Towards a Computational Theory of Textual Discourse

Rob't F. Simmons

Technical Report NL37

Department of Computer Sciences

University of Texas at Austin

September 1978

ABSTRACT

An outline for a computational theory of textual discourse is presented and illustrated by the analysis of an example paragraph. A three year program of research is proposed, aimed toward the further development of the theory, and the testing of discourse grammars for a wide range of textual discourse including narrative, essay, editorial, technical and instructional, and dialogue forms. Existing programs will be used to test and debug the discourse grammars and these programs will be further developed to result in a unified discourse modelling system.

Towards a Computational Theory of Textual Discourse

I Introduction

A computational theory of textual discourse, in our view, would define a process for generating and analyzing a sequence of text sentences that comprise a well-formed discourse. Such a theory would be required to account for:

1) the selection and ordering of a sequence of propositions that represents some human's experiential data,
2) the transformation of a string of propositions into a corresponding sequence of natural language sentence derivations,
3) the realization of each sentence derivation as a morphological string in some particular natural language.

A computational interpreter for such a theory would be required to analyze and generate textual discourses using such data structures as networks of propositions and rules, with various inference algorithms for parsing, establishing connectivity, proving the truth of assertions, answering questions, etc.

A theory of discourse is necessarily interdisciplinary in its nature. Language philosophers and logicians have contributed very important notions of speech acts, conversational postulates, and many analyses of word meanings and rhetorical patterns of argument (Austin 1962, Searle 1969). Linguists have developed initial theories for assigning systems of rhetorical labels and grammar forms for producing well-formed sequences of propositions. (Grimes 1975, Halliday & Hasan 1976, Beaugrande 1978). Psychologists have studied what human readers remember of discourse and have developed methods for measuring the relative importance of particular propositions to what is recalled. (Kintsch 1974, van Dijk 1975, Meyer 1975, Thorndike 1977). Computer scientists have presented prototype systems for modelling the understanding of dialogue (Winograd 1972, Groszh 1977, Cohen 1978), and of narratives, (Meehan 1977, Cullingford 1978, Simmons & Correira 1978). In fact, modelling of discourse and dialogue understanding was a major topic of the recent conference on Theoretical Issues in Natural Language Processing. (TINLAP-2 1978)*.

The problems in developing a fairly complete computational theory of discourse are complex. At the base of the theory, there must be a formal system which can represent rules of discourse and propositions that encode understanding of experiential events, and which can apply these data to the generation or analysis of well-formed discourses. The scope of the theory must be wide, eventually encompassing:

Narrative-- stories, newspaper articles, folk-tales, etc.

Encyclopedia Descriptions -- technology, biography, natural history,

Technical and Instructional Manuals,

Argument -- editorials, essays, advertisements,

Dialogue,

Poetics.

The theory must also have the depth to account for morphological, syntactic and semantic analysis of utterances in the context of particular discourses. It must be internally consistent and consistent if possible with psychological studies that suggest the structures that humans use to organize or understand texts and dialogue.

* The articles and references in this proceedings encompass the literature relevant for natural language understanding.

The value of such a theory is manyfold. When realized as a computer program, it can serve to test the consistency of rule systems for defining various forms of discourse. It will generate psychological hypotheses about how people may understand various forms of language. Eventually the computer realization of such a system will be able to analyze, summarize, translate and generate ordinary textual materials.

II Prototype Theory of Discourse

We define the primitive elements for a theory of discourse as Experience, Symbols, Thoughts and Propositions. Experience is an intangible, uniquely organic phenomenon that can only be pointed to as the organism's participation in one event after another. Strictly speaking, experience is known by an individual human and inferred of other organisms. We define Symbols as uniquely public experiences; people can agree on the presence or absence of a particular symbol or mark. A Proposition is an ordered set of symbols. A Thought (or concept) is defined as the association of a proposition and an experience. Experience can occur without thought (e.g. deep sleep, etc.) A Discourse occurs when some human organizes a series of thoughts and utters the propositions associated with them. Although the utterance may be accomplished by natural language, gesture, painting, sculpture, etc., our concern with discourse is limited to written texts.

A discourse can occur without thought if some mechanical system is used to order symbols and transform the result into natural language. Such a discourse can be meaningful -- i.e. a human experience -- but only if some human thinks about it. Computers can apply transformations to preserve or change the meaning of propositions in someone's experience, but the computer itself is a communication device that can transmit propositions (with or without change) from and to humans and other machines.

Although the purpose of discourse is to communicate human experience, the discourse itself is a physical linguistic object composed only of symbols and propositions. We might be tempted to characterize a well-formed discourse as one that can communicate a set of propositions so that the experience of the listener corresponds in a predictable way with that of the sender; but this correspondence of experiences can only be inferred by measuring the sender and receiver's agreement on propositions included in or implied by the discourse.

But the same tests can be applied to a computer without any notion of experience. If a computer analyzes a given text to obtain a sequence of propositions, we can measure the agreement between this sequence and those that humans -- or even other computers -- assert to be true of the text. The point of this discussion is to define a notion of objective discourse as a logical system which may be associated with "meaning" and "experience" in humans, but in itself is composed simply of propositions.

A well-formed discourse is defined as a sequence of propositions that are terminal nodes in the derivation tree of some discourse grammar.

A grammar is defined broadly as a system of propositions composed of rules and terminals. A rule is composed of a pattern and a series of operations so that when the pattern is matched, the operations are undertaken. An operation is either a rule or a terminal. A terminal has no operation. This definition of a grammar is very general and it covers phrase structure forms, production systems, and the highly abstract systems of Horn clauses that may include functions and quantified variables.

A discourse grammar is defined as a system of rules that terminate in propositions.

A derivation tree can be generated from any rule in a grammar by setting that rule as the root and realizing its operations as derivation trees recursively until terminal elements are achieved. The sequence of terminals in such a derivation is a well-formed string with respect to the grammar. If we truncate the derivation by removing its terminals, the foreshortened tree is a summary derivation.

A grammar has the summary property, if for every summary derivation, the leaves form a summary of the truncated terminals. The summary property is achieved by associating with each rule in the grammar, a transformation that summarizes its descendent nodes.

A rhetorical grammar is defined as a system of rules that transform a sequence of propositions into a (usually shorter) sequence of sentence derivations.

A sentence is defined as a sequence of words or morphemes that are terminal elements in a sentence derivation tree generated by some sentence grammar. By including appropriate transformations at each node, (as is done in deep case structure analysis) the sentence grammar has the summary property.

An Interpreter is a procedure that starting with a rule or a set of terminals, develops from a grammar one or more derivation trees as constrained by the rule or set of terminals.

This outline of a discourse theory claims that a person has experiences, some of which may be associated with propositions to form thoughts. When thought is to be communicated, its propositions are ordered by selecting a derivation from a discourse grammar. That is, a discourse grammar is used to organize a set of propositions into a sequence dominated by a discourse derivation. Elements of this series of propositions are combined, deleted and reordered according to a rhetorical grammar to form derivation trees for sentences. We visualize these derivation trees as the semantic structures of sentences, e.g. deep case structures. At this point a sentence grammar is applied to the semantic structure to produce a sequence of morphemes or words in some natural language.

Understanding a discourse requires an interpreter to analyze strings of words, using a sentence grammar to produce a semantic derivation that represents the propositional structure of the sentence. The rhetorical grammar is then applied to a sequence of sentence derivations to transform them into a sequence of discourse propositions. This operation is required to establish reference for each anaphor (i.e. ellipsis, pronoun or definite noun phrase) and to restore propositions that are implied by the context. The resulting connected series of completely instantiated propositions is interpreted with the aid of a discourse grammar to provide a discourse derivation tree.

Each grammar is so organized that its derivation trees have the summarizing property so that truncation of the terminal nodes of the tree results in a summary derivation whose terminals generate a summary of the discourse or sentence. This summarizing property is in accordance with psychological evidence of what humans understand and remember from texts, and it provides a most desirable artifact in that for any text generated or analyzed, summaries are immediately available.

So far the computational theory calls for three levels of derivation:

1) Discourse level -- Applies ordering rules via a discourse grammar to produce a derivation tree whose terminals are a discourse ordered sequence of propositions.

2) Rhetorical level -- Applies rules to delete, combine and reorder propositions into semantic descriptions of sentences.

3) Sentence level -- Applies case grammar rules to produce a derivation tree whose terminals are words in some natural language.

We have developed computational systems for each of these three levels and we are satisfied with the interpreter and rule forms for levels one and three. Our experience with level two is limited so far to a single study that delineates some of the parameters of rhetorical transformation and we must still define the grammar form and the interpreter to accomplish level two.

At the discourse level we have so far only studied narrative grammars and a wider range of discourse may cause revisions to the theory. Further, we have studied the three levels in isolation but there are important interactions that can only be appreciated when all three levels are integrated. Clearly when a human writes discourse the choice of words, the structure of sentences and the sequencing of propositions are all interdependent, and an adequate theory of discourse should account for the inter-relatedness of these choices.

The methodology for developing this theory reduces to the following steps:

1) Select texts

2) Analyze them into propositional structure

3) Abstract the propositions into such notions as Topic, Class, Attribute, Act, Result, Setting, Motivation, Goal, Complication Resolution, Conclusion, etc.

4) Observe the sequences of these abstractions in various forms of discourse,

5) Characterize the discourse by rules that define sequences of these abstractions,

6) Test the rules for their capability to generate, analyze and summarize the discourse,

7) Test for psychological validity by determining if human readers agree that the propositions that the grammar infers in establishing coherence of discourse are true,

8) If steps 4 and 5 fail to produce understanding of the discourse, resort to psychological experiments to determine what people remember of the text and what summaries they produce. This procedure clarifies the linguistic analysis.

There is also a computational approach to improving the rule forms and interpreters. This is the fairly common method of measuring where the computing time is being consumed and improving the programs. But the real gains in improved computing systems come into being, when one realizes that there is an underlying similarity of computations in several programs, and a reformulation can find a simplifying expression for the several.

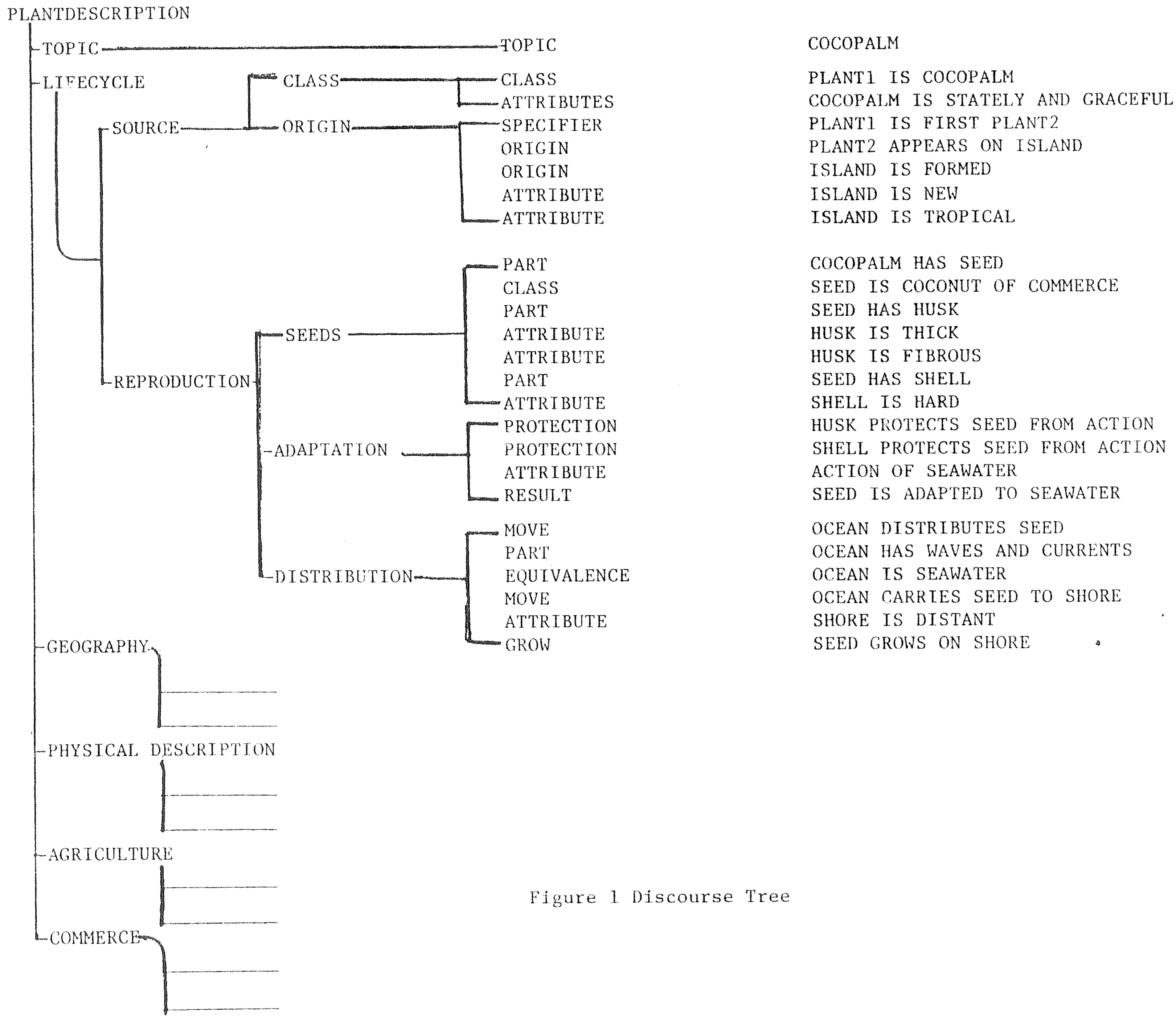

Figure 1 Discourse Tree

III Examples of Discourse Analysis

The attached paper, "Rule Forms for Verse, Sentences and Story Trees", includes several example analyses of narratives and narrative grammars, and detailed discussion of the nature of story trees (i.e. discourse trees for stories). Of particular interest is the relation of the derivation tree for solving a block stacking problem, to a corresponding discourse tree. An interpretive program is also described for generating stories and summaries from a discourse grammar.

In this section we present a discourse analysis for a paragraph from an expository text about coconut palms from an encyclopedia article. (See Compton's 1962) The article begins as follows:

> Coconut Palm
>
> One of the first plants to appear on a newly formed tropical island is the stately and graceful coconut palm. The seed which is the coconut of commerce, is protected by its thick fibrous husk and hard shell from the action of seawater. It is thus peculiarly adapted for distribution by ocean waves and currents and may be carried over a thousand miles from the parent plant to grow on some distant shore.

Figure 1 shows a discourse tree for the propositions of this paragraph in the context of following (unanalyzed) paragraphs abouts its geography, physical description, agriculture and commercial uses. (Certain propositions, namely, "peculiarly adapted" and "over a thousand miles from the parent plant" are deleted for abbreviation of the presentation.)

In Figure 1, the propositions are listed on the right, and the classifying name for each proposition is to its immediate left. Groups of these class names are further classified using names that relate to the topic of the propositions that are grouped. The classifying process is continued until everything is gathered under the node, PLANTDESCRIPTION. It is this process of naming and classifying that results in the summarizing property associated with discourse trees. Reading from the left, we can see that a plant description is outlined and thus summarized, as a TOPIC, LIFECYCLE, GEOGRAPHY, PHYSICAL DESCRIPTION, AGRICULTURE, and COMMERCE. The LIFECYCLE is further subdivided into SOURCE and REPRODUCTION, each of which is further divided until finally the propositions on the right are achieved. The analysis was derived following the method described in Section II (p. 8). Since there are many ways of classifying propositions, we cannot be sure that the example is a good discourse analysis; only the requirement that it form the basis for summarizing the discourse at various levels assures us that it is at least minimally adequate.

Figure 2 shows a propositional tree for the discourse. In this figure, the classifying names from Figure 1 are taken as predicate names, and variables are used to communicate shared arguments from proposition to proposition. Variables are either single alphabetic characters from M to Z or such characters associated with a number e.g. M1, M2, ... Z1. The variables eventually correspond to words or phrases in the discourse so the X of (TOPIC X) corresponds to COCOPALM in the righthand proposition, (TOPIC COCOPALM). Communication of variables from proposition to proposition can be seen clearly in Figure 2b under the node REPRODUCTION. The propositions, (SEEDS X Y U Z W) and (ADAPTATION Y Z W S1) share the variables Y, Z and W. Under SEEDS, these variables are bound as follows: Y: SEED, Z:HUSK and W:SHELL. The values of these variables are passed to the ADAPTATION proposition to give, (ADAPTATION SEEDS HUSK SHELL S1), and S1 is later bound to SEAWATER to be passed on to the proposition (DISTRIBUTION SEED SEAWATER T M). This binding and passing of variables in propositions models a major aspect of coherence in discourse.

```
(PLANTDESCRIPTION)
        (TOPIC X)                      (TOPIC COCOPALM)
        (LIFECYCLE X)
             (SOURCE X)
             (REPRODUCTION X)
        (GEOGRAPHY X)
             (...
        (PHYSICALDESCRIPTION X)
             (...
        (AGRICULTURE X)
             (...
        (COMMERCE X)
```

FIGURE 2a Top Level of Discourse

```
(LIFECYCLE X)
   (SOURCE X)
        (CLASS1 X Y)
             (CLASS Y X)               (CLASS PLANT1 COCOPALM)
             (ATTR V X)                (ATTR (STATELY GRACEFUL) COCOPALM)
        (ORIGIN1 Y)
             (SPECIFY Y Z)             (SPECIFY PLANT1 (FIRST PLANT2))
             (ORIGIN R1 Z W)           (ORIGIN APPEAR PLANT2 ISLAND)
             (ORIGIN R2 W)             (ORIGIN FORMED ISLAND)
             (ATTR N W)                (ATTR NEW ISLAND)
             (ATTR S W)                (ATTR TROPICAL ISLAND)
   (REPRODUCTION X)
        (SEEDS X Y Z W)
             (PART X Y)                (PART COCOPALM SEED)
             (CLASS Y U)               (CLASS SEED (COCONUT OF COMMERCE)
             (PART Y Z)                (PART SEED HUSK)
             (ATTR Z Z1)               (ATTR HUSK (THICK FIBROUS))
             (PART Y W)                (PART SEED SHELL)
             (ATTR W W1)               (ATTR SHELL HARD)
        (ADAPTATION Y Z W S1)
             (PROTECTION Z Y S)        (PROTECTION HUSK SEED ACTION)
             (PROTECTION W Y S)        (PROTECTION SHELL SEED ACTION)
             (ATTR S R1 S1)            (ATTR ACTION OF SEAWATER)
             (RESULT Y R4 S1)          (RESULT SEED (IS ADAPTED TO)
                                                           SEAWATER)

        (DISTRIBUTION Y S1)
             (MOVE T R1 Y1)            (MOVE OCEAN DISTRIBUTES SEED)
             (PART T V)                (PART OCEAN (WAVES CURRENTS))
             (EQUIV T S1)              (EQUIV OCEAN SEAWATER)
             (MOVE T R2 Y M)           (MOVE OCEAN CARRY SEED SHORE)
             (ATTR M M1)               (ATTR SHORE DISTANT)
             (GROW Y R3 M)             (GROW SEED GROW SHORE)
```

Figure 2B Detail of Lifecycle Node

Figure 2 Example Tree of Discourse Propositions

The illustrated tree of discourse propositions is at once the abstraction of a derivation tree for the discourse, and a propositional grammar that can be used to generate or analyze the set of propositions. The method for computing with such structures is shown in the attached paper. (Analysis is described in Correira 1978). The discourse tree is translated into Horn clauses as in the following example:

```
(LIFECYCLE X) < (SOURCE X) (REPRODUCTION X)
(SOURCE X < (CLASS1 X Y) (ORGIN1 Y)
(CLASS1 X Y) <(CLASS Y X) (ATTR V X)
```

In this form, the interpreter can use the clauses as rules to generate or analyze a discourse to instantiate a discourse tree with summarizing properties.

When the interpreter, starting with (PLANTDESCRIPTION), has instantiated the entire tree of Figure 2, all variables are bound and the tree has as leaves the terminal propositions on the right hand column. The resulting tree has the summary property. For example, the briefest summary is, PLANTDESCRIPTION of COCOPALM. Another summary of the entire article would be, PLANTDESCRIPTION of COCOPALM, its LIFECYCLE, GEOGRAPHY, PHYSICAL DESCRIPTION, AGRICULTURE and COMMERCE. It can be noticed that these summaries actually require transformation of the propositions to produce English expressions; this is accomplished by associating a transformation with each node. The method is described in the attached paper.

We can also see some similarities of this notation to notions of frames (Minsky 1975), to scripts and plans (Schank and Abelson 1977) and to templates and paraplates (Wilks 1976). The top level tree of Figure 2a resembles a frame named PLANTDESCRIPTION with slots for

TOPIC, SOURCE, GEOGRAPHY etc. In Figure 2b, the node ADAPTATION is more like a script that says if a seed has parts that protect it from some medium, (i.e. seawater), the seed is adapted to the medium. And Wilks' templates are propositions that can easily be written in Horn clauses.

It is currently our impression that discourse grammars derived from analyses such as that of Figure 1 will prove adequate to express the larger units of computation that are encoded in frames, scripts and plans, or templates and paraplates. The use of propositions with variables at once provides for maintaining coherence and the establishment of topics and summaries for each level of the discourse tree.

IV Natural Language Software*

In the past four years we have developed several programs that have allowed us to study sentences and discourse empirically. Generally, our method has been to develop interpreters and one form or another of rule system (generally referred to as grammars) to accomplish a logical, linguistic or psychological task.

One of the earliest of these was Hendrix' mathematical exploration of the nature of semantic networks (Hendrix 1975) which showed very clearly that the network formalism was a convenient computational representation for a set theoretic modelling scheme and was theoretically sufficient to model most describable events including inference rules. The interpreter for semantic networks was a question answering algorithm that evolved through five versions to result finally in a system that could accomplish fully quantified inferences

*Programs in this section prepared cooperatively by Dan Chester, Alfred Correira, Robert Simmons and M. Kavanaugh Smith in semantic network representations of knowledge. (Simmons & Chester 1977).

Contemporary with this effort we were also exploring the translation from short English descriptions into line drawings. (Simmons & Novak 1975, Smith 1975, Novak 1976). The immediate result of this work was to provide for us a new perspective on the problem of analyzing sentences into deep case structures. This perspective led us to a general purpose, all-paths, chart-based, bottom-up parser which has since become an important teaching tool and a component of many of our language study systems. (Simmons 1978). It has accumulated about 350 rules and has been tested on about 50 sentences.

At that time we also programmed three experimental systems to generate English sentences from deep case structure analysis. One of these was brought to a high degree of development and is used as a teaching system and included in other language processing systems. It has been used successfully with small grammars of up to fifty or so rules, to generate dozens of well-formed sentences.

Early this spring, this system was augmented with the capability to interpret context sensitive lexical transfer rules. This augmentation gave it the capability to paraphrase and to translate between languages. The parser and the paraphraser were then integrated into a system that translates sentences, and the translator has been tested with a paragraph of Danish translated to French, a paragraph of Thai to English, and a paragraph of German to English. Within the limits of single sentence analysis the system is thoroughly satisfactory (but obviously incomplete with respect to translating discourse).

Late in 1976 we were inspired by a van Dijk study of what people understood of text (van Dijk 1975) to construct a discourse understanding system that would reproduce his findings. By June of 1977 we had developed a prototype system for generating stories and their summaries and since then have further developed the system to generate and to analyze sequences of propositions that form narrative discourse. (Simmons & Correira 1978, Correira 1978).

The primary discourse analysis system accepts a grammatical ordering of Horn clauses and either generates or analyzes a string of case relations dominated by a derivation tree that has the summary property and provides for the system to produce summaries of the stories it has generated or analyzed. This sytem has generated and analyzed several short tales and one comparatively long Bocaccio story of 150 propositions.

Since this system generates a series of propositions we were motivated to initiate a study of how to transform from such a series into the complicated English representations that include embeddings, ellipsis, conjunction and pronouns. A prototype program by Correira & Hare (1978) translated the ordered set of propositions representing a brief story about Camelot into reasonable English using conjunctions, ellipsis and pronouns. This study showed us a need and a direction of research for a system to apply rhetorical rules to a sequence of propositions for planning sentences.

An analysis system that resolves anaphor and ellipsis to reduce sentence analyses to ordered sets of connected propositions is currently under development. So far it is successful in using rules associated with the verb to resolve such sequences as the following:

John had a book. Henry wanted it. He gave it to him.

" " He stole it from him.

It is presently being modified to account for a much larger series of examples of anaphor.

This inventory of software is presented to show a capability to represent and test linguistic and psychological formulations with computational rigor. When these systems reach a satisfactory level of development, we propose to develop a unified system that will generate and analyze texts from the morphological to the discourse level.

V Proposal

We propose a three year program of discourse study to the end of developing and testing discourse and rhetorical grammars that describe a variety of forms of discourse including narrative, essay, editorial, argument, technical and instructional manuals, and dialogues. The grammar form will probably remain the Horn clause organization that we have used successfully for narrative grammars. Existing programs will be used to test and debug the grammars.

We propose to study a wide range of text samples attempting to determine the discourse structure by linguistic analysis. Where this is not clear, we propose to undertake psychological experiments of the type described by Kintsch (1974), to determine the macrostructures that human readers obtain. These macrostructures reveal the discourse structure that people impose on the material they have read, and tend to bring clarity to the linguistic analysis.

The discourse grammars will be tested for consistency by using them to generate, analyze and summarize the texts they describe. They will be further tested for psychological validity by determining if human text readers agree with the propositions that the grammar infers in its task of establishing coherence of discourse.

Rules for a rhetorical grammar will be developed for combining elements of a series of propositions in such a manner that ellipsis, conjunction and the assignment of pronouns are marked so that the sentence generator may simply realize them as it develops a natural language string.

The reference analysis system that relates words to contexts, resolves anaphor and restores ellipsis will be developed to a satisfactory level. The development of this analytic system will be aided by the rhetorical work on generating conjunctions, pronouns and ellipsis.

We propose to construct a unified discourse modelling system that will use a single interpreter to analyze or generate from morphological to discourse level. This system will use an indexed taxonomic network of propositions and rules; it will incorporate a context that can be used like the chart parser to avoid redundant computations, and it will provide single or all-paths derivations with bottom-up or top-down control.

Throughout the accomplishment of these tasks we will be concerned with the interactions of word choice, sentence structure and discourse sequence, and seek to discover how to model their inter-relations.

Scientific papers will be published communicating the computational theory of discourse, associated software, and grammars. The unified system for discourse modelling will be made available to the scientific community.

REFERENCES

1. Austin, J. L. How to Do Things With Words, Oxford Univ. Press, 1962.
2. Beaugrande, R., Toward a General Theory of Creativity Poetics, 1979, (In Press).
3. Cohen, P.H., On Knowing What to Say: Planning Speech Acts, Ph.D. Thesis, Dept. Comp. Sci., Univ. of Toronto, 1978.
4. Compton's Pictured Encyclopedia, F.E. Compton & Co., Chicago, 1962.
5. Correira, A., Generating, Summarizing and Analyzing Narrative Discourse, Tech Report NL-36, Dept. Comp. Sci., Univ. of Texas, Austin, 1978.
6. Correira, A. and Hare, D., Generating Connected Natural Language from Case Predicate Story Trees, mss. from authors, Dept. Comp. Sci., Univ. of Texas, Austin, 1978.
7. Cullingford, R. E., Script Application: Computer Understanding of Newspaper Stories, Ph.D. Thesis, Dept. Comp. Sci., Yale Univ., 1978.
8. Grimes, J., The Thread of Discourse, Mouton, The Hague, 1975.
9. Grosz, B., The Representation and Use of Focus in Dialogue Understanding, SRI, Menlo Park, Calif., Tech. Note 151, 1977.
10. Halliday, M.A. and Hasan, R., Cohesion in English, Longman Press, London, 1976.
11. Hendrix, G.G., Partitioned Networks for Modelling Natural Language Semantics, Diss., Univ. of Texas, Dept. Comp. Sci., 1976.
12. Kintsch, W., Representation of Meaning in Memory, Wiley, New York, 1974.
13. Meehan, J., The Metanovel: Writing Stories by Computer, Dissertation, Yale, Dept. Comp. Sci., 1976.
14. Meyer, B., The Organization of Prose and its Effects on Recall, North Holland, Amsterdam, 1975.
15. Minsky, M., A Framework for Representing Knowledge, in Winston, P. (ed.) The Psychology of Computer Vision, McGraw-Hill, New York, 1975.
16. Novak, G., Computer Understanding of Physics Problems Stated in Natural Language, AJCL Microfiche 53, 1976.
17. Schank, R. and Abelson, R., Scripts, Plans, Goals and Understanding, Wiley, New York, 1977.
18. Searle, J.R. Speech Acts: An Essay in the Philosophy of Language, Cambridge Univ. Press, 1969.
19. Simmons, R. and Chester, D., Inferences in Quantified Semantic Networks, Proc. 5IJCAI, 1977.
20. Simmons, R. F. and Correira, A., Rule Forms for Verse, Sentences and Story Trees, Tech. Report NL-35, Dept. of Comp. Sci., Univ. of Texas, Austin, 1978.
21. Simmons, R. & Novak, G., Semantically Analyzing an English Subset for the Clowns Microworld. AJCL Microfiche 18, 1975.

22. Simmons, R. F., Rulebased Computations on English, in Hayes-Roth, R. and Waterman, D. (eds.) Pattern-Directed Inference Systems Academic Press, New York, 1978, (in press).
23. Smith, M.K., Procedural Automata in the Clown's World, Proc 4th Tex. Conf. on Comp. Sys., Univ. of Texas, Austin, 1975.
24. Thorndike, P., Cognitive Structures in Comprehension and Memory of Narrative Discourse, Cognitive Psychology, 9, 1, pp. 77-110, 1977.
25. Tinlap-2, Theoretical Issues in Natural Language Processing-2, Proceedings, ACM, PO Box 12105, Church St. Station, New York City, 1978.
26. Van Dijk, T.A., Recalling and Summarizing Complex Discourse, mss., Univ. Amsterdam, Dept. of Gen. Lit. Studies, 1975.
27. Winograd, T., Understanding Natural Language, Academic Press, New York, 1972.
28. Wilks, Y., Parsing English, in Charniak, E and Wilks, Y. (eds.) Computational Semantics, North-Holland, New York, 1976.