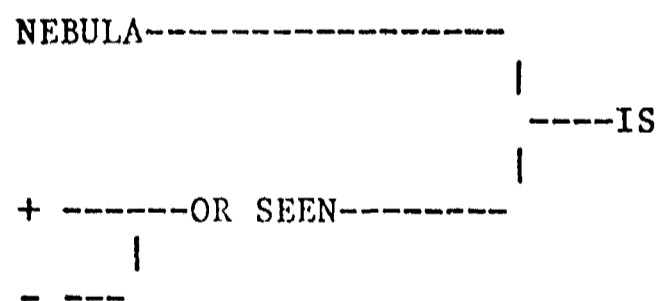
are adequate for only a restricted class of texts. The following studies all propose some type of multi-level representation of content.

1.2.1 Bartlett -

It seems appropriate to begin a discussion of memory structure with Bartlett [32] since he is often referenced as the originator of several currently popular hypotheses. It should be remembered that Bartlett was concerned with memory in general and the phenomena associated with recall, and did not restrict himself to the study of discourse. His principle conclusions were that memories are not stored in isolated, static units, and that exact recall is very rare. In fact, he often found cases of gross distortions in the recall of his subjects. He suggests that instead, memory is composed of a number of active, organized masses of past reactions and experiences, which he designates schemata, and a small amount of outstanding detail. Remembering is seen as a constructive process strongly affected by memories other than the one being retrieved. His ideas of organization beyond the sentence, of storage of something other than actual input sentences, and of active processes that modify text prior to its reproduction in recall are all currently enjoying wide acceptance.

1.2.2 Crothers

Crothers [72] is concerned with the recall of short, expository paragraphs containing material not likely to be familiar to his subjects. He presents results obtained with paragraphs about nebulae. He proposes the existence of a semantic representation underlying any particular discourse which contains the meaning of the discourse, but does not reflect the details of the surface form of the text. Thus, a single semantic representation might underlie numerous actual discourses. His semantic representation assumes a conceptual taxonomy showing the relationship of each known concept to its superordinate concept (i.e. the familiar semantic hierarchy), but he does not provide details on this. He also proposes an additional set of hierarchies showing the relationships of the concepts in a particular discourse. For example, the following:



means that a nebula is either seen or it is not seen, where all of the concepts are defined on the conceptual graph. Crothers does not carefully define his notation, so it is not clear exactly what may or may not occur at a node. Primarily, concepts (words) or connectives are used (e.g. NEBULA. IS AND, OR, WHY). Since he deals only with expository, descriptive material he finds no need to represent actions or time. Crothers performed recall experiments on two different versions of the same material, and arrives at three conclusions. First, he concludes that the surface paragraph is not a

significant factor in recall, thus arguing for a surface-independent semantic representation for the paragraphs. His last two conclusions are negative rejecting hypotheses he had previously suggested. The first of these is that superordinate nodes in the meaning of the paragraph will be recalled more frequently than subordinate nodes. The second rejected hypothesis is that information not directly connected to the most superordinate node of the paragraph is less likely to be recalled than information that is connected. Since his results do not support either of these hypotheses, Crothers suggests that other variables, such as frequency of concepts, are probably also of importance.

1.2.3 Grimes And Meyer -

Grimes [75] has proposed an extended case grammar supplemented with what he calls rhetorical predicates, which are higher-order predicates that take other propositions as their arguments. He subdivides rhetorical predicates into three groups: paratactic, hypotactic and neutral. Paratactic predicates always take arguments of equal weight (i.e. no argument is subordinate to any other argument). Hypotactic predicates have one dominant argument, to which the others are subordinate. Neutral predicates may be used as either paratactic or hypotactic predicates. The following is a list of some of the rhetorical predicates proposed by Grimes, and their basic meaning:

Paratactic predicates

- ALTERNATIVE - Options, or
- RESPONSE - Question and answer, problem and solution

Hypotactic predicates

- ATTRIBUTION - Gives qualities of the dominant proposition
- EQUIVALENT - Gives restatement of the dominant proposition
- SPECIFIC - Gives more specific information about a general dominant proposition
- EXPLANATION - Gives an abstract explanation for the specific dominant proposition
- ANALOGY - Gives an analogy to support the dominant proposition

Neutral predicates

- COLLECTION - List of elements related in some unspecified way
- COVARIANCE - Causality, an antecedent and consequent

Meyer [75] uses Grimes' rhetorical predicates and most of his case grammar to create analyses of paragraphs she uses in recall experiments. For example:

RESPONSE

PROBLEM

COLLECTION

need to generate electric power

protect environment

rational utilization of natural resources

SOLUTION

breeder reactors

would be the essentials of an analysis of a paragraph which stated that the need to generate electric power while protecting the environment and rationally utilizing natural resources is a problem which breeder reactors solves. Meyer indicates the hierarchical structure through the use of indentation. Since the rhetorical predicates in the above example were all paratactic, no indentation was used. However, the following example:

finite reserves of natural resources

SPECIFIC

COLLECTION

coal

oil

gas

shows the subordination of the specific set of resources to the general idea. Meyer's analyses of expository paragraphs using this scheme always follow the author's organizational structure and always result in a purely hierarchical structure like those depicted above. Meyer also recognizes a class of sentences that can occur in expository paragraphs, but do not contribute any

content to the passage. She calls these signalling because their function is to explicitly indicate the structure of the passage. She notes four types of signalling:

1. Specification of the structure (e.g. "Two options exist.")
2. Prematurely revealed information abstracted from the remainder of the passage (e.g. "The alternatives are solar energy, nuclear energy, geothermal energy and laser fusion energy.")
3. Summary statements
4. Pointer words (e.g. "unfortunately", "an important point is")

Meyer conducted recall experiments varying the position of certain material in the content hierarchy of passages, and including or omitting certain signalling information. Her principle conclusion is that material structurally higher in the content hierarchy is remembered substantially better than identical information placed lower in the hierarchy in another passage. She also notes that passages with identical structure but totally different content exhibit very similar patterns of recall at the higher levels. However, at the lower levels, the pattern of recall varied, indicating content dependence. In regard to signalling, she concludes that it has very little effect at the top level, but seems beneficial at the middle levels.

In regard to a theoretical explanation of why the highest ideas are recalled best, Meyer suggests three proposals and points out the weaknesses of each. First, lower propositions might be subsumed by higher ones with the passage of time. Immediate recall experiments show that the phenomenon occurs even without a time lapse, however. Secondly, it is possible that all propositions are stored but that retrieval is easier at the higher levels. She criticizes this proposal because even cued recall experiments were unable to retrieve the lower propositions. Finally, perhaps only higher ideas are ever stored. But her results show that the loss of lower level propositions begins immediately, but continues with time at a more rapid rate than for the higher level propositions. Meyer suggests that probably some combination of these processes is occurring, as well as other processes sensitive to the structure of the passage as a whole.

1.2.4 Kintsch And Van Dijk -

Kintsch and van Dijk (van Dijk [74,75a,76], van Dijk and Kintsch [forthcoming] and Kintsch and van Dijk [76]) present a model for the organization of discourse as a whole, and a number of (sometimes informal) experiments attempting to validate the model. Beginning at the lowest level, a discourse representation consists of a set of propositions. A distinction is made between two different types of representation, called text bases, as to whether all implied information is made explicit or whether it is left implicit. The notion of coherence is introduced, which is the property that distinguishes a discourse from a random set of sentences. Referential

identity has been suggested as a major test of coherence, but it is not an adequate definition (see discussion of coherence and Bellert's proposals in Section 3). Kintsch and van Dijk argue that coherence is more accurately captured by the requirement that each proposition of the text base be connected with one or more preceding propositions. Propositions are connected if one is a condition for the other, with the strength of the connection ranging from possible to necessary. Thus, an explicit text base is one containing all of the propositions necessary for coherence, while an implicit text base is one with some of these propositions deleted. The deleted propositions are those that can be assumed known or which normally would be inferred by the understander. Different types of discourse would have different rules governing the deletion of propositions. For example, casual conversation would allow more deletion than careful argumentation.

The text base is organized hierarchically under macro-structures, which are higher order propositions. Macro-structures may be related to their propositional arguments by a number of macro-rules. Four of these rules are suggested. The first, information reduction, is a rule of generalization which would explain the existence of the macro-structure

John is ill.

which had as its arguments propositions like

John has a fever.

John has the flu.

The second rule, deletion, would explain the relationship between

Peter saw a ball.

and its subordinate propositions

Peter saw a ball.

The ball was blue.

The third rule, integration, combines a central event with its normal pre-conditions, results and component actions. Thus

John went to Paris.

might have as its arguments

John took a cab to the station.

He bought tickets.

He went to Paris.

The fourth rule, construction, explains the relationship between a complex fact and its component parts. For example, the macro-structure

Peter built a house.

might have as its arguments the event sequence

Peter laid a foundation.

Peter built walls.

Peter built a roof.

In general, two conditions hold for all macro-structures. First, the

macro-structure must be semantically implied by its micro-structure (i.e. its propositional arguments). The exact meaning of the term implication is not clearly stated by Kintsch and van Dijk. They sometimes refer to it as entailment and treat it as a formal logical relation, but at other times say that it is not logical in the strict sense. In any case, it is clear that there are semantic rules or structures that allow creation of macro-structures such as those above. The second condition is that the sequence of macro-structures representing a coherent text base must itself be coherent. Recalling the definition of coherence, this implies that no macro-structure may delete information contained in its micro-structure which is a condition for another macro-structure. An important consequence of this is that the macro-structures of a text, taken by themselves, form a coherent summary of that text.

A final component of a text representation is the specification of the form or structure of the discourse, but discussion of this proposal will be postponed until Section 2. Their investigations are primarily with narrative text, and they call the narrative structure a schema.

Kintsch and van Dijk suggest that the concept of narrative structure combined with the idea of macro-structures leads them to the following hypotheses regarding comprehension and recall of narrative. First macro-structures are primarily what is stored when a text is understood. Recall uses macro-structures as a starting point in retrieval, and summaries directly reflect the macro-structures. Secondly, since macro-structures are essential to comprehension, they must be constructed at the time of reading.

Finally, a narrative schema is necessary for the organization of the text representation. Kintsch and van Dijk have done a number of summary, recall and other experiments to test these hypotheses. Most use a 1600 word text from The Decameron. In recall experiments, the propositions corresponding to the macro-structures were found to be recalled most often, and were very unlikely to disappear in delayed recall (nine days later). In contrast, many other propositions recalled immediately were omitted in the delayed recall. The propositions most likely to be recalled have the following functions:

1. Introduce main characters
2. Give major goals of characters
3. Describe actions leading to these goals
4. Describe events occurring to these characters leading to or from their goals

while propositions having the following functions are likely to be forgotten:

1. Setting description
2. Preparatory actions
3. Mental actions
4. Component actions

5. Probable consequences of an action or an event

6. Meta-narrative statements by the author

Summarizing experiments showed that summaries of the story were very much like delayed recall. When a summary was written after presentation of each successive part of the story, followed immediately by a complete summary, propositions were included in the partial summaries that were omitted in the final summary. A final group of recall experiments compared recall of a 70 word paragraph in isolation, to cued recall of the same paragraph imbedded in an 850 word text. Surprisingly, the recalls were almost identical. But when the entire text was recalled, the reproduction of the particular paragraph was much smaller and less accurate than in the first two cases. Their conclusions from these results are that the recall of small amounts of text is a different process from the recall of a long text, that recall of a long text relies on the macro-structure of the text as a means of organizing the text, that the micro-structure is forgotten much more easily and that summarizing is based on the macro-structure.

Finally, some experiments were done in which incorrect summaries were presented prior to the presentation of the story. These were found to have practically no influence on the final understanding of the story, and the subjects were unable to accurately recall the incorrect summaries.

1.3 Some Considerations Regarding Memory Experiments

Before attempting to describe an informal model based upon these hypotheses and experimental results, some general criticisms of memory experiments expressed by Spiro [75] deserve consideration. Spiro's principle argument is that recall consists of active reconstruction processes, rather than passive processes which merely reproduce that which was stored at the time of comprehension. His general position is thus much like that of Bartlett, and he observes that almost all experimenters since Bartlett have failed to replicate his findings of significant errors in recall. As a result, psychologists have tended to concentrate on the process of understanding, or construction, and have treated recall as something of a simple retrieval process. Spiro argues that for reconstructive errors to occur, it is necessary that the input be understood in terms of some pre-existing schema. Any text that results in the creation of a new schema would not be subject to the same kind of effects of previous knowledge. The interference caused by the pre-existence of schemata is even more pronounced if other uses of the schema are made between the time of comprehension and the time of recall. Spiro also points out that any experimental subject would be unlikely to integrate material read or heard in an experiment into his general knowledge since its truth and source are unknown. This coupled with the desire to perform well, could easily result in a considerably modified process of understanding text in experimental settings. Spiro performed a number of recall experiments which support his hypothesis. He used text about interpersonal relationships and instructed the subjects that the experiment was concerned with their reactions to the incidents described, thus maximally

involving their pre-existing structures in the understanding process. Spiro found substantial errors in the subjects recalls, which he explains as the effect of interaction with pre-existing structures for interpersonal relationships. (Kintsch and van Dijk [75] found similar results using stories from the Bible.) Spiro concludes that most text recall experiments have not shown these effects because they did not meet the required conditions, but frequently involved the presentation of material that was totally new and which was kept isolated by the subject. These criticisms not only are relevant to the question of construction versus reconstruction, but also to inferring discourse organization from recall experiments, and indicate that one must be very cautious in generalizing from the results of text recall experiments due to the many interactions between the subject's knowledge, the experimental setting, the type and content of the text and the extra-experimental effects (or lack of them) in long term experiments.

1.4 Discussion And Conclusions

A great deal of agreement is seen in the preceding studies. The idea that the meaning of a text differs from the meaning of its individual sentences, in that prior knowledge is used to infer implicit information during comprehension, is generally accepted. The idea that this meaning must be well structured, or coherent, is also expressed directly or indirectly by most of the researchers. However, explicit disagreement exists on both the nature of content structure, and its relationship to recall. Crothers and Meyer disagree as to whether a neutral structure reflecting pre-existing

knowledge is built, or the author's structure used. Kintsch and van Dijk essentially use both approaches since the time ordering in narrative is author defined (although the concept of time is pre-existent), but the macro-structures which hierarchically organize the text are pre-existent. Meyer, Kintsch and van Dijk all agree that propositions higher in the structure are more likely to be recalled, but Crothers evidence did not support this conclusion. Clearly, this hypothesis, if true, could be confirmed only if it is tested against the appropriate content structure. Spiro's comments regarding the use (or lack of use) of pre-existing structures are relevant. In the case of unfamiliar expository material, such as that used by Crothers and Meyer, it seems reasonable to suppose that the author's organization would be used to organize the semantic representation due to the lack (or non-use) of any appropriate pre-existing content structures. Thus Meyer's results support the importance of propositional level, and Crothers results do not. However, a familiar structure, like an action sequence, is understood in the same way regardless of presentation setting. Thus, Kintsch and van Dijk found that the role of a proposition in the pre-existing structure (the macro-structures) was primary in determining its importance. It should be noted that even Meyer found variations in the recall of mid-level propositions that she could not explain by reference to the author's organization. It is plausible to assume that these variations were caused by individual differences in prior knowledge, resulting in differing content structures.

In light of this, it seems plausible to propose a general model that distinguishes the following component processes of text understanding (realizing that like almost all distinctions, such as syntax versus semantics, they are really inadequate, but useful, oversimplifications). First there is the process of comprehension of the surface structure which builds a representation that is independent of the surface form of the text. This representation contains as much inferred information as is necessary to meet some minimal level of understandability, or coherence. This representation is organized hierarchically by higher level content structures which summarize or generalize over a larger amount of detailed information. These higher-level structures would normally be pre-existing (i.e. known to the understander) although a text could result in the creation of new structures. Different types of texts could certainly differ in the complexity of each of these tasks, explaining the variation in difficulty of understanding.

The process(es) explaining forgetting, and the integration of new information into prior knowledge, operates on this text representation. This process is poorly understood, and there are no well defined hypotheses about its operation. It appears to operate primarily on the content structure of a text. So, for example, Spiro's subjects integrated information from the stories they read into their general expectations about love and courtship, and lost the specifics of the text, even recalling it with radical modifications. This process would explain the almost total loss of detail information with time, as well as the interactions and distortions caused by previous knowledge.

Finally, recall is a retrieval process which operates on the current form of the representation. If that form is relatively unchanged since comprehension, recall may be almost distortion-free. But if significant changes have occurred, recall will be distorted. As suggested by Kintsch and van Dijk, summary is seen as recall to a controlled depth. Retrieval accesses the higher-level organization first, and uses these structures to access more detailed information. As retrieved information is generated as output text, a test of coherence would be made. If something has been forgotten or changed so that a necessary explanation or link is missing, the retrieval process would supply a probable piece of information from the general pre-existing structures. This, along with modifications from integration into prior knowledge, accounts for the reconstructive aspect of recall.

This general model does not account for the recall of very specific surface details (e.g. words and constructions used in the input text) but as suggested by Kintsch's work, it seems best to treat this as a process separate from the general understanding of text. One other inadequacy of the model is that it specifies complete comprehension of all input propositions, including assignment to an appropriate role in a higher level structure. Frederiksen suggests that overgeneralization is a loss of detail which reduces the processing load. The idea that partial processing is sometimes done on input information is also suggested by Bobrow and Norman [75]. They describe such processes as resource-limited. This kind of processing seems to be indicated by facts such as loss of detail in immediate recall and the subjective impression of partial comprehension in a number of situations (such as reading very rapidly, reading while partially distracted or reading very complex

material). However, while the idea of discarding comprehended information could be described computationally, selective comprehension is a process that requires further investigation.

One of the principle contributions of computational investigations with this sort of model will be specification of the processes of comprehension, which can only be done by specifying the nature of pre-existing memory structures. Meyer's work is an example of the use of undefined basic structures. She criticizes Crothers for requiring two graphs to represent a text one of which is his fundamental hierarchy of known concepts. She, instead, simply doesn't define any of her predicates first or higher order, apparently taking them as primitive. Assuming that one attempted to define even the argument roles for the rhetorical predicates, enormous difficulty would be encountered. Advocates of case systems have always had difficulty defining the exact role of cases, or delimiting the class of entities that could fill a case role. Consider how much more difficult it would be to define what pairs of things could exist in the RESPONSE or COVARIANCE relations, or exactly what the characteristics of these arguments are. Kintsch and van Dijk attempt to set out much more well defined structures, but even they have little to say about the set of structures that must exist, what information they must contain and how the comprehension algorithms use the text and the structures to build higher-level structures. The section on computational models will consider algorithmic attempts at text understanding, and should clearly indicate the complexity of these issues.

2.0 THE FORM OF CONNECTED DISCOURSE

The general model of text understanding suggested in the last section describes the structured content of text representation. This section contains work investigating a similar, yet distinct, type of structure - the abstract, underlying form of a text. (Not everyone would accept this distinction, but it seems analogous to the distinction between the syntactic structure of sentences and the semantic structure of their underlying meaning.) Much of this work has been done by cultural anthropologists or by those interested in literary analysis. Little of the traditional work has been done with cognitive or computational models in mind, and hence tends to rely on intuitive understanding of the terms. This less computational work will be presented quite briefly, since the presentation is intended to suggest the kinds of results that have been obtained, rather than the details of these conclusions. It is necessary to define these ideas in a computational way that can be integrated into the models of text understanding, and some recent work has begun this task.

2.1 Structural Analysis Of Text

Most modern work in the structural analysis of texts has roots in the work of the Russian structuralist, Propp [68]. Propp was concerned with the form of Russian folktales. and developed a method for describing the similarities he found in a corpus of 100 folktales. The major structural unit that he developed, he designated a function. These functions act like meta-actions in that they describe major events or event complexes that were

repeatedly found in the tales. For example, Absentation is the function that describes the situation in which one of the members of the family absents himself from home, and Trickery is the function which describes a villain's attempt to deceive his victim in order to take possession of him or his belongings. Propp found thirty-one such functions to be adequate to describe the events in the tales that he studied. Importantly, the ordering of the functions was found to be fixed. Some functions are optional, but if they are present their position is fixed with respect to the other functions. All of the functions are defined in terms of actors that participate in the situations they describe. Propp did some analysis of the restrictions on the assignment of various characters to these functional roles. He notes that the same character is likely to play a fixed set of roles, which he designates a sphere of action. These spheres of action include such familiar story roles as hero, villain and false hero. Finally, Propp notes that a single folktale may be composed of a sequence of basic tales. He designates the basic tale a "move". Thus, a folktale may be a one-move or a multi-move tale, with each move conforming to the definition of a tale. These rules capture the similarity of the folktales which could differ widely in the details of characters, setting, specific actions, etc. It is clear that Propp's rules could be expressed formally, and that is what Klein et al [74] did by writing a computer program to generate the function sequences of a number of folktales, using Propp's definitions.

The ideas of Propp have been refined and generalized by Barthes, Bremond, Greimas and Todorov (their work is not generally available in English but is reviewed in van Dijk [72] and Weinold [72]). Their work has been directed

toward the redefinition of Propp's functions as propositions, and of actors as case relations of these propositions. They have also attempted to generalize the functions by making them less specific, and to apply them to other forms of texts. However, Hendricks [72] demonstrates the flexibility of Propp's original functions by using them to analyze part of the Bradbury novel Something Wicked This Way Comes.

Chafe [72, 75] discusses the process of "verbalization", by which he means the translation of non-verbal knowledge into verbal output. One of the proposed component processes of verbalization is breaking the knowledge into smaller chunks according to patterns, which are called "schemata". These schemata are grammar like descriptions of possible verbal sequences. Thus, he notes that a certain class of stories will always be of the form:

PLOT + MORAL

Chafe also mentions such schemata as TRICK and VISIT as occurring in the stories he has analyzed. He explicitly states that, at this time, schematic analysis of a story is done by "imagination and intuition". He discusses aspects of the schematic analysis of several simple tales, but these are mostly illustrative rather than complete, in any sense.

Colby [73] analyzes Eskimo folktales, identifying "eidons", which are similar to Propp's functions, and producing a grammar capable of generating some of these tales. Colby observes that most native speakers are probably not very familiar with story grammars of this type (or the knowledge they represent), since only a few individuals are able to generate such tales. Nonetheless, it seems reasonable to assume that the same type of structure can

be found in more conventional stories

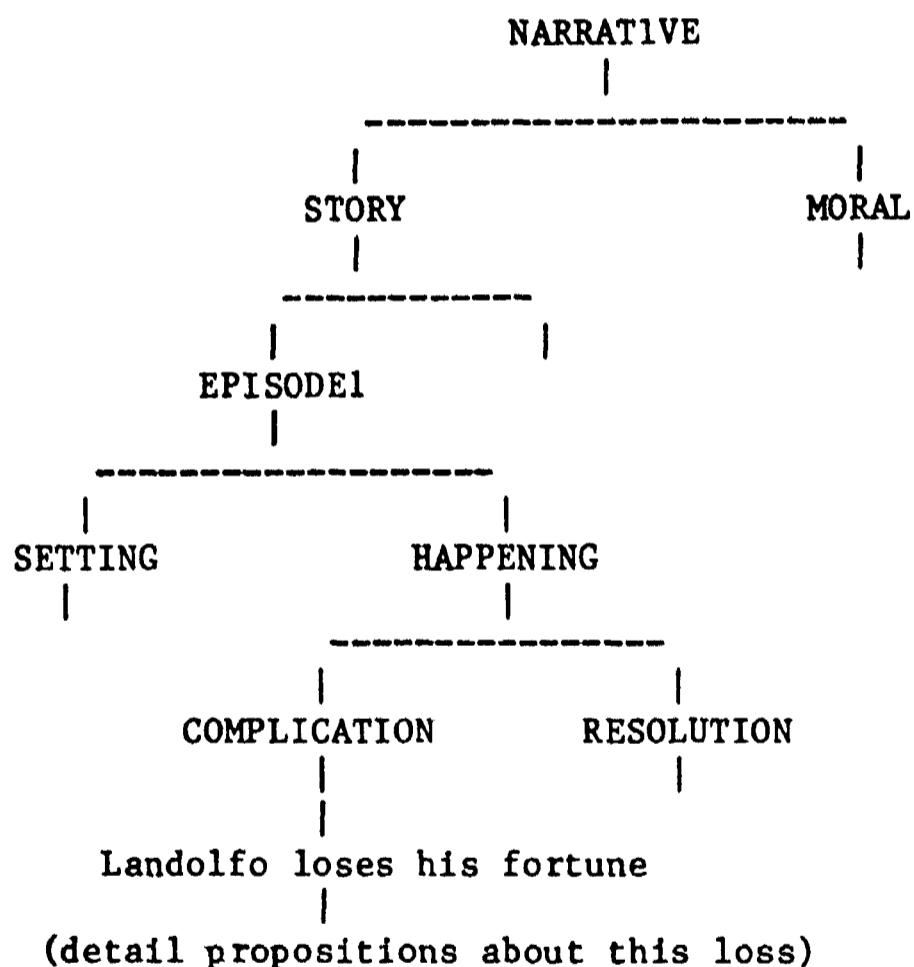
Although most work in this area has involved analysis of some form of traditional narrative, Labov and Waletzky [67] present analyses of oral versions of personal experiences. Even in this informal type of narrative, they find five regular components: an orientation, a complication, an evaluation, the resolution and a coda or moral. The last two components are optional. A totally different type of material is examined by Becker [65] who discusses patterns in expository paragraphs.

All of these analyses rely on the intuitive understanding of the analyzer to grasp the meaning of the structures from informal descriptions and to find these structures in the text being analyzed. For this reason, the particular structures that have been suggested are of less interest than is the general result that quite regular high level patterns are found in texts of many types, which may be characterized independently of the specific textual content.

2.2 Kintsch And Van Dijk

In addition to the content macro-structures (discussed in section 1.2.4) Kintsch and van Dijk [75] and van Dijk [75a, 76] discuss even higher level "super-structures", used to organize the content of a text according to the type of the text, which are derived from the work in structural analysis of texts. They limit their work to narratives, which they define as a specific type of action discourse. Narrative categories such as Episode Setting,

Complication, etc. are super-structures under which are organized either more narrative categories or the content macro-structures of the text. Van Dijk [75a] presents a fairly complete example. It includes a 1600 word text from The Decameron, a propositional analysis of the story, the content macro-structures of the story and the narrative categories under which the macro-structures are organized. The following abbreviated fragment should give an indication of the kind of analysis being proposed.



Some experiments were done to test the role of the understander's notion of narrative structure. An Apache myth and three Decameron stories, which were of comparable difficulty at the sentence level, were presented. They differed in that the Apache myth had an organization not familiar to the subjects. There was much greater variety in the propositions different people

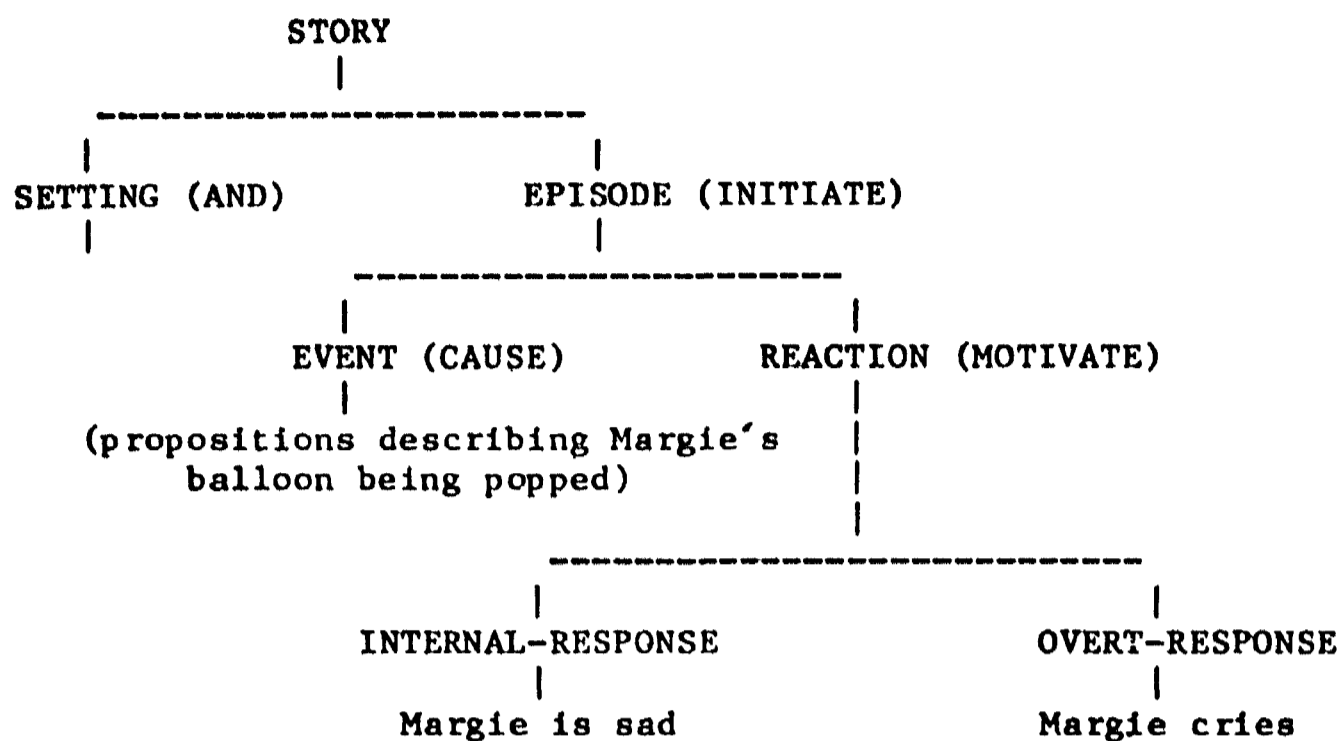
used in the recall of the myth than in the other three stories. Another experiment compared recalls of subjects who had read a normal story to those of subjects who had read scrambled versions of the same story. The recalls were indistinguishable by judges. Both of these groups of experiments indicate the active role of one's expectations about the form of a discourse in organizing the representation of that discourse. When this is impossible, as in the Apache myth, recall is less organized and more random.

2.3 Rumelhart

Rumelhart [74, 75] proposes a story grammar that is capable of producing structures similar to those described by Kintsch and van Dijk (section 2.2) His grammar contains rules like the following:

1. STORY -> SETTING + EPISODE
2. EPISODE -> EVENT + REACTION
3. REACTION -> INTERNAL-RESPONSE + OVERT-RESPONSE

Using these rules, he describes the syntactic structure of simple stories. The following fragment illustrates the types of structures derived (ignoring the parenthesized predicates for now):



Associated with most syntactic story rules are semantic rules which are used to generate a corresponding semantic structure for the story. The semantic rules for the syntactic rules given above are:

1. ALLOW (SETTING, EPISODE)
2. INITIATE (EVENT, REACTION)
3. MOTIVATE (INTERNAL-RESPONSE, OVERT-RESPONSE)

where the semantic predicates are intended to mean what is suggested by their names (e.g. the SETTING ALLOWS the EPISODE to occur). The figure given above contains the appropriate semantic predicate, in parentheses, at each node in the syntactic structure. (Rumelhart uses a completely separate semantic structure.)

Rumelhart discusses some elementary summarization rules that operate on the semantic structure of a story. Some illustrative rules are:

1. MOTIVATE (thought, response) => response
2. INITIATE (X, Y) => Y
3. INITIATE (X, Y) => Y (because | when | after) X

These rules could produce either of the following summaries (for the EPISODE in the story fragment given above):

1. Margie cried.
2. Margie cried because her balloon had popped.

Rumelhart's ideas are intended to be primarily illustrative ("a tentative beginning of a theory"), and deal with very simplified stories. They do argue clearly for the notion of both a syntactic and a semantic structure for the representation of a story's meaning. The syntactic structure allows production of the semantic structure, which is justified by its usefulness in producing summaries.

2.4 Mandler And Johnson

Mandler and Johnson [77] discuss an extended form of Rumelhart's story grammar, which they propose as a basis for text recall studies. The extended

grammar allows multiple episodes and does not utilize the additional structure containing the semantic interpretation which Rumelhart suggests. Mandler and Johnson feel that this second structure is unwieldy and frequently redundant. Conjunction, time-ordering and causality are included as categories in the grammar. They observe that restrictions imposed on a story by this type of grammar include allowing only a single protagonist per episode, and prohibiting the realization of higher level nodes in the actual text. They demonstrate the utility of their grammar on several stories, including "The War of the Ghosts". They also note that a transformational component remains to be developed which would provide for deletions and reorderings of the ideal story structure. They discuss a set of predictions for recall studies, which are suggested by this definition of story structure. These include:

1. Accuracy will be better as the story is complete and ordered as defined by the ideal structure.
2. Elaborations will be poorly recalled.
3. Optional nodes' realizations will be less well recalled.
4. Causality will be better recalled than simple time-ordering.
5. Inversions should be fewer as the story is closer to the ideal structure.
6. Omissions and violations of the ideal structure will be reflected by additions and distortions.

Unfortunately, the only experimental work reported is a comparison of recalls of students from first grade, fourth grade and university levels, on a set of stories which are very near the ideal structure. Hopefully, this is only the beginning stage of validation of the proposed model. An interesting result from this experiment is the probability of recall of a proposition as a function of its role. For university students, Settings (introduction of characters) and Beginnings (initiation of event sequences) were best remembered. Attempts and Outcomes (which together formed Goal Paths) were next best recalled. Reactions (mental events) and Endings (which were usually either very predictable or omitted in the stories used) were least well recalled. These results are in harmony with the results reported by Kintsch and van Dijk (which are discussed in Section 1.2.4).

2.5 Thorndyke

Thorndyke [77] attempts to validate the basic notion of underlying text form as a necessary part of a text understanding. He applies a story structure grammar, which is nearly identical to that proposed by Rumelhart, to two stories, "Circle Island" (analyzed by Frederiksen [75b]) and "The Old Farmer" (given by Rumelhart [74]), each of which results in a straightforward hierarchical structure. The initial rule of the grammar is

STORY -> SETTING + THEME + PLOT + RESOLUTION

Thorndyke wishes to test the effects of violating this rule. He does so by performing recall, summary and recognition experiments on four variations of

the stories. These are a normal form of the story, a story with the Theme moved to the end, a story with the Theme omitted and a descriptive version which omits causal and temporal continuity, using only stative or single action sentences. In comprehensibility judgements and recall tests, the normal form story was best, the Theme-after form next, with the last two cases more dependent on the particular story. Recalls of the Theme-after passages also showed a strong tendency to relocate the Theme to its normal position, near the beginning of the story. There was also a tendency in the more structured passages for higher-level propositions to have a higher probability of recall. Summarizations showed that, of the recalled propositions, the higher-level ones were much more likely to be included as part of the summary. Summaries of the descriptive presentations yielded a wider selection of propositions. In recognition tests, the more structured passages produced more recognition errors when the test proposition was consistent with the story, while the less structured passages produced more accurate recognition.

In attempts to separate the effects of content and structure, Thorndyke tested the four stories obtained by using each of the original plots with two character / object sets. He found that the Farmer plot was always more comprehensible. He also found that presentation of a second story using the same plot structure improved later recall of the first story.

Thorndyke concludes that text structures are basic to comprehension, although some are inherently more complex than others. The distinction between structure and content is supported by the positive effect of repeated structure prior to recall. And finally, the structural position of a

proposition has significant influence on both recall and summarization.

2.6 Text Grammar

"Text grammar" is the designation used for another currently active line of investigation into the structure of text. The fundamental ideas of text grammar are discussed by van Dijk [72], Ihwe [72], Kummer [72], Petofi [72] and Petofi and Rieser [73]. (Unfortunately, a great portion of the work is not available in English.) The principle concern is the formulation of generative descriptions of texts, rather than individual sentences, and is usually approached using a model similar to that of the generative semantics school of linguistics. Van Dijk suggests the following five components of a text grammar:

1. Semantic formation rules for the meanings of texts, as a whole.
2. Transformations which operate on this text meaning.
3. Transformations which produce a sequence of sentential semantic representations from the text meaning.
4. Transformations which produce a sequence of sentential syntactic representations (including lexical items) from the sequence of sentential semantic representations.
5. Rules pairing syntactic representations with morphonological representations.

These components would operate in the order given to generate a text. He emphasizes this is only the outline of a theory, and that most of the hard details remain to be specified. A significant amount of work has been done on components 1 and 2, formally defining representations of text. This has included numerous examinations of the use of some extended predicate calculus as a basic representation (primarily addressing the meaning of individual sentences and objects) but has also led to psychologically motivated investigations (as discussed in Section 1.2.4) and to studies in the general structure of textual types (as discussed in Section 2.2). Discussions of components 3 and 4 have been mostly descriptive, indicating phenomena that must be accounted for without actually defining them. Many of these phenomena seem best characterized as operations operating on two or more underlying propositions. These phenomena include anaphora (including pronominalization and article selection), sentence stress, contrast, use of clausal conjunctions, and verb tense determination. Of course, certain literary styles may alter the "normal" rules, choosing repetition over pronominalization omission of causal connection, repetition of certain sentence forms, etc. Operations which occur on a single proposition include: semantic transformations, such as personification and the use of metaphor; syntactic transformations, such as inversions and other stylistic operations; and phonological transformations, such as those producing rhyme, alliteration and meter. The set of realization rules used to produce a particular text from its generated underlying structure is assumed to be limited by the type of text that is being generated (e.g. mystery, narrative, etc.), which is indicated by some underlying abstract type marker. The selected rules will

also determine the textual characteristics usually referred to as style or literary merit.

It seems that any complete text understanding model must have a set of rules relating the underlying structure of the text - individual and groups of propositions together with global aspects of the text - to its surface realization. Although using a generative model, text grammarians are concerned with identifying and characterizing these relationships, and the results of their investigations should be of definite value in describing text understanding, once their research has reached the point of clearly defining the nature of these surface phenomena in terms of the meaning of text.

2.7 Discussion And Conclusions

The principal hypotheses of these investigators is that there is are very high level structures which organize content structures in the representation of a text. This section was begun with the warning that the distinction between these two different types of structures might be difficult to define. It seems clear that although each proposal that has been considered claims to be concerned with this high level structure, some confusion exists. The interactions between the abstract structures and the content structures are not characterized at all. Thus, Kintsch and van Dijk do not describe the restrictions on what type of events may realize a COMPLICATION category. Similarly, Rumelhart does not explain how the correct semantic rule would be selected when a choice is possible (e.g. Two sequential EVENTS may be related by either CAUSE or ALLOW). Mandler and Johnson, and to some extent Thorndyke,

create additional confusion by mixing syntactic categories with categories that would usually be considered semantic. Thus, conjunction time ordering and causality may appear in the "syntactic" structure of a story. Although this difficulty in distinguishing between the two types of structures may suggest that there is no distinction, the evidence provided by the existence of describable classes of texts seems to indicate that some type of high level structures do exist. It remains to clearly define the nature of these structures, and to explain their relationship with content structures, without repeating the fallacy of creating a multitude of subcategories (e.g. many different EVENT subtypes) which are alleged to be purely syntactic categories.

These structures fit neatly into the model discussed in the conclusion of the content section (section 1.4). They are the highest level of structure, under which content structures are organized. This is consistent with the hypothesis of Kintsch and van Dijk (discussed in section 1.2.4) that the structural description of a text is one of the important aspects of its representation. They specifically suggest that these structures organize macro-structures, which provide the content organization of the text. This is quite consistent with the results of structural analysis which, as Hendricks [73] points out, ordinarily uses a summary or synopsis, not the actual text, as a basis for analysis. (Recall that macro-structures are proposed as the basis for summary generation). Of course very simple or very novel texts might not have any structure of this type, since no appropriate structure would be pre-existent. Of course, the problem of learning these structure is as difficult to explain (if not more difficult) as the learning of content structures. Such acquisition is simply not well understood. The work of the

text grammarians suggests the complexity and diversity of linguistic phenomena that may be related to underlying textual representation. This work also suggests that many aspects of a text, other than its literal meaning and general type, may require explicit representation. Although a sound computational representation of stylistic aspects of texts may seem a distant goal, it is still important to retain it as part of the goal of understanding text, and to realize that the richness of human use of language will be only partially accounted for without this component.

3.0 COMPUTATIONAL MODELS OF TEXT UNDERSTANDING

This section will survey computational work on the process of text understanding. Computational models will be seen to be of value in at least two ways. First, many of the features postulated in other models will be found in these computational models, frequently motivated primarily by computational concerns. This tends to support these hypotheses. Secondly, many points passed over very quickly by other workers are seen to offer formidable problems when one attempts to set out a full computational description. The discovery of inadequate description is essential in developing sound models. "Computational" will be taken to mean any model that is actually programmable or that is formally defined. The work selected for discussion in this section is generally treated in more detail than that in previous sections. The reason for this is that if the complexity of computationally specifying certain representations and processes is to be made clear, it is impossible to treat these matters cursorily. This, in turn, has necessitated greater selectivity on the part of the author in choosing work that seems to best convey certain types of problems.

Bellert [70] attempts to define the notion of the coherence of a text, and in doing so suggests some of the same conclusions reached by computational linguists in their text understanding work. Coherence refers to the property of a set of utterances which make it a connected discourse rather than a random collection of utterances. Referential identity has often been suggested as an indication of coherence, but it is clearly insufficient. For example:

John drinks a lot of coffee. John married a blonde.

Both sentences refer to John, but do not form a coherent discourse. Furthermore, referential identity is unnecessary. Consider the two sentences:

There has been a drought. People are starving.

These are coherent without any explicit referential identity. Bellert defines a coherent text as one in which the semantic interpretation of each sentence is dependent on the semantic interpretation of the preceding sentences. The semantic interpretation of a sentence is defined as the set of conclusions that can be inferred from that sentence. She suggests that there are two types of conclusions that may be drawn:

1. those drawn only from knowledge of the language
2. those drawn from knowledge of the world

Both types of conclusions are absolutely necessary in understanding text and may be appropriately drawn when the coherence of the text requires. She concludes that "an utterance has meaning only in the entire context and through our knowledge of the world". The rest of the work discussed in this section will strongly reinforce these conclusions.

The discussion is divided into two sections. The first contains earlier work that supports the claim that world knowledge is essential to the understanding of text. The closely related problem of making implicitly conveyed information explicit is also a principle concern. The second section

describes later work in which the organization of world knowledge is recognized as a critical question for text understanding.

3.1 The Necessity Of World Knowledge

3.1.1 Charniak -

Charniak's work [72, 74, 76] is probably the first attempt to set out in a well defined fashion the dimension of information processing that must be carried on in the understanding of a story. The term "understanding" is necessarily vague, but Charniak suggests an intuitive definition. Consider the following story fragment:

Fred was going to the store. Today was Jack's birthday and Fred was going to get a present.

It should be clear that if a human had read and understood this fragment that he would be able to answer such questions as

1. Why is Fred going to the store?
2. Who is Fred buying the present for?
3. Why is Fred buying a present?

Charniak claims that a semantic representation of this fragment (i.e. an understanding of it) should explicitly contain the answers to ordinary questions such as these. Note the important point that this type of

understanding could be attained only through the use of general knowledge of the world, along with the explicit statements of the text. Knowledge required to answer the above questions would include such facts as:

1. A person having a birthday is likely to receive presents.
2. Presents are often bought at stores.

So Charniak's goal becomes outlining an answer to the question of how common sense knowledge may be incorporated into the process of understanding natural language. A closely related goal is the determination of how much knowledge of this type is required by the basic problems of natural language, such as the resolution of pronominal reference.

Charniak breaks the problem of processing natural language into two parts. The first part is the translation of natural language into a form that is convenient for use in making deductions. This internal representation is like the understanding that a person would be capable of obtaining without context. The internal representation is a canonical propositional form. Thus, either of the sentences

Jack caught a cold.

Jack came down with a cold.

might be represented by the proposition

(BECOME-SICK-WITH JACK COLD)

which represents the explicit meaning of either of the sentences. The second

part of the problem is what Charniak calls Deep Semantic Processing (DSP). This is the processing that makes explicit the implicit information conveyed by the story. Charniak elects to examine only DSP.

The function of the system that Charniak proposes is to read a story which has already been translated into internal representation and output a data base of propositions that explicitly represent all of the information conveyed by the story. In order to do this, the system uses common sense knowledge to make implicit information explicit. This knowledge is coded into the system initially, and is not learned or modified.

Charniak's goal of understanding stories leads directly to the questions of what kind of implicit information is conveyed, what kinds of common sense knowledge are required to explicate that information, how should this common sense knowledge be represented and when and how is it used. In answering the question about the types of implicit information Charniak discovered that resolution of pronominal reference frequently required common sense knowledge. Consider Charniak's most discussed example [Charniak, 74]:

Today was Jack's birthday. Penny and Janet went to the store. They were going to get presents. Janet decided to get a top. "Don't do that" Penny said. "Jack has a top. He will make you take it back."

The "it" in the last line is normally understood to refer to the top that Penny would buy, but any purely syntactic procedures would probably select the top that Jack already has, on the basis of recency. The correct choice of referents seems to be based on the common sense knowledge that if a present is purchased and given, it may be returned or exchanged. The fact that if a person receives a duplicate present he may not wish to keep it is also

relevant. This realization led Charniak to concentrate on the problem of pronominal reference. Instances of reference are clearly recognizable in the input, and the need to utilize common sense knowledge seems fully present.

Charniak decided to represent common sense knowledge as inferences. That is, implicit knowledge is made explicit by having rules which infer the implicit fact upon presentation of the necessary explicit information. For example, if one knows that it is raining, the inference that anyone outside will get wet is valid. The presentation of the sentence

John went outside.

is sufficient to trigger the common sense inference

John got wet.

in a situation in which it is known to be raining. Charniak's choice of inferences as a representation was probably strongly influenced by the availability of MICRO-PLANNER. (For a discussion of MICRO-PLANNER, see Winograd [74] and Charniak [76]) Before examining his inferences in more detail, one other question must be considered. Since an inference is drawn at some particular time, the question of when inferences should be made arises. Three possible answers to the question are suggested:

1. Make no inferences until the story is accessed (e.g. to answer a question about it, summarize a part of it, etc.)

2. Make inferences as the story is read, but only those that are necessary to solve some particular problem (e.g. to solve problems of ambiguity or reference)
3. Make non-problem inferences as the story is read, making explicit as much information as possible

Charniak rejects the first possibility because if it is necessary to make deductions from all previous propositions when accessing some proposition, there is no theoretical difference from making them as the story is read. And he argues that there is no way to be sure that you have made the correct and necessary inferences without examining all previous propositions. An argument against both the first and second possibilities is that the meaning of any proposition may be context-sensitive. Consider a story in which Janet wants to trade with Jack for his paints. The sentence

"Those paints make your airplane look funny" Janet said.

should probably be understood as part of a bargaining strategy and not an expression of Janet's true feelings. Since no explicit problem is presented by this sentence, no inferencing would be done, and hence the correct understanding would be missed. So Charniak selects the third possibility, making inferences whenever possible as the story is read.

Charniak implements inferences essentially as MICRO-PLANNER antecedent theorems. He distinguishes two types of common sense inferences which he designates base routines and demons. Base routines represent knowledge that

should always be available and does not need to be triggered by a specific context. For example, knowledge about trading should always be available so that any statement about a trade that has occurred will cause the inferences that the ownership relations have been reversed. Base routines are implemented as antecedent theorems matched against each input proposition. Some inferences are not always appropriate, such as the previous example about getting wet in the rain. The inference would be incorrect if made in a non-raining situation. This type of inference, called a demon, is implemented as antecedent theorems that are not always active. Demons are activated by base routines. Thus, there would be a base routine about raining which could make immediate inferences and activate demons appropriate to the context, rain. The previous inference rule would be a demon of this type.

Charniak refers to the set of propositions which match a base routine's pattern as its topic concept. He notes that a topic concept may occur either before or after a proposition which would match one of the topic concept's demons. Consider the two sequences:

1. It was raining. Jack was outside.
2. Jack was outside. It was raining.

In the first sequence, the rain base routine is matched which activates the outside-implies-wet demon. The demon matches the next proposition causing the inference that Jack is wet. Charniak calls this looking forward, and it is handled correctly by antecedent theorems. In the second sequence, however, the demon is not activated until after the proposition so the inference is

missed. This is called looking backward. Charniak proposes an extension to antecedent theorems so that when a demon is activated, the data base is searched for matches. Then, in the second sequence, the inference would be made when the the demon was activated.

A final observation about demons concerns deactivation. Obviously, if demons are context dependent, they must be deactivated when the context is no longer present. This is illustrated by

It was raining. When it quit, Jack went outside.

It should not be inferred that Jack got wet. Although this simple example could probably be handled by the rain base routine, in general the problem is quite difficult and Charniak offers no real solutions. He assumes that deactivating demons after some fixed number of intermediate propositions would be a satisfactory first approximation.

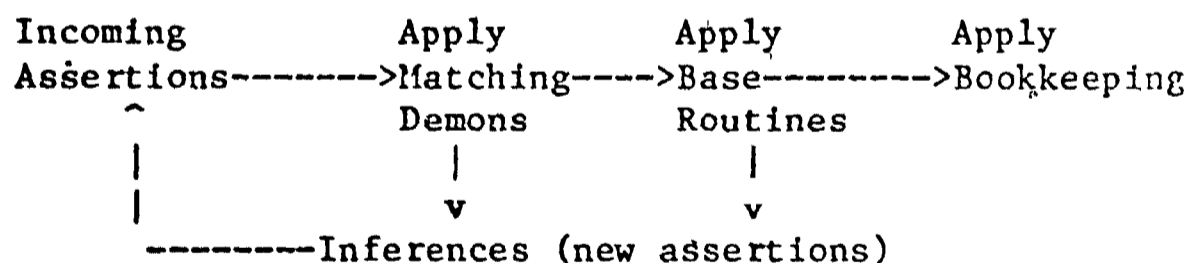
Charniak's model includes two other components. He suggests a bookkeeping component to keep the data base updated and consistent. When inferences are made, they will frequently replace other assertions previously true. For example if Jack is inside, then goes outside, bookkeeping would mark the original fact as no longer true, but would keep it for historical purposes (such as answering the question, "Was Jack inside?"). The fourth component is made up of fact finders. These are necessary as a result of Charniak's decision to make all possible inferences which are expressed as demons as the story is read. Clearly, many possible inferences should not be made to avoid clogging the system with a huge number of assertions. To avoid

this, these unnecessary inferences will not be realized as demons. Inferences such as the facts that John is in his house, his neighborhood, etc. if he is in his kitchen fall into this category. Some situations may require the availability of these facts. The sequence

Jack was in the house. Later he was in the kitchen.

should not be interpreted as a change in location, from the house to the kitchen. Fact finders are implemented as MICRO-PLANNER consequent theorems, with patterns which are matched against desired goals. They thus make certain information available through deduction that is not sufficiently important to infer as soon as possible. Fact finders are always available, not activated and deactivated like demons.

Charniak's complete model is depicted below. The model shows that the inferences are treated like additional propositions and are thus subject to DSP.



3.1.2 Rieger -

In some ways Rieger's work [74. 75] seems closely related to Charniak's, but it is not clear that this is completely true. Rieger's system accepts sentences as input which are already analyzed into their semantic representations, and makes explicit additional information that he claims is

implicitly in the input. He designates these additional pieces of information inferences, and agrees with Charniak that they should be made whenever possible. Rieger discusses the use of inferences in handling problems such as reference, but the bulk of his work is specification of the inferences, themselves, and it is here that the main value of his work is found. Rieger identifies sixteen classes of inferences which will be discussed in three broad categories (not Rieger's categories).

3.1.2.1 Causal Connection -

First, there are inferences which are concerned with the causal connections between states and acts (Rieger's treatment resembles the use in robotics work (e.g. Fikes and Nilsson [71]) of preconditions - states that must be true for an act to occur - and postconditions - states that result from an act occurring). Given that a state or act is true or has occurred, what inferences may be made?

1. Causative inferences suggest the likely cause
input: Mary has the diamond ring.
inference: Someone must have given or sold the ring to Mary.
2. Enablement inferences suggest states that were necessarily true
input: Mary gave John a book.
inference: Mary had the book just before she gave it to John.
3. Resultative inferences suggest results that followed
input: Mary gave John the book.
inference: John has the book.
4. Missing enablement inferences explain why something cannot occur
input: Mary couldn't see the horses finish.
inference: Something must be blocking her view.

5. Intervention inferences explain how something may be stopped or prevented

input: Mary was hitting John with a bat.
inference: Taking the bat away from Mary would stop her.

3.1.2.2 Missing Information -

The second category of inferences concerns supplying common knowledge about familiar objects or actions that is not in the input.

1. Specification inferences fill in missing parts
input: John hit Mary.
inference: John used his hand to hit Mary.
2. Function inferences supply the normal role of objects
input: John got a book.
inference: John will read the book.
3. Normative inferences supply information about what is normally true
input: Pete is a human.
inference: Pete probably has a gall bladder.
4. State duration inferences suggest how long some state will persist
input: John started eating at 6:00.
inference: He is probably still eating at 6:15.
5. Feature inferences connect features of objects with the objects
input: Fred wagged his tail.
inference: Fred is a non-human animal.
6. situation inferences supply other likely aspects of a situation
input: Mary is going to a masquerade.
inference: Mary is probably wearing a costume.

3.1.2.3 Motivation And Knowledge -

The third category of inferences concerns human motivations and knowledge, and their relation to one's actions.

1. Motivation inferences suggest reasons for an actor to do something
input: John hit Mary.
inference John wanted Mary to be hurt.
2. Action-prediction inferences suggest a possible course of action from a person's wants
input: John wants some nails.
inference: John is likely to go to a hardware store.
3. Enablement-prediction inferences suggest reasons for a person bringing about a certain state
input: Mary put on her glasses.
inference: Mary probably wants to look at something.
4. Utterance-intent inferences supply information intended, but not actually stated
input: Mary couldn't jump the fence.
inference: Mary wanted to jump the fence.
5. Knowledge propagation inferences predict what else a person would know if given that he knows certain things
input: Bill knew that Mary hit John with a bat.
inference: Bill knew that John had been hurt.

One point that should be clarified is that the use of the term inference to designate the implicit information a sentence conveys is partially misleading. What is implicit, and thus might need to be made explicit, is directly a function of the semantic representation. Consider the previous example

Mary gave John the book.

If the fact that John possesses the book after being given it is an inference, it is not clear what meaning the representation of the sentence captures. Surely, the transfer of possession is the meaning of the sentence, not just an implicit addition. At the other extreme, the example

John hit Mary.

surely does not necessarily mean that John wanted to hurt Mary. This is clearly conveyed implicitly, if at all.

It is very difficult to evaluate the completeness or adequacy of such a large set of inference types, but Rieger certainly suggests how large the class of information is that can be implicitly conveyed by text and, therefore which might need to be made explicit during the processing of that text. Any proposed scheme for representing and understanding text should be examined to see how each of these types of information is stored and in what way the information is asserted to be part of the meaning of the input.

3.1.3 Schank -

Charniak's model of story comprehension primarily discusses connections between propositions of a story that alter the sequence of events involved in understanding the story. For example, the understanding of some sentence of a story may be correct because a demon had been previously activated, thus allowing the correct interpretation of that sentence. Charniak does not say much about explicitly representing the connection between the proposition responsible for activating the demon, and the correctly understood

proposition. However, since he mainly investigates reference, obtaining the correct referent is an explicit representation of the understanding. Clearly, many kinds of connections might exist between propositions of a connected discourse. Schank ([73b], Schank and Abelson [77]) investigates causality as one of the primary discourse connectives. Schank's proposals are discussed in terms of his Conceptual Dependency theory [Schank, 73a], but are largely independent of it. The term conceptualization is used to refer to either a proposition (using one of the small set of primitive predicates or acts) or to a state (i.e. an object having some value for some attribute). Schank notes that the sentence

John cried because Mary said he was ugly.

asserts a connection between

Mary said John was ugly.

and

John cried.

But closer consideration reveals that some of the links in this connection are unspecified. John must have found out that Mary had said it and this knowledge must have made him unhappy, which was manifested in his crying. In order to be able to recognize instances of missing links and to be able to supply them, Schank develops a classification of types of causality and characterizes their form and meaning. (It should be clear that here, as in all Conceptual Dependency, there is no simple mapping from English to the

underlying representation.)

The first type of causation is a result in which an act may bring about a change of state. For example

John went to New York. -RESULT-> John is in New York.

The second type of causation is enablement. This is the situation in which a change of state brings about the conditions necessary for some act to occur.

John has a ring. -ENABLE-> John gave it to Mary.

The third causal type, initiation, is the relationship between any conceptualization and the act of someone thinking about that thing.

John -INITIATE-> I think about Bill.
(John reminds me of Bill.)

The fourth and final type is reason causation. This is the relationship that holds between the act of deciding to do something and actually doing that thing.

John decided to leave town. -REASON-> John left town.

Schank also suggests the utility of a non-specific causal connection, which he designates "lead-to". This would be used to represent causally connected events in situations where the specific type of causality or the specific chain of causes is unknown.

Although Schank explores the purely syntactic expansion of English causals into his formal causals, he concludes that the syntactic approach is too limited. Consider the sentence

The hurricane caused my depression.

Treating this as an instance of initiation causation (the hurricane initiated depressed thinking) makes it causal-syntactically correct, but there is clearly something missing that explains what led to the depression. Schank suggests that other knowledge is used to find the connection, and this suggestion seems to indicate a far more general approach to understanding English expressions of causality than any syntactic approach.

Schank suggests that a person would, if possible, find a reasonable connection such as that the hurricane probably blew down the speaker's house which caused him to lose money which caused him to become depressed. In order to accomplish this elaboration a person would have to use a great deal of world knowledge. Schank proposes that a number of different kinds of knowledge are involved here. First, there are axioms about the way people feel, such as

A1. Bad results can cause depression.

Associated with an axiom would be a number of more idiosyncratic beliefs, such as

IB1. Less money is a bad result.

Finally, general knowledge of the world would also be necessary. This would

necessarily include facts such as

- WK1. Objects can have monetary value.
- WK2. Changes in an object's physical state can cause a change in its worth.

A person making the above connections would essentially be engaged in problem solving. Given the final state, depression, he would find an axiom which explains it then find an idiosyncratic belief which would meet the axiom's preconditions (bad results), then use world knowledge to establish that a negative change of state could be the real culprit, then notice that a hurricane is able to destroy objects. The real point of this discussion seems to be that much additional knowledge is required to establish connections that people are able to (and do) make when understanding connected sentences. This is an affirmation of Bellert's hypothesis, and confirms Charniak's conclusions, but is reached after work on a different text understanding problem.

Schank [75b] attempts to use these types of causality as a basis for defining the semantic representation of a paragraph. He argues that the collected representations of the individual sentences of the paragraph do not form a representation of the paragraph as a whole. Much additional implicit information could be made explicit, but Schank suggests that undirected explication is not plausible. He offers a solution to the problem of what information should be made explicit by defining a paragraph representation as a causally connected sequence. Inferences are limited to those items of information required to find the causal connections.

The states that are usually true when an act occurs are called the necessary conditions of that act. These states would be connected to the act by an enabling causation. Schank divides necessary conditions into two types: absolutely necessary conditions and reasonable necessary conditions (ANCs and RNCs). The former must always be true when an act occurs. RNCs, however, are normally true, but may be violated without creating an anomalous situation. For example

John was working in his yard.

has as an ANC (among others) that John has a yard. An RNC might be that it was nice weather, but this could be violated. It would only indicate something unusual, and perhaps significant. Schank illustrates the use of necessary conditions to establish causal chains by analyzing several stories. A brief description of one of these should indicate his methods. He considers the following paragraph:

John began to mow his lawn. Suddenly his toe started bleeding. ... When he cleaned off his toe, he discovered that he had stepped in tomato sauce.

Mowing the lawn has as ANCs such things as the existence of John, the lawn, a mower, etc. and as RNCs that it is good weather and the grass needs mowing. These conditions would be inferred when the first sentence was processed since they were not explicitly stated. If the first sentence had been preceded by an explicit statement of either or both of these RNCs, they would have been connected as enabling the mowing. (There is great difficulty in finding a general scheme which allows inferring of normal necessary conditions, but prohibits inferring unusual ones. In this case, it seems reasonable to infer

all of them.) The second sentence presents a more difficult problem. Some sequence of inferences would have to be made which produces the chain: mowing involves turning blades, which could hit a toe, which could cause it to bleed. The final sentence of the story illustrates the point about normality considerations. To explain the tomato sauce, one must infer that John got it on his toe, but this seems like an abnormal inference to make with no other explanation. However, if the paragraph was preceded by the sentences:

John was eating pizza outside on his lawn. He noticed that the grass was very long and he got out his mower.

Then the final sentence seems much more reasonable. One would infer that pizza has tomato sauce, and John dropped some on the lawn and later stepped in it.

In summary, Schank states the following conditions for the representation of a paragraph:

1. Each input sentence is represented.
2. These representations should be conceptually connected, primarily by causal chains.
3. The necessary conditions for every conceptualization must be explicitly represented, and may originate as input sentences or may be inferred either from previous conceptualizations or because they are normal.

4. A story is the joining of causal chains that culminate in the "point" of the story. Other paths are of less interest.

3.2 The Organization Of World Knowledge

A characteristic common to all of the preceding work is that a great deal of world knowledge is required, but there is no clear organization of this knowledge. It is clear that any system with large amounts of knowledge represented simply as demons or inference rules would become bogged down searching for relevant knowledge and would quite probably draw incorrect inferences because knowledge would be applied in inappropriate contexts (recall Charniak's concern about the deactivation of demons). Charniak [75, 76, 77] himself suggests that more organized knowledge would be superior to the demon approach. These realizations have led to a number of proposals for the organization of knowledge. Minsky [75] introduced the term "frame" for knowledge structures, but Winograd [75] Bobrow and Norman [75] (using the term "schema") and Rumelhart and Ortony [77] discuss closely related ideas. Recently, Bobrow and Norman [77] have described a language for knowledge representation (KRL) in which the frame concept plays a central role. These knowledge structures are intended to organize conventional or encyclopedic knowledge, rather than definitional features or characteristics. This knowledge is described in terms of roles or slots which participate in the situation. These are the variables of the frame. Frames must also have the ability to reference other frames. A detailed discussion of the nature of

frames is inappropriate here, primarily because it is outside the scope of this survey, but in part because many of the computational questions regarding frames are not completely answered. However, recent computational work has begun the attempt to incorporate such organization of knowledge into text understanding systems.

3.2.1 Schank And Abelson -

3.2.1 1 Scripts -

Schank ([75b], Schank et al [75] and Cullingford [75]) extends his ideas with the addition of large, pre-existing knowledge structures that he calls scripts. He says that it is necessary to have large amounts of specific knowledge about known situations, since otherwise it is difficult or impossible to recognize the causal relationships between events. A script is a sequence of actions that provide knowledge about the typical occurrence of some situation. The following is a partial description of the script for going to a restaurant:

```
Script: Restaurant
Track: Coffee shop
Roles: Customer, Waitress, Chef, Cashier
Props: Tables, Menu, Food, Check, Money
Reason: To get food to eat
Entry conditions: Customer (is hungry; has money)
Results: Customer (is not hungry; has less money;
           is pleased)
           Cashier (has more money)
```

```
Scene 1: Entering
         Go to restaurant
         Find an empty table
```

Decide where to sit
Go to table
Sit down [MAINCON]

Scene 2: Ordering
Receive menu
Read menu
Decide what to eat
Tell Waitress what is wanted [MAINCON]

Scene 3: Eating
...

Scene 4: Exiting
...

The scenes are the main episodes of the event and each is defined in terms of a sequence of specific actions. Schank suggests that scripts would need to be divided at the top level into different tracks which distinguish the sequences for different types of restaurants.

The usefulness of a script is seen by considering the sequence

John went to a restaurant. He ordered a hamburger from the waitress. He paid and left.

Many details of this sequence have been omitted and connections between the actions would be impossible to establish without knowledge of what constitutes going to a restaurant. Furthermore, the definite reference to the waitress is meaningful only because all participants in a script are automatically introduced by any reference to that script. The script also contains the conditions that must be true for successful execution of the script, as well as the results of successful execution. Each scene has one act that is considered the essential act of that part of the script. These acts are

called the MAIN CONceptualizations of a scene.

Scripts represent the knowledge about conventional situations that any reader would be assumed to know. But the separate question of how to represent a text which contains script usages must be answered. If the representation of the restaurant story above was merely an elaborated causal chain containing many events inferred from the restaurant script, the unifying description that these events together constitute going to a restaurant would be lost. Schank proposes that that the elaborated causal chain is only one level of the story's representation - the Conceptual Dependency (CD) level. There is also a second level of representation - the Knowledge Structure (KS) level. At the KS level the above story would have a representation like:

```
Script=Restaurant
Customer=John
Food=Hamburger
```

Additionally, the script representation would be linked to each of the events at the CD level which was part of that script instance. Schank suggests that the causal chain at the CD level should contain all the events explicitly mentioned plus the MAINCONS of any scene that is mentioned. If any event is encountered which is not understandable in terms of the current script, the event will be represented at the CD level, but will also be placed on a Wierd list. For example Schank discusses the processing of a story in which John has his pocket picked while riding a subway. Later, he has no money with which to pay the bill after eating at a restaurant. The first sequence of events is understood in terms of the subway script, but the pocket picking event is placed on the Wierd list. When applying the restaurant script, the

inability to pay is encountered in the text and is inexplicable from the **restaurant script**. The script applier asks the monitoring program to look for **any unusual occurrences** that could result in John having no money. The **monitor** finds the pocket picking event, and returns the information to the **script applier** which uses it as the required explanation.

Of course, the problems in general are much more difficult than this **simple example** indicates. For example, an unusual event within a script could **interrupt** the script's normal continuation. Obstacles, such as not being able to **get what you want** at a restaurant, may cause altering or abandoning the **restaurant script**. Distractions, such as a robbery in a restaurant may lead to **suspension** of the restaurant script for a sub-story, or to its abandonment. Furthermore, simultaneous, independent scripts are possible. The situation of **eating in a dining car** involves both the train script and the restaurant script. Finally, the activation and termination of scripts is a very complex problem. For example, consider the sequence:

John went to a restaurant. After eating lobster, he bought a watch.

Does this describe a restaurant event followed by a purchasing event, or the **more unusual case** of buying a watch in a restaurant? Either is possible, and the story in which this sequence is found would determine the most probable **interpretation**.

3.2.1 2 Goals And Plans -

Schank and Abelson [75, 77] (also see Meehan [76] and Wilensky [76]) have recently suggested that all connected event sequences are not appropriately represented by scripts. Consider the sequence:

John wanted to become king. He went to get some arsenic.

It seems that treating the poisoning-the-king situation as so conventional that a script for it exists is unrealistic. The solution lies in realizing that while scripts handle well known situations, mechanisms must exist which are able to handle novel situations using general knowledge. "Plans" are suggested as the appropriate mechanism. A plan is a general course of action intended to realize some goal(s). A possible hierarchy of high-level goals is shown in Figure 3-1(a). Motivated actions are always associated with the satisfaction of some high-level goal. Frequently, more specific sub-goals exist which are motivated by a high-level goal. For example, someone may want to go to the train station (instrumental goal) so he can go to New York (a specified goal) so he can "enjoy pleasure" (high-level goal). The purpose of clearly defining a goal hierarchy is to enable understanding of situations in which an actor faces goal conflict and elects to pursue the highest goal (e.g. "preserve health" rather than "enjoy pleasure").

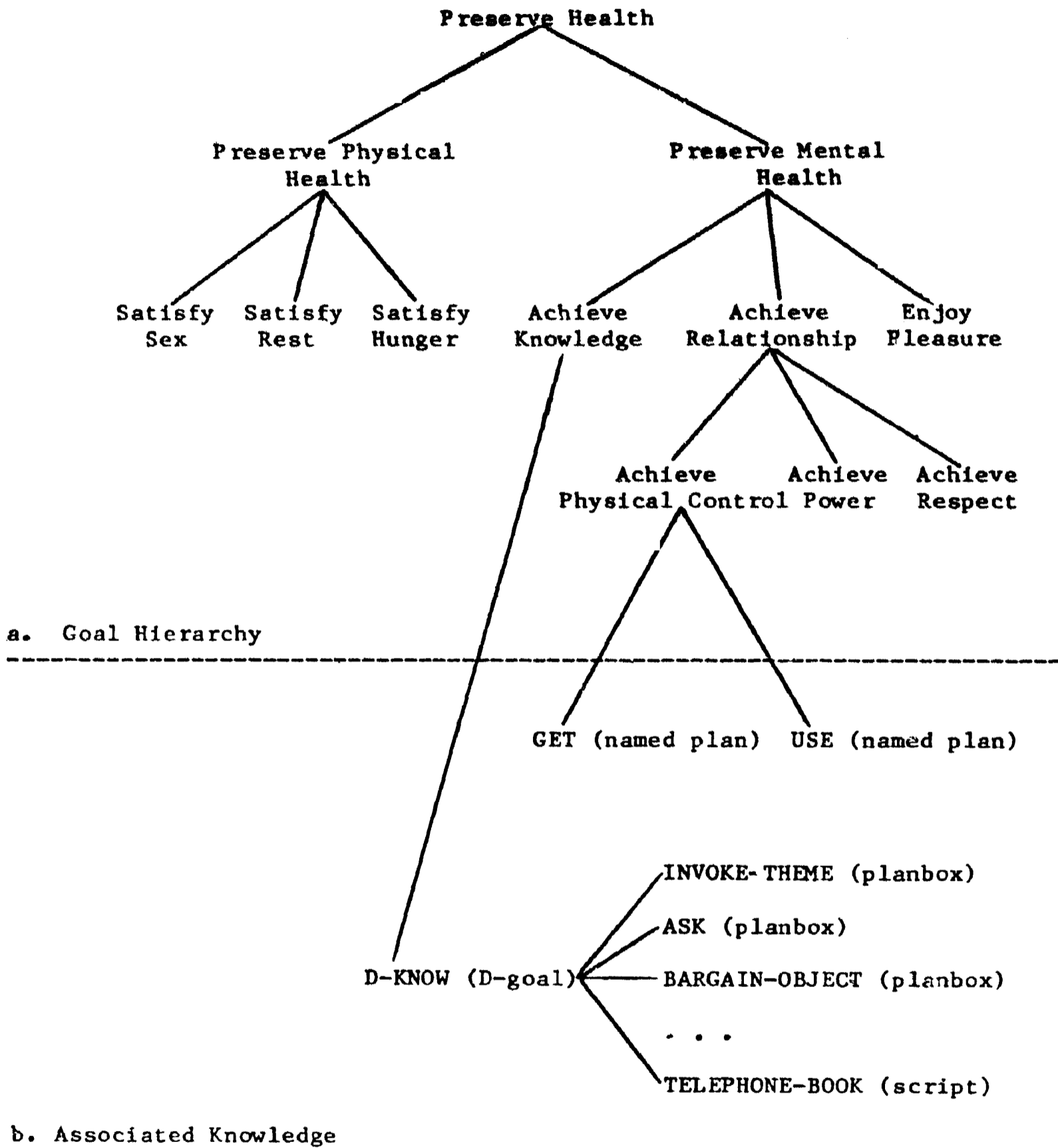


Figure 3-1

But general knowledge about how goals can be achieved must also be available. The "D-goal" is proposed as the fundamental unit of organization for such information. A D-goal is a point of access to planning information for the realization of some goal. For example, D-KNOW is the D-goal for the goal of knowing something, D-CONT for the goal of physically controlling something, and D-PROX for the goal of being in proximity to something. The D-goals' associated knowledge is an ordered list of "planboxes", each of which provides detail information on one method for achieving the goal. The ordering provides the sequence in which the methods are likely to be considered. For example, D-KNOW has an ASK planbox (as a highly likely method) which specifies the actual act of asking the appropriate question, the intended result (i.e. getting the answer), and the pre-conditions necessary to successfully ask. Some of the pre-conditions are: communication is possible; the person being asked knows the answer; the person being asked is disposed to answer the question. When a planbox is extremely specific, it becomes a script. Using the telephone book is a method of acquiring certain kinds of knowledge that is so conventional that it is a script. Certain recurring sequences of D-goals are called "named plans". The named plan, USE(x), stands for:

D-KNOW(the location of x)
 D-PROX(x)
 D-CONT(x)
 perform preparations and do action appropriate to x

Figure 3-1(b) shows the relationship of D-goals, planboxes, scripts and named plans to each other and to the goal hierarchy.

The representation of the plans of actors in a text would be at the KS level (with scripts). The D-goals would be explicitly represented and would be connected to actual acts at the CD level which (attempted to) implement the plan. For example, the sentence:

John tried to find out who ate the candy.

would be represented as:

<u>KS Level</u>	<u>CD Level</u>
D-KNOW(John,?.ate the candy)	plan ---implementation--> DO Failure <-----result-----

That is, the plan of trying to know who ate the candy was implemented as some unspecified act (DO) which failed to achieve the plan's goal. Acts (at the CD level) which meet pre-conditions of planboxes would be linked to the D-goal they enable.

Additional information is suggested as being appropriate and necessary for the KS level. This includes, for each character, his goals, the current status of each goal, strategies which could be used to achieve each goal and facts (true information as opposed to actual occurrences) relevant to goal understanding. This information is maintained on "Goal Fate Graphs", which also contain associations between characters and any "Themes" (large goal complexes such as BECOMING-RICH) in which he is participating.

Schank and Abelson appear to be committed to the development of a very complex system for the organization of knowledge, and the representation of such knowledge when it occurs in stories. It should be apparent that they are attempting to model goal related human knowledge and actions, since such knowledge and actions are common in stories. Since this knowledge may be required to understand particular stories it is not clear at what point (if any) one ceases to study text understanding and begins modelling personality.

3.2.2 Phillips -

Phillips' [75a,75b] presents probably the most comprehensive computational model of text, in that he is concerned with the representation of all types of knowledge and textual relationships. This breadth is informative, but necessarily results in a lack of depth in some areas. He presents a representational scheme which he uses for the various required types of knowledge, which include world knowledge, linguistic knowledge and the knowledge conveyed by the text. Phillips world knowledge, "the encyclopedia", consists of both the static data structures and the dynamic processes which operate upon these structures. The static data structure is a fairly conventional but very well defined, semantic network (closely related to the suggestions of Hays [73]) which use nodes for entities and events, and arcs for the relationships between nodes. The set of hierarchical relationships between entities (the taxonomic structure) is called the paradigmatic structure. The syntagmatic structure is the set of relationships between events and the event participants (case or argument relationships).

Phillips introduces a Modality node attached to every event which is always used to refer to the event and its participants as a whole. (In many systems this is represented only indirectly.) The discourse relationships of causality, time ordering and spatial relations are represented by arcs between the Modality nodes of events. Phillips defines another type of relationship which he designates the Metalingual organization of knowledge. This allows him to represent a concept as a unit, and yet have a complete subnetwork elaborating the meaning of that concept. For example, the unitary concept "poison" would have as its composition "someone ingesting something which causes that person to be ill". This representational technique provides the fundamental capabilities of frames or scripts - expansion of something into its parts, and knowledge about those parts.

The processes that operate on the semantic network are divided into two classes. The first, path-tracing, involves only following paths through the paradigmatic structure. This type of process would be used to find that "Mary gobbled caviar" was a more specific instance of "People eat food". The second type of process is pattern-matching, and involves constraints between the components. For example, determining that "John killed himself" is suicide while "John killed Bill" is not, requires a coreference constraint in the definition of suicide. Phillips observes that path-tracing is computationally equivalent to finite state automata, while pattern-matching is necessarily more complex.

Discourse, or text, is represented using the same representational scheme, but may be characterized by properties not applicable to discontinuous knowledge. The first of these properties is connectedness. Two propositions are paradigmatically connected if each has an argument such that the two arguments have a common paradigmatic superordinate node, or if the first is the immediate superordinate of the the second. For example, "lions" and "tigers" both have "animal" as superordinate, so the following are connected paradigmatically:

Lions are indigenous to Africa. Tigers have stripes.

And the two propositions:

Man is a hunting animal. Modern man hunts for sport.

are also connected since the second is a more specific proposition ("modern man" is immediately subordinate to "man") than the first. Two propositions are discursively connected when discursive relations (e.g. causality) exist between them.

The second attribute of discourse is thematicity. A theme is a prescribed pattern (represented by a Metalingual construction) to which a discourse may conform. A theme may be "contentive" like "accidental drowning" which is represented as existing when "a person is caused to be in the water and is unable to act, the combination of which causes him to drown". This is contentive since it specifies both the parts (events) of the theme and their interrelationships. A "non-contentive" theme is one that provides only a structural pattern or the interrelationships between some unspecified

entities "Clue" is defined to be "an unobserved act that causes the existence of something". The exact events and results are unspecified. A discourse is thematic if its propositions can be matched to a theme or hierarchy of themes (i.e. a theme which matches several propositions in a discourse, such as "accidental drowning", may in turn be a part of another theme, such as "tragedy"). Phillips then defines a coherent discourse as one which is both connected (each proposition is connected to at least one other proposition) and thematic.

Phillips points out that although there is no simple one-to-one mapping between his representations and the constructs used in text analyses like those of Grimes (section 1.2.3) and Propp (section 1..2), his structures do provide for the representation of most of the proposed relationships. Some of Grimes rhetorical predicates, such as Attribution, Specific and Collection, correspond to paradigmatic connectedness, while others, such as Covariance, correspond to discursive connectedness. Still other rhetorical predicates, such as Response and Analogy and Propp's patterns of functions in a move (as well as the functions, themselves) correspond to thematic structure. Hence, Phillips claims to have presented computational interpretations for the principle text phenomena.

As a test of his proposed model of text, Phillips presents a model of text understanding embodied in a computer program. His program inputs a discourse in the form of parse trees of the sentences and builds the knowledge representation of the discourse. An interesting use of the Metalingual construction is its function in replacing non-cognitive surface words by the

appropriate cognitive structure. For example, the preposition "through" is replaced by the cognitive structure for "in-contact-with" in a sentence like:

The Abominable Snowman walked through the snow.

Thus, the same representational mechanism is used for such divergent types of knowledge as syntactically-related information and thematic structures.

Once the parse tree has been converted into the knowledge representation of the proposition, the understanding process involves two major steps. The first is begun by matching the input proposition (IP) against the encyclopedia to find a corresponding generalized proposition (GP). Thus, the IP

The boat contains Horatio Smith.

is matched to the GP

Things contain people.

Notice that Phillips uses GPs to capture the same knowledge that features and selectional restrictions capture in many other systems. So the correct interpretation of a GP is that it is a plausible proposition, not a necessary truth. Once the IP has been matched to a GP, three additional checks are made starting from the matched GP. One is to determine if any of the terms of the proposition have Metalingual definitions. If so, new knowledge corresponding to the definition is added to the discourse representation. For example, if the discourse included that "John was poisoned" the elaborated information that "John ingested something that caused him to be ill" would be added. A second check is to see if the matched GP is discursively connected to any

other GPs. If so, new knowledge is added to the discourse which corresponds to the related GP and the discursive connection. If the matched GP were "People are injured" and it had a causal link to "People are unable to act", then the IP "John was injured" would result in the addition of "John was unable to act" with a causal link from the original IP. It is by using these interrelationships of GPs that Phillips accomplishes inference. The third test is to determine if the matched GP is a part of any contentive theme. (The encyclopedia's GPs have pointers to all contentive themes which contain them.) If so the theme is matched against the discourse as a whole to see if all components are present and all constraints satisfied. If they are, the theme is added to the discourse representation. "Accidental drowning" would be added to a discourse representation which had matched the GPs and connections " ("People contact water" and "People cannot act") cause ("People drown") " with a coreference constraint on "people".

After processing IPs and adding the related structure to the discourse representation, the second major step is testing the discourse for coherence. This involves two tests. The first determines if the discourse is connected. Then, thematicity is tested by checking to see if a single undominated theme has been found. It should be noted that non-contentive themes, since they have no component actions, and thus cannot be pointed to by GPs, must be tested for in a serial fashion. If the discourse passes both tests, then it is judged coherent.

To illustrate the process, a very brief description of the understanding of a story will be given. Given the story:

- (IP1) A boat contains Horatio Smith.
- (IP2) The boat overturns.
- (IP3) Horatio Smith drowns.

the following events would occur:

1. IP1 and IP2 would be matched to GPs which are parts of a complex event that has as a causal result, when instantiated, the added propositions:
 - (AP1) Horatio Smith is in the water. (caused jointly by IP1 and IP2)
 - (AP2) Horatio Smith is injured. (caused jointly by IP1 and IP2)
2. The GP corresponding to AP2 results in the addition of:
 - (AP3) Horatio Smith cannot act. (caused by AP2)
3. The GPs corresponding to AP1 and AP3 are part of a complex event that has the causal result "People drown". But this is the GP which matches IP3. So a causal link is added from AP1 and AP3 to IP3.
4. The discourse is tested for connectedness, and passes the test.
5. The discourse matches the "accidental drowning" theme, which is the only theme it matches, so it is thematic, and thus coherent.

Phillips principle goal was a computational model of text. His model of text understanding was intended to demonstrate that the proposed representations could actually be built from an input text. His text model does seem to present a reasonable representation for a number of text phenomena, several of which have not been considered by other computational models. Paradigmatic connectedness and some types of thematic structure seem particularly important. However, several objections must be raised to his model of text understanding. The first objection is that his tests for coherence are applied only to the complete discourse, and are not formulated in such a way as to suggest strategies to avoid incorrect interpretation of propositions and relationships. If this were done, the sequential processing of input propositions would continually test for coherence, and incoherence would immediately suggest that some misinterpretation might have been made, or some inference omitted. A second objection is that the definitions of connectedness and thematicity are inadequate. A common superordinate node is simply not sufficient to explain paradigmatic connectedness. The example of sentences about lions and tigers would be connected only in some context that explained why these statements were being made (e.g. "All I know about lions and tigers is ..."). Thematicity is defined without respect to how much of the discourse the theme accounts for. A theme cannot account for all of the propositions in a discourse unless the discourse is very trivial, and yet if a theme matched only the first three propositions of a one hundred proposition discourse, it could hardly be called the theme of the discourse. Finally, a number of aspects of the understanding processes are not convincingly shown to be computationally practical. The problem of avoiding incorrect additions to

the discourse representation from a rich encyclopedia is ignored. Each example has exactly the right information available. Also, it is not clear that the use of generalized propositions would work when there are several levels of generalization possible. For example, the generalized proposition "People contact water" has associated knowledge, but the higher level proposition "People contact things" would also need to be present with its associated knowledge. And finally, the methods of accessing themes - by pointers to all occurrences of generalized propositions or by serial search - both seem computationally unacceptable. Propositions like "People go" would result in a combinatoric explosion of possibilities. For all of these reasons, Phillips' understanding model is useful more in suggesting the kinds of problems involved, than in providing an actual model of text understanding.

3.3 Other Work

Space limitations and the narrowly defined scope of this survey have combined to eliminate certain interesting work from detailed consideration. Wilks [75, 76, 77] has described a text understanding (and translation) system which uses a meaning representation called Preference Semantics. The system normally operates in a basic mode, dealing with sentences individually, but is capable of entering an extended mode when a reference problem occurs. An example is "it" in

John drank the whiskey from the glass and it felt warm in his stomach.

Several processes are involved in the attempt to determine the correct referent. "Extraction" is the addition of logically true propositions, only

implicit in the text, but available from the meanings of the units. Thus, both

"It" was in John's stomach.
The whiskey is in a part of John.

would both be extracted. These two propositions can be identified, thus resolving "it" as "the whiskey".

However, had the above example been "... and it was good." the original propositions, as well as relevant extractions, would then be subjected to "common sense inference rules" such as

1. (1 drink 2) -> (1 judges 2)
2. (1 is good) <-> (2 wants 1)

which would be used to find the shortest inference chain identifying the referent. For the example, when "the whiskey" is tested, the established connections would be

(John drank whiskey) -1-> (John judges whiskey)
-subset-> (John wants whiskey) <-2- (Whiskey is good)

where "want" events are a subset of "judge" events.

A recent addition to Wilks' system is large knowledge structures (like frames) called "pseudo-texts". They are suggested as containing other knowledge which might be required in solving reference problems (although they have other important use in understanding individual words).

At the level of text, Wilks differs from Rieger and Charniak primarily in his insistence upon entering an extended mode only when a problem requires it. Reference problems are the only problems he discusses as triggering this mode. His use of pseudo-texts in establishing textual connections is still too briefly described to critically evaluate.

Rieger ([75b, 75c, 76a, 76b], Rieger and Grinberg [77]) has described a complex system for the organized representation of cause and effect knowledge, and plans utilizing this knowledge. A set of decision trees ("selection networks") are postulated which, given a goal, select certain "common sense algorithms" capable of realizing that goal. Rieger asserts that the same knowledge structures used for planning, should be used for understanding the intentional acts of others. To do this in text, he requires knowledge which allows prediction of goals and actions likely to be made in response to the occurrence of some event or state. A subsequent sentence is tested to see if it confirms any prediction by being a step in an expected action or in an algorithm to achieve an expected goal. This testing requires that all algorithms be indexed by step. Thus, it must be possible to find all occurrences of (X GOTO Y) in some set of algorithms. A confirmation is used to more confidently predict the course of action being followed. Clearly, Rieger is dealing with the same problems as Schank and Abelson (section 3.2.1.2), but has not yet clearly demonstrated the utility or practicality of his approach to this aspect of text understanding.

Hobbs [76, 77] has discussed analyses of various texts based on his own system of semantics. Possible intersentential relations are described by pattern - action pairs which, when matched by input sentences modify the text representation tree appropriately. These relations include causality, time ordering, paraphrase, example, contrast, parallel construction and violated expectation. These relationships indicate that Hobbs does not make a distinction between content and form (author imposed) relations. The complexity of Hobbs' system (in the large number of rules and their interactions) makes evaluation difficult until he has completed his computer implementation.

Other work that should be mentioned includes Schmidt ([76], Schmidt and Sridharan [77]), who discusses the problem of recognition of plans from actions, as well as Novak [76] and Bobrow, et al [77] who both use frame like knowledge structures in specialized language understanding systems.

3.4 Discussion And Conclusions

What have the discussed computational models added to the previously described model of text understanding? Charniak, Rieger and Wilks have demonstrated how it is possible to infer information only implicit in a text. Schank has especially examined the richness and complexity of causal connections, which are often implicit. This capability raises many additional questions, however, including when, and how many, inferences should be made. While Charniak and Rieger suggest making many inferences whenever possible, Wilks argues for making them only as required. Neither approach has yet been

applied to sufficiently large texts and knowledge bases to convincingly demonstrate its validity.

Pre-existent knowledge structures have been suggested by Schank and Philips to avoid the extremely difficult problem of making many-step inference chains to establish implicit connections. Although use of these structures has demonstrated their ability to meet this goal, as well as their usefulness in generation of summaries and paraphrases, problems have appeared. Complex knowledge structures and real texts present many difficulties in matching an input proposition with an element in the structure. When the match is imperfect, or when many choices are (computationally) possible, it is very difficult to perform the required matching correctly.

In an attempt to deal with this difficulty (as well as others), and to recognize the fact that novel situations are also understandable, Schank, Abelson, Rieger and Schmidt have been concerned with recognizing the motivations and intentions of actors. This aspect of understanding is clearly important, for the stated reasons, but the approaches described have almost certainly been too simplified for the understanding of actual texts.

Thus "the requirement that an understanding of a text "contains as much inferred information as necessary to meet some minimal level of ... coherence" (section 1.4) is seen to be a computationally difficult problem. The organizing content structures are useful in dealing with this problem, but create new problems. Knowledge and representation of plans and intentions is introduced to handle some of these problems. It should be noted that the previously discussed model has not specifically distinguished this type of

knowledge. Coherence has been defined in these computational models as connectivity established by inferencing, matching a knowledge structure or being a step in a plan. The exact distinctions between these three types of knowledge are still blurred, and it remains to be demonstrated that these distinctions are appropriate and adequate.

Only Phillips and Hobbs have addressed representing the form of texts, but these suggestions have not been computationally adequate, nor has there been any clear distinction between form and content structures. A better defined notion of text form should certainly play a role in the ongoing determination of textual coherence that occurs during comprehension in the text understanding model.

With regard to forgetting, little has been added, although higher-level knowledge structures could be used to summarize their more detailed components (which could then be "forgotten"). Schank has also suggested that the least connected propositions in a representation are those most likely to be forgotten. However, these ideas have not been seriously investigated. In general, computational models understand texts perfectly (if at all), and do not contain any imperfect retrieval processes. Permanent learning and integration with prior knowledge has not been investigated in these systems, nor has any explanation been offered for occasional recall of surface text. These systems generally do not keep the surface text at all, but could easily do so. However, it would result in perfect recall of this information.

The principal value of computational models has been the demonstration of the great difficulty in actually specifying comprehension algorithms. It is all too easy, when explaining how a particular result is achieved, to ignore the problem of avoiding incorrect paths. The fact that higher-level content structures are computationally useful, and also psychologically indicated, is an important confirmation. The investigations of more, actual texts by systems having more knowledge available (not just the relevant knowledge) is an important step in validating the models being proposed.

4.0 FINAL OBSERVATIONS

The following points are accepted by most researchers concerned with text understanding:

1. Much information implicit in texts is explicit in an understanding of that text.
2. There is some type of hierarchical or multi-level organization(s) of the understanding of a text.
3. Organized, pre-existent knowledge is required to achieve this type of understanding.

The study of knowledge representation, and the nature of knowledge structures is a primary concern. Many computational problems have not been resolved. A particularly important question for text understanding concerns the necessity of redundantly copying general knowledge for specific instantiations. Fahlman [77] discusses this problem in detail, and offers some ways in which to avoid unnecessary redundancy. The relationship of the content and the form of texts needs additional clarification. Similarly, the relationship of general content structures to representations for plans and intentions needs study to see how distinct these are. And of course, the large areas of learning and forgetting are important, but missing, components of a text understanding model. In addition to these open problems, attempts to apply the combined insights of the diverse research perspectives to actual texts is a necessary step in evaluating the adequacy of current text understanding models.

REFERENCES

- Anderson, R. C., Reynolds, R. E., Schallert, D. L. and Goetz, E. T. 1976. Frameworks for Comprehending Discourse. Technical report no. 12. Laboratory for Cognitive Studies in Education, University of Illinois at Urbana-Champaign.
- Anderson R. C., Spiro, R. J. and Montague, W. E. (eds.) 1977 (in press). Schooling and the Acquisition of Knowledge. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Barclay, J. R. 1973. The Role of Comprehension in Remembering Sentences. Cognitive Psychology, 4, pp. 229-254.
- Bartlett F. C. 1932. Remembering: A Study in Experimental and Social Psychology. London: Cambridge University Press.
- Becker, A. L. 1965. A Tagmemic Approach to Paragraph Analysis. College Composition and Communication, XVI, 5, pp. 237-242.
- Bellert, I. 1970. On a Condition of the Coherence of Text. Semiotica, II, 4, pp. 335-363.
- Bobrow, D. G. and Collins, A. (eds.) 1975. Representation and Understanding. New York: Academic Press.
- Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A., Thompson, H. and Winograd, T. 1977. GUS; A Frame-Driven Dialogue System. Artificial Intelligence, 8, 2, pp. 155-173.
- Bobrow, D. G. and Norman, D. A. 1975. Some Principles of Memory Schemata. in Bobrow and Collins [75].
- Bobrow, D. G. and Winograd, T. 1977. An Overview of KRL, A Knowledge Representation Language. Cognitive Science, 1, 1, pp. 3-46.
- Bransford, J. D., Barclay, J. R. and Franks, J. J. 1972. Sentence Memory: A Constructive versus Interpretive Approach. Cognitive Psychology, 3, pp. 193-207.
- Bransford, J. D. and Franks, J. J. 1972. The Abstraction of Linguistic Ideas: A Review. Cognition, 1, 2, pp. 211-249.
- Bransford, J. D. and McCarrell, N. S. 1974. A Sketch of a Cognitive Approach to Comprehension. in Weimer, W. B. and Palermo, D. S. (eds.). Cognition and the Symbolic Processes. Hillsdale. N. J.: Lawrence Erlbaum Associates.

- Chafe W. 1972. First Technical Report, Contrastive Linguistics Project. Department of Linguistics, University of California at Berkeley.
- Chafe, W. 1975. Some Thoughts on Schemata. in Schank and Nash-Webber [75].
- Charniak, E. 1972. Towards a Model of Children's Story Comprehension. AI TR-266, Artificial Intelligence Laboratory, Massachusetts Institute of Technology.
- Charniak, E. 1974. "He will make you take it back": A Study in the Pragmatics of Language. Technical report. Istituto per gli Studi Semantici e Cognitivi, Castagnola, Switzerland.
- Charniak, E. 1975. Organization and Inference in a Frame-like System of Common Sense Knowledge. in Schank and Nash-Webber [75].
- Charniak, E. 1976. Inference and Knowledge. in Charniak and Wilks [76].
- Charniak, E. 1977. Ms. Maloprop, A Language Comprehension Program. in Proceedings of the Fifth International Joint Conference on Artificial Intelligence.
- Charniak, E. and Wilks, Y. (eds.). 1976. Computational Semantics. Amsterdam: North Holland.
- Colby, B. N. 1973. A Partial Grammar for Eskimo Folktales. American Anthropologist, 73, pp. 645-662.
- Crothers, E. J. 1972. Memory Structure and the Recall of Discourse. in Freedle, R.O. and Carrol, J.B. (eds.). Language Comprehension and the Acquisition of Knowledge. New York: John Wiley and sons.
- Cullingford, R. 1975. An Approach to the Organization of Mundane World Knowledge: The Generation and Management of Scripts. in Proceedings of the 13th Annual Meeting of the Association for Computational Linguistics.
- van Dijk, T. A. 1972. Some Aspects of Text Grammars. The Hague: Mouton.
- van Dijk, T. A. 1974. Models of Macro-structures. unpublished mimeograph.
- van Dijk, T. A. 1975a. Recalling and Summarizing Complex Discourse. unpublished mimeograph.
- van Dijk, T. A. 1975b. Discourse Meaning and Memory. to appear in Journal of Reading Behavior.
- van Dijk, T. A. 1976. Macro-structures and Cognition. contributed to the Twelfth Annual Carnegie Symposium on Cognition.

- van Dijk, T. A. and Kintsch, W. 1977 (in press). Cognitive Psychology and Discourse: Recalling and Summarizing Stories. in Dressler, W.U. (ed.) Trends in Text-Linguistics. New York: de Gruyter.
- Fahlman, S. 1977. A System for Representing and Using Real-World Knowledge. unpublished PhD dissertation. Massachusetts Institute of Technology.
- Fikes, R. E. and Nilsson, N. J. 1971. STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving. Artificial Intelligence, 2, pp. 189-208.
- Franks, J. J. and Bransford, J. D. 1974. Meaning for Syntactic Form as a Function of Semantic Context. Journal of Experimental Psychology, 103, 5, pp. 1037-1039.
- Frederiksen, C. H. 1975a. Acquisition of Semantic Information from Discourse: Effects of Repeated Exposures. Journal of Verbal Learning and Verbal Behavior, 14, pp. 158-169.
- Frederiksen, C. H. 1975b. Effects of Context-induced Processing Operations on Semantic Information Acquired from Discourse. Cognitive Psychology, 7, pp. 139-166.
- Frederiksen, C. H. 1975c. Representing Logical and Semantic Structure of Knowledge Acquired from Discourse. Cognitive Psychology, 7, pp. 371-458.
- Grimes, J. 1975. The Thread of Discourse. The Hague: Mouton.
- Hays, D. G. 1973. Types of Processes on Cognitive Networks. in Proceedings of the 1973 International Conference on Computational Linguistics.
- Hendricks, W. O. 1972. A Structural Study of Narration: Sample Analyses. Poetics, 3, pp. 100-123.
- Hendricks, W. O. 1973. Methodology of Structural Narrative Analysis. Semiotica, VII, 2 pp. 161-184.
- Hobbs, J. 1976. A Computational Approach to Discourse Analysis. Research report 76-2, Department of Computer Science, City College, CUNY.
- Hobbs, J. 1977. Coherence and Interpretation in English Texts. in Proceedings of the Fifth International Joint Conference on Artificial Intelligence.
- Ihwe, J. 1972. On the Foundations of a General Theory of Narrative Structure. Poetics, 3, pp. 5-14.
- Kintsch, W. 1974. Representation of Meaning in Memory. Huntsdale, New Jersey: Lawrence Erlbaum.

- Kintsch, W. and van Dijk, T. A. 1975. Recalling and Summarizing Stories. unpublished English translation of "Comment on se' Rappelle et on Re'sume des Histoires". Langage, 40, pp. 98-116.
- Klein, S. 1974. Modelling Propp and Levi-Strauss in a Meta-Symbolic Simulation System. Technical report no. 226, Computer Science Department, University of Wisconsin.
- Kummer, W. 1972. Outlines of a Model for a Grammar of Discourse. Poetics, 3, pp. 29-55.
- Labov, W. and Waletzky, J. 1967. Narrative Analysis: Oral Versions of Personal Experiences. in Helm, J. (ed.). Essays on the Verbal and Visual Arts. Seattle: University of Washington Press.
- Mandler, J. M. and Johnson, N. S. 1977. Remembrance of Things Parsed: Story Structure and Recall. Cognitive Psychology, 9, 1, pp. 111-151.
- Meehan, J. 1976. The Metanovel: Writing Stories by Computer. unpublished PhD dissertation, Yale University.
- Meyer, B. J. 1975. The Organization of Prose and Its Effects on Recall. Amsterdam: North Holland.
- Minsky, M. 1975. A Framework for Representing Knowledge. in Winston, P. (ed.) The Psychology of Computer Vision. New York: McGraw-Hill.
- Norman, D. A. and Rumelhart, D. E. 1975. Explorations in Cognition. San Francisco: Freeman.
- Novak, G. 1976. Computer Understanding of Physics Problems Stated in Natural Language. American Journal of Computational Linguistics, Microfiche 53.
- Petofi, J. 1972. Syntactico-Semantic Organization of Text Structures. Poetics, 3, pp. 56-99.
- Petofi, J. and Rieser, H. (eds.) 1973. Studies in Text Grammar. Dordrecht: Reidel.
- Phillips, B. 1975a. Topic Analysis. unpublished PhD dissertation, State University of New York at Buffalo.
- Phillips, B. 1975b. Judging the Coherence of Discourse. in Proceedings of the 13th Annual Meeting of the Association for Computational Linguistics.
- Propp, V. 1968. Morphology of the Folktale. Austin: University of Texas Press.
- Rieger, C. J. 1974. Conceptual Memory. unpublished PhD dissertation, Stanford University.

- Rieger, C. J. 1975a. Conceptual Memory. in Schank [75a].
- Rieger, C. J. 1975b. Conceptual Overlays: A Mechanism for the Interpretation of Sentence Meaning in Context. in Advance Papers of the Fourth International Joint Conference on Artificial Intelligence.
- Rieger, C. J. 1975c. The Commonsense Algorithm as a Basis for Computer Models of Human Memory, Inference, Belief and Contextual Language Comprehension. in Schank and Nash-Webber [75].
- Rieger, C. J. 1976a. An Organization of Knowledge for Problem Solving and Language Comprehension. Artificial Intelligence 7, 2, pp. 89-127.
- Rieger, C. J. 1976b. Spontaneous Computation in Cognitive Models. Technical report no. 459, Department of Computer Science University of Maryland.
- Rieger, C. J. and Grinberg, M. 1977. The Declarative Representation and Procedural Simulation of Causality in Physical Mechanisms. in Proceedings of the Fifth International Joint Conference on Artificial Intelligence.
- Rumelhart, D. E. 1974. Notes on a Schema for Stories. unpublished manuscript.
- Rumelhart, D. E. 1975. Notes on a Schema for Stories. in Bobrow and Collins [75].
- Rumelhart, D. E. and Ortony, A. 1977. The Representation of Knowledge in Memory. in Anderson, Spiro and Montague [77].
- Scallert, D. L. 1976. Improving Memory for Prose: The Relationship Between Depth of Processing and Context. Journal of Verbal Learning and Verbal Behavior, 15, 6, pp. 621-632.
- Schank R. 1973a. Identification of Conceptualizations Underlying Natural Language. in Schank and Colby [73].
- Schank, R. 1973b. Causality and Reasoning. Technical report no. 1, Istituto per gli Studi Semantici e Cognitivi, Castagnola, Switzerland.
- Schank, R. (ed.). 1975a. Conceptual Information Processing. Amsterdam: North Holland.
- Schank, R. 1975b. The Structure of Episodes in Memory. in Bobrow and Collins [75].
- Schank, R. and Abelson, R. 1975. Scripts, Plans and Knowledge. in Advance Papers of the Fourth International Joint Conference on Artificial Intelligence.

- Schank, R. and Abelson, R. 1977. Scripts, Plans, Goals and Understanding. Hillsdale N.J.: Lawrence Erlbaum Associates.
- Schank, R. and Colby, K. M. (eds.) 1973. Computer Models of Thought and Language. San Francisco: Freeman.
- Schank, R. and Nash-Webber, B. (eds.) 1975. Theoretical Issues in Natural Language Processing. Association for Computational Linguistics.
- Schank, R. and the Yale A. I. Project. 1975. SAM: A Story Understander. Research report no. 43, Department of Computer Science, Yale University.
- Schmidt, C. F. 1976. Understanding Human Action: Recognizing the Plans and Motives of Other Persons. in Carrol, J. and Payne, J. (eds.). Cognition and Social Behavior. Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Schmidt C. F. and Sridharan, N. S. 1977. Plan Recognition Using a Hypothesize and Test Paradigm: An Example. in Proceedings of the Fifth International Joint Conference on Artificial Intelligence.
- Spiro, R. J. 1975. Inferential Reconstruction in Memory for Connected Discourse. Technical report no. 2, Laboratory for Cognitive Studies in Education, University of Illinois at Urbana-Champaign. (to appear in Anderson, Spiro and Montague [77]).
- Steinberg, D. D. and Jakobovits, L. A. (eds.). 1971. Semantics. Cambridge: University Press.
- Thorndyke, P. W. 1976. The Role of Inference in Discourse Comprehension. Journal of Verbal Learning and Verbal Behavior, 15, 4, pp. 437-446.
- Thorndyke, P. W. 1977. Cognitive Structures in Comprehension and Memory of Narrative Discourse. Cognitive Psychology, 9, 1, pp. 77-110.
- Weinold, G. 1972. On Deriving Models of Narrative Analysis from Models of Discourse Analysis. Poetics, 3, pp 15-28.
- Wilensky, R. 1976. Using Plans to Understand Natural Language. in Proceedings of the Annual ACM Conference.
- Wilks, Y. 1975. Semantics for Natural Language Inference. Artificial Intelligence, 6, 1, pp. 53-74.
- Wilks, Y. 1976. Parsing English. in Charniak and Wilks [76].
- Wilks, Y. 1977. Knowledge Structures and Language Boundaries. in Proceedings of the Fifth International Joint Conference on Artificial Intelligence.
- Winograd, T. 1974. Five Lectures on Artificial Intelligence. AI Memo no. 246, Stanford University.
- Winograd, T. 1975. Frame Representations and the Declarative / Procedural Controversy. in Bobrow and Collins [75].