SP-3132

A COMPUTATIONAL MODEL OF VERBAL UNDERSTANDING

Robert F. Simmons, John F. Burger, Robert M. Schwarcz

30 April 1968

SP-3132

SP *a professional paper*

A Computational Model of Verbal Understanding

Robert F. Simmons
John F. Burger
Robert M. Schwarcz

30 April 1968

SYSTEM
DEVELOPMENT
CORPORATION
2500 COLORADO AVE
SANTA MONICA
CALIFORNIA
90406

SDC

A 1162 REV

ABSTRACT

This paper presents a theory of verbal understanding based on a formal model of conceptual structures that represent verbal meanings expressed in English sentences. Verbal understanding is defined as the capability for disambiguation, paraphrase, question answering, translation, etc., with regard to natural language sentences. The model has been implemented as Protosynthex III in LISP on the Q-32 time-shared system. Experimental results from the system include examples of the analysis of complex sentences, disambiguation of multisensed words via sentence context, question answering via logical inference, and meaning-preserving paraphrase generation. The authors conclude that sophisticated natural language processing by computers is a realistic goal that has been partly achieved. The rate of progress toward complete achievement is seen to be proportional to the amount of developmental support available.

ACKNOWLEDGMENTS

The research reported in this paper results from a continued long-term effort--the Synthex Research Project--begun at SDC in the summer of 1960, financed by combinations of funds from SDC and, until recently, from ARPA. During this period of eight years many language processing systems and subsystems have been constructed including Protosynthex I, II and the Protosynthex III described in this paper. In developing these systems we are indebted directly to dozens of programmers and researchers who at one time or another have contributed to the project. We are further indebted to many colleagues at SDC and throughout the country who formally or informally have served as consultants but most grateful of all to the stable, well-financed, research environment and the psychological support provided for us by SDC and its research management.

BLANK PAGE

A Computational Model of Verbal Understanding

Robert F. Simmons
John F. Burger
Robert M. Schwarcz

I. Introduction and Background

The long-term goal for computational linguistics is to increase our understanding of linguistic and conceptual structures and to formally describe them so that computers can deal effectively with natural languages in such applications as question answering, stylistic and content analysis, essay writing, automated translation, etc. The eventual realization of this goal requires not only a satisfactory model of linguistic structures, but also models for verbal understanding and verbal meaning. In this paper we outline a theory and a model of verbal understanding and describe Protosynthex III, an experimental implementation of the model in the form of a general-purpose language processing system. The effectiveness of the model in representing the process of verbal understanding is demonstrated in terms of Protosynthex III's capability to disambiguate English sentences, to answer a range of English questions and to derive and generate meaning-preserving paraphrases.

Background: Computational linguistics is fortunately a field in which there is no dearth of state-of-the-art surveys. Over the last three years, Bobrow, Fraser and Quilliar [1967], Kuno [1966] and Simmons [1966] have independently reviewed recent relevant literature in structural linguistics, semantics, psycholinguistics and computer language processing. A critical survey is even now in press by Salton [1968] to cover most recent trends. A survey of question-answering systems by Simmons [1965] describes the earlier developments in that area.

Several very recent lines of research by Quillian [1965, 1966], Colby [1966, 1967], Bohnert [1966], Abelson [1965], Green and Raphael [1967] and Simmons, et al. [1966, 1967, 1968] have introduced ideas of deep logical and/or conceptual structures to represent understanding of phrases and sentences from natural language. Theoretical papers by Katz [1967], Woods [1966] and Schwarcz [1967] and experimental work by Kellogg [1967a, b] have advanced our understanding of how to accomplish various forms of semantic analysis. Recent papers by Kay [1966, 1967] have been of great value in explicating and generalizing computational methods for syntactic analysis with particular reference to various forms of transformations.

These surveys and recent lines of research lead to the conclusions that the field of computational linguistics is a very active one, developing computational techniques at a rate that keeps pace with the advances in structural linguistic theory. Unfortunately, excepting for the Abelson and Colby models and cognitively oriented works by Miller et al. [1960], Deese [1967] and Reitman [1965], there appears still to be a significant lack of psychological theory of verbal understanding to guide computational experimentation.

II. A Representation of Deep Conceptual Structure

Such operations as semantic analysis, question answering, paraphrase and mechanical translation each require the explication or transformation of concepts that are signaled or communicated by sentences in natural language. The concepts being communicated via language are not the words nor the phrases nor any other explicit structure of a discourse. Instead, what is being communicated is some set of relations among cognitive structures (i.e., ideas) that are held in common between a speaker and a hearer of the language. The linguistic notion of deep syntactic structure is a partial recognition of this fact, but for computers to demonstrate "verbal understanding" and manipulate "verbal meanings," an even deeper level of conceptual structure must be represented. This deep conceptual structure serves as a partial model of verbal cognition, i.e., of how a human understands and generates meanings communicated by language. The effectiveness of a model of verbal understanding can be evaluated in terms of how well it supports such criterial operations as disambiguation, question answering, paraphrase, verbal analogies, etc. Whether the model truly represents the operations that humans actually use is another question and one to be studied by psychological experiment.

We thus define verbal understanding as the capability of a system to disambiguate, paraphrase, translate and answer questions in and from natural language expressions. Verbal meaning is defined as the set of interrelations in the model among linguistic, semantic and conceptual elements that provides this competence.

Our general model of understanding derives from a theory of structure proposed by Allport [1955] in the context of psychological theories of perception. Our models also owe a conceptual debt to such widely varying sources as Chomsky's [1965] theory of deep syntactic structure, Quillian's [1966] semantic nets and most recently, to Fillmore [1966, 1967] who proposes a significant variation to the Chomsky deep structure.

The primitive elements of our general model are <u>concepts</u> and <u>relations</u>. A concept is defined either as a primitive object in the system or as a concept-relation-concept (C-R-C) triple. In the model of verbal understanding, a concept that is a primitive object corresponds to a meaning or word sense for a word. But even these "primitives" can be defined as a structure of C-R-C triples that can be transformed to a verbal definition. A relation can also be either a primitive object or a C-R-C triple. Ideally, all relations should be primitive and well-defined by a set of properties such as transitivity, reflexivity, etc. Since each property corresponds to a rule of deductive inference, well-defined relations are most useful in making the inferences required for answering questions or solving verbal problems. Any relation, primitive or complex, can be defined in extension by the set of pairs of events that it connects. However, unless the relation is definable intensionally by a set of deductive properties, its use in inference procedures is generally limited to the substitution of equivalent alternate forms of expression.

Any perception, fact or happening, no matter how complex, can be represented as a single concept that can be expanded into a nested structure of

C-R-C triples.* The entire structure of a person's knowledge at the cognitive or conceptual level can thus be expressed as a single concept or event; or at the base of the nervous system, the excitation of two connected neurons may also be conceived as an event that at deeper levels may be described as sets of molecular events in relation to other molecular events.

Meaning in this system (as in Quillian's) is defined as the complete set of relations that link a concept to other concepts. Two concepts are exactly equivalent in meaning only if they have exactly the same set of relational connections to exactly the same set of concepts. From this definition it is obvious that no two nodes of the concept structure are likely to have precisely the same meaning. A concept is equivalent in meaning to another if there exists a transformation rule with one concept as its left half and the other as its right. The degree of similarity of two concepts can be measured in terms of the number of relations to other concepts that they share in common. Two English statements are equivalent in meaning either if their cognitive representation in concept structure is identical, or if one can be transformed to the other by a set of meaning preserving transformations (i.e., inference rules) in the system.

English sentences can be mapped onto the deep conceptual structure of this model of verbal understanding by considering prepositions, conjunctions and verbs as relational terms, and nouns, adjectives and adverbs as conceptual objects. Thus, a sentence such as "The angry pitcher struck the careless batter" can be expressed in the following set of relational triples:

---

*From a logician's point of view, the C-R-C structure can be seen as a nested set of binary relations of the form R(C,C); the referenced statement is a claim that any event can be described in a formal language of such triples.

A. (((pitcher MOD angry) TMOD the) struck ((batter MOD careless) TMOD the))

As it stands, this is simply a form of syntactic diagramming of the sentence (where MOD and TMOD are modificational relations). However, by using the semantic analysis procedure to be described in Section III, the selection of word sense meanings is made explicit as follows (SUP means "has as a semantic superclass"):

B. (((pitcher SUP player) MOD (angry SUP emotion) TMOD the)
(struck SUP hit)
((batter SUP player) MOD (careless SUP attitude) TMOD the)).

The particular sense of "pitcher" is the one that is "a kind of player"; the sense of "strike" is "to hit" and the sense for "batter" is "player." The complex element (struck SUP hit) is the relational term for the larger triple

((pitcher, etc.) (struck SUP hit) (batter, etc.)).

When the triple structure B is embedded in the conceptual model, it can be roughly represented by the graph of Figure 1.

The result of embedding the sentence in the conceptual structure is to make explicit many aspects of verbal meaning that were implicit in the selection and ordering of words in the English sentence. Without any analysis or context the example sentence would answer only the question "Is it true that the angry pitcher struck the careless batter?" With such a relational analysis and embedding in the conceptual structure a whole range of new questions can be answered--for example:

Is a pitcher a person?
Is a batter a baseball player?
Did a baseball player hit a person?
Do persons have attitudes?
etc.

However, Figure 1 is only an approximate representation of the actual conceptual structure. The subscripts on each word in Figure 1 represent the word sense and concept selection appropriate to the sentence. In the actual structure a concept number occurs for each word on the graph. Each unique sense of meaning for a word corresponds to exactly one concept number; but each concept number may map onto more than one word sense <u>and</u> onto a defining structure of concepts. For example, the words "young" and "youthful" share a sense meaning in common, viz., "having the characteristics of youth." In each case this sense meaning corresponds to a concept number, say C72. C72 might be defined by the structure (C72 EQUIV (CO C42 C55)) which translates into "C72 is equivalent to something having youth."

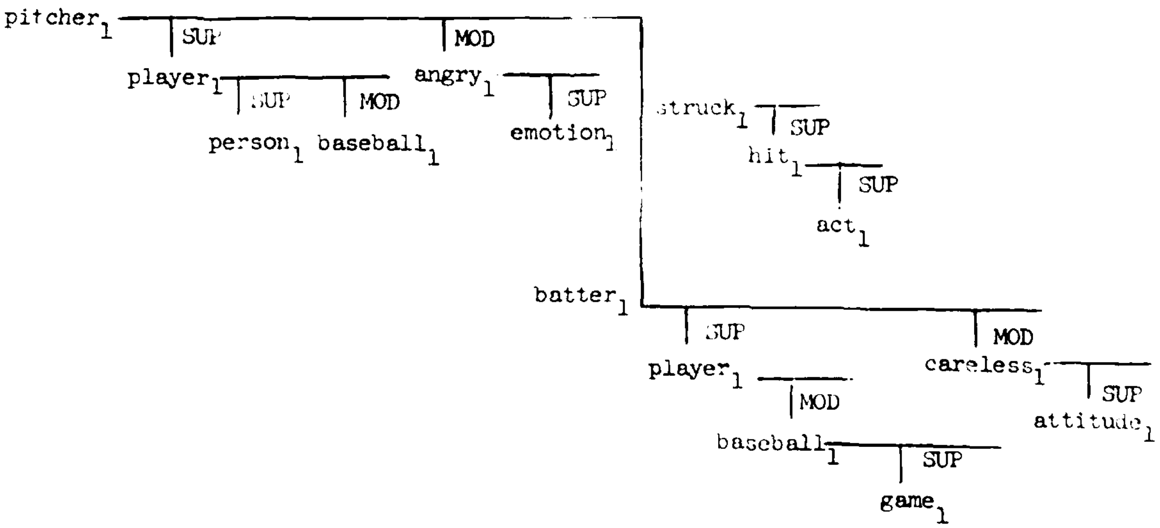


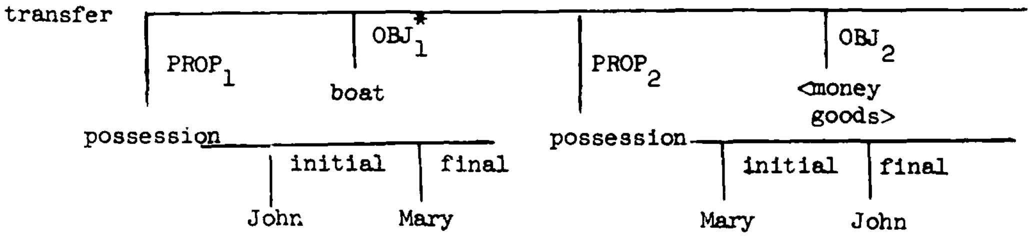
Figure 1. A Graph of Conceptual Structure

What the conceptual structure does is to allow word meanings to be represented by a single conceptual object but, at the same time, to allow a conceptual object to be expressible in many different verbal forms. The conceptual level is necessary for paraphrase and translation operations. For example, the English expressions "old man" and "ancient" and the French word "vieux" can all be expressions of a single concept which we will label C37. The structure (man_c MOD old_c)--where the subscripted "c" means the concept number--is a defining term for C37 which is one of the word senses for "ancient" and for the French word "vieux." When the semantic analysis system produces (man_c MOD old_c) it tests to discover whether the triple can be expressed, as in this case it can, as a single concept. In the generation system that concept, C37, can be expressed by any of its mappings onto word senses and thus onto words.

This particular version of a structure for verbal understanding is our current model. It has shown itself strong enough to support many kinds of verbal understanding operations, but in our experimentation with it, we have found that it is not as deep a structure as we would like to have. In this structure, for example, the equivalence of the two statements (a) "Mary bought a boat from John," and (b) "John sold a boat to Mary" can only be discovered by a transformation rule of the following form:

((X (buy from Z) Y) EQUIV (Z (sell to X) Y)) .

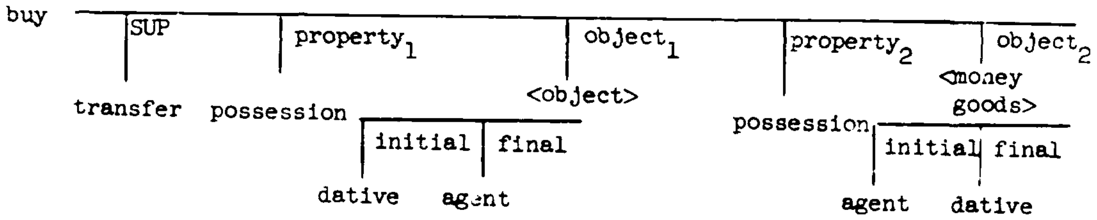
Using lexical data in this form would allow analysis of the example sentence, b, into something like the following deep conceptual structure:

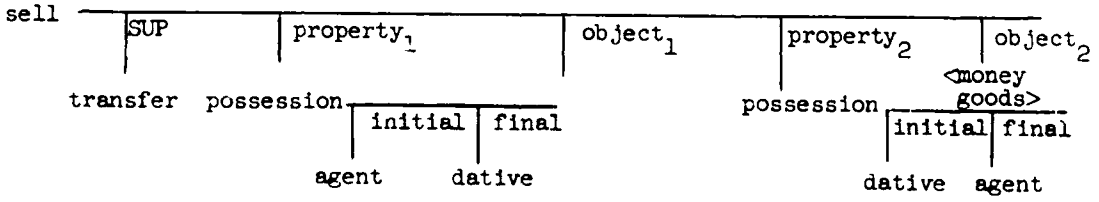


By changing the relative positions of Mary and John in the structure, the transfer can be expressed as "buy" or "sell" and the identity of meaning for the two expressions a and b is made explicit. The transfer concept could equally well express ideas of "get," "take," "give," "exchange," "borrow," "steal," etc., as partially analyzed by Bendix [1966]. The additional advantage of this kind of structure is that it suggests that all relations in the cognitive model can indeed be well-defined by contrast to our present unsatisfactory mix of well- and poorly-defined relations. However, the detailed definition of such improved structures remains still to be done as a later piece of research than the one reported here. It is mentioned in passing to demonstrate that the deeper the conceptual structure used, the more explicit become the meanings expressed in English sentences.

Assuming for the moment our assertion that the model we have first described does support the criterial operations of verbal understanding, the important question is: By what means can we transform English sentences into such a conceptual structure? Section III immediately following, describes a method of syntactic and semantic analysis that accomplishes the transformation. Section IV describes experiments to test the system's capabilities for disambiguation, paraphrase and question answering.

Following a line of thought suggested by Fillmore [1967], we believe a deeper conceptual structure than we are now using might very well express the concepts of "buy" and "sell" as examples of "transfer" somewhat as follows:





III. Analytic Method

The representation of the underlying deep conceptual structure described in the preceding section makes possible various meaning-preserving operations on the concepts that are communicated by natural language statements. The problem of transforming from English statements and questions into that structure requires both syntactic and semantic analysis. In an earlier paper [Simmons and Burger, 1968] we described a method of semantic analysis that worked well for fairly uncomplicated sentences but, for lack of powerful transformational rules, was weak with regard to certain complicated structures. The method described here introduces transformational machinery that has proved adequate for the most difficult sentence structures that we have experimented with.

The method of analysis requires a lexicon, a grammar that includes transformations, a set of semantic event forms (SEFs), and a modified Cocke algorithm to actually carry out the analysis. In brief, the method finds immediate constituents of the surface structure of the sentence, transforms these into the form of deep conceptual triples and tests each such triple for semantic well-formedness. The resulting analysis is a bracketed structure of triples with each element marked for its selection of word sense meaning or concept. All analyses that are allowed by the grammar and SEFs are produced. A person operating the system is given the opportunity to select any one or several interpretations to be stored in the conceptual model.

The Lexicon: The lexicon is composed of word and concept entries. With each English word entry, a set of word sense meanings is associated; each word sense, in turn, is associated with a syntactic class, a set of syntactic features, a chain of semantic word classes and a concept. For each concept entry, there may be a pointer to one or more word senses that may be used to express that concept verbally and an equivalence relation to one or more concept structures that represent its meaning. Some concepts, however, are not expressible as single word senses and are only verbally expressible by deriving the word senses for a concept structure to which they are equivalent. In addition to these elements, each concept entry has pointers to its tokens in the data structures where it has been used.

The semantic word classes that characterize each word are a chain of concepts that are in a linguistic superset relation. To explain by example, the word "pitcher" is characterized by two word senses and thus two different chains of semantic classes as follows:

pitcher ... N, player, person, mammal
... N, container, physical object, object

The first superset chain (or SUP-chain) means that "pitcher is a kind of player is a kind of person is a kind of mammal." This is usually expressed as "pitcher SUP player SUP person...." Actually in place of the words for semantic classes, the lexicon contains concept numbers that usually refer to particular word senses. A more complete example of dictionary structure is presented in Figure 2.

	SENSE NBR	SYNTACTIC CLASS	SYNTACTIC FEATURES	SEMANTIC CLASSES	CONCEPT NBR
BOYCOTT	41	N	SING	123,200	53
	42	V	PL, PR	123,200	53
DISCOVERY	55	N	SING	98	67
DISCOVER	56	V	PL, PR	98	67
FIND	34	N	SING	---	98
	35	V	PL, PR	---	98
REBEL	150	V	PL, PR	---	123
	151	N	SING	---	124
STRIKE	207	N	SING	67, 98	100
	208	V	PL, PR	53, 123	55

ABBREVIATIONS:
SING = Singular
PL = Plural
PR = Present Tense

Figure 2. A Fragment of Lexical Structure

A concept is created for the system for each new word sense and for each occasion when an equivalence relation occurs. Since every word sense can be defined by a dictionary definition that can be substituted in contexts where the word in that sense is used, it follows that every word sense concept is in an equivalence relation to some other concept structure that expresses its meaning. In the actual system, not every concept need be so defined, although the power of the system for verbal understanding obviously increases with the number of concepts that are defined.

The concept entry for the second sense of "strike" in Figure 2, i.e., concept number 55, would appear as follows:

Concept	Word Senses	Meaning	Used/in
C55	208	(C0 C89 C251)	G42, G45 ... etc.

This example shows that C55 may be expressed verbally by sense 208 that corresponds to the singular, present tense, verb "strike." By looking up C0, C89 and C251, it can be discovered that the meaning of C55 can also be expressed by the words "to stop work." The list of G-prefixed numbers in the Used/in column are simply pointers to data structures in which the concept C55 has been used to make factual statements.

The aim of this form of lexical structure is to distinguish clearly between linguistic and conceptual information. Syntactic classes and features* are defined as those elements which are required by the grammar and are clearly linguistic in nature. Semantic classes are expressed as concepts and are in a borderline area between the linguistic structure and the deep conceptual structure. Semantic classes are elements of the semantic event forms, but are also concepts that can occur anywhere in the deep conceptual structure.

<u>The Grammar</u>: For discovering immediate syntactic constituents for a sentence and transforming them directly into the conceptual structure, we use a form of rule that combines phrase structure rewrite rules with a transformation. The form of this grammar can be understood by a simple example.

(a) adj + noun —> (B MOD A) NP = (NP (noun MOD adj))

---

*Although the lexical and conceptual structures provide for treatment of tense and agreement based on features, the analysis, generation and question-answering algorithms do not yet use this information.

The phrase structure component states that an adjective followed by a noun can be rewritten as a Noun Phrase. The transformation requires that the Bth or second element of the left side be written first followed by the term MOD followed by the Ath or first element.

A more complex example to account for a certain type of discontinuity is illustrated below.

(b) adv + S —> (BA (BB MOD A) BC) S
= (S (Subject (verb phrase MOD adverb) object))

The transformation of (b) states that the BAth element of the left side is to be written first. The Bth element is S; S always breaks down into a triple whose Ath element is a noun phrase, whose Bth element is a verb phrase, and whose Cth element is an object noun phrase or an explicit null symbol. Thus the BAth element is the Ath element of the Bth element, or the subject of the S term. Similarly the BBth element is the verb phrase and the BCth element is whatever is in the object position.

A simple grammar to account for the sentence "the angry pitcher hit the careless batter" is presented in (c) below:

(c) adj + noun —> (B MOD A) NP
art + NP —> (B TMOD A) NP
verb + NP —> (D A B) VP
NP + VP —> (A BB BC) S

The string of syntactic word classes corresponding to these words is as follows:

art + adj + noun + verb + art + adj + noun

The analysis that results from this grammar is as follows:

(S (((noun MOD adj) TMOD art) verb ((noun MOD adj) TMOD art)))

In a previous paper [Simmons and Burger, 1968] we showed that this type of structure can be obtained by applying transformations to the elements of a phrase structure analysis of a sentence. That is precisely what the combined phrase structure and transformation rules of this type of grammar accomplish as each constituent of the sentence is discovered.

There is no theoretical limit to the depths to which the transformational notation can refer; strings such as ABBBCAB can be written to refer to the first, second, or third element of the nth level of depth of structure. Certain elements of the transformations such as MOD, D (an explicit dummy marker) and the brackets are taken literally; only combinations of the terms A, B, and C refer to the structure of the left-hand side. The elements in a rule can be semantic classes on which the transformation can operate, and the resulting constituent can be a composition function of the semantic classes. For example, a rule might be written to analyze the phrases "park bench" and "wooden bench" as follows:

place + furniture —> (B LOC A) furniture-LOC

and material + furniture —> (B TYPE A) furniture-TYPE

Compositions such as furniture-LOC imply a controlled combination of the SUP-chains for the two elements in a manner such as that described by Katz [1967] or used by Kellogg [1967a].

Rules of this kind would eliminate the set of SEFs and the separate check for semantic acceptability. The disadvantage would be an enormous increase in the number of rules. Consequently, we have so far preferred to keep separate the syntactic and semantic components of the system.

We have understood the transformational component of these rules as a program that operates an interpreter whose data is the structure identified by the names of the two constituents on the left-hand side. Recently, a paper by Kimball [1967] has proved that a certain modified form of Chomsky's transformations can completely imitate the operation of a Turing machine. Because of the limit to two elements on the left side, our form of transformation is less powerful than this.

In applying this form of grammar to a surface string of syntactic information we make the assumption that just the information cued by syntactic word classes, and by syntactic and semantic features, is sufficient to allow transformation to the underlying deep conceptual structure. It has so far been possible to write transformations that account for very complex sentences (as will be seen in Section IV), but our only defense against counterexamples is to attempt to demonstrate that a grammar can be built to account for each challenge.

Semantic Event Forms: As each constituent is discovered and transformed according to the grammar, the result of the transformation is tested for semantic well-formedness. The SUP-chain of semantic classes and a set of semantic event form (SEF) triples whose elements are semantic class terms are required for making this test. By considering our example sentence again, the elements and method of this test can be explained.

A. The angry pitcher hit the careless batter.

When "pitcher" was looked up in the lexicon, two word senses were discovered and both of these were nouns; for "angry" there was only one. Thus, two constituents of the form "adj + noun" were discovered to represent "angry pitcher." The SUP-chain of semantic classes that represented each sense of the words was then called into use to form the following pair of complex triples:

pitcher	MOD	angry
player		emotion
person		feeling
animal		sense
mammal		

and

pitcher	MOD	angry
container		emotion
physical obj		feeling
object		sense

Thus a complex triple is one whose elements are SUP-chains of the elements in a simple triple. From the total set of SEFs, the possibly relevant ones are those which contain one or more elements that are included in any of the complex triples of the sentence. This subset of SEFs include among others the following:

(ANIMAL MOD EMOTION)
(PERSON MOD ATTITUDE)
(PHYSOBJ MOD QUALITY)
(PERSON HIT PERSON)
(OBJECT HIT PERSON)
(PERSON BOYCOTT ORGANIZATION)
ETC.

The test for semantic well-formedness is to discover whether any triple of elements, selected one from each SUP-chain in a complex triple, corresponds to an SEF. In the present example, the combination (animal MOD emotion) from the first complex triple does correspond to an SEF in the list. No combination of elements from the second complex triple corresponds to an SEF, so the sense of "pitcher" as a "container" does not apply to the constituent (N MOD Adj) for that sense and it is rejected. For the acceptable sense, "pitcher" as "person," the constituent is kept and elsewhere it is stored as ((pitcher SUP player) MOD (angry SUP emotion)). In subsequent constituents using this complex constituent, the SUP-chain of semantic classes for the head element "pitcher SUP person" is used to stand for the entire constituent.

The result of these semantic tests is to reject many syntactic constituents that would otherwise lead to multiple interpretations of the sentence. For example, if we consider the number of common meanings for "pitcher," "struck," and "batter" to be respectively 2, 3, and 2 there would be 12 possible interpretations of the sentence. By use of the three SEFs (ANIMAL MOD EMOTION) (PERSON MOD ATTITUDE) and (PERSON HIT PERSON) only the one interpretation presented below survives the analysis process.

(((pitcher . person) MOD (angry . emotion)) (struck . hit)
((batter . person) MOD (careless . attitude)))

The dot pairs are used for conciseness in representing (concept SUP concept).

A whole series of questions arises at this point: What is an SEF? How many will be required to deal with a large subset of English? How does one select the level at which to write them? These and others are questions that we have considered at length and we will try to summarize our present understanding.

It appears to us that an SEF is an abstraction of some element of lexical information that should (in a more sophisticated system) be directly a part of the lexicon. It appears to be an abstraction expressed in terms of semantic classes of the set of features that characterize a word's combinatorial possibilities in ordinary usage in the language. For example, the SEF (ANIMAL MOD

SENSE) indicates the relationship expressed by linguists in terms of a restriction on sensory verbs and adjectives to co-occurence with subjects marked by the feature " + animate." We believe, for the present state of computational linguistics, we can better represent such linguistic data in the form of acceptable combinations of semantic classes for words--i.e., SEFs--and later, from the useful SEFs work out underlying features.

We have no answer to the question of how many SEFs would be required to cover a large subset of English. A related device, the semantic message forms [Wilks,1968] are based on approximately fifty semantic classes and believed by CLRU* researchers to allow sufficient combinations to account for all English forms. We are currently tending toward the belief that although the separate SEF set provides adequate machinery for relatively small subsets of English, this information must eventually become an integral part of the lexicon to avoid very large space and time requirements in semantic analysis of large sets of English.

Selecting the level at which to write an SEF is hardly more easily dealt with. Considering each SEF as a rule of semantic combination, the task is very much like that of preparing a grammar. One attempts to obtain the minimal number of SEF rules that will distinguish acceptable and nonacceptable combinations of word senses. The elements of each rule are selected at the highest level of semantic abstraction that will successfully distinguish all word senses that are in a superset relation to--i.e., subclasses of--those elements. Thus, in coining the SEF (ANIMAL MOD EMOTION) we are stating our understanding that the nature of these concepts is such that anything that is an emotion is

---

*Cambridge Language Research Unit, Cambridge, England restricted to modifying only those things that are animals. Similarly in (PERSON BOYCOTT ORGANIZATION) we restrict the concepts that are kinds of "boycott" to co-occurrence with things that are persons as subjects and things that are organizations to receive the action.

In favor of the SEF approach, we have found them simple to build and use and of the same functional utility as the semantic markers and selection restrictions of Katz's [1967] current semantic theory. Something approaching the function of his projection rules can be seen in our use of the semantic class of the head of a construction to stand for the semantic classes of the whole. However, we claim only that the SEF approach is a first approximation to expressing semantic information that should be an integral part of a lexical entry for a word sense.

The Analysis Algorithm: After several experiments in producing various forms of recognition algorithms, we finally concluded that the Cocke algorithm was superior in respect to conciseness, completeness and efficiency of computation. This algorithm has been presented in ALGOL and described in detail by Kay [1967]. Our modifications have been only to add more tests on each constituent for agreement and semantic well-formedness and to introduce transformations into the operation of the grammar.

The essential operation of the algorithm is to test--exhaustively, but efficiently--each adjacent pair of elements in a sentence structure to discover if they form constituents acceptable to the grammar. If they do, the pair of constituents are rewritten according to the grammar rule. The process continues until all elements of the sentence are encompassed by at least one single constituent usually named S. All interpretations acceptable to the grammar are so formed.

IV. Results

The complete language processing system that has been described has been programmed as Protosynthex III in LISP 1.5 for the SDC Q-32 time-shared computer. The semantic analysis system has also been programmed in JOVIAL and used to prepare the examples on pp. 33, 34. It includes the capability to syntactically and semantically analyze single sentences into the formal language of the conceptual structure. From the resulting conceptual structure, the system is able to answer a range of English questions using logical inference procedures based on properties associated with the well-defined relations. It is also able to paraphrase by finding equivalence relations among concepts and to generate English sentences in accordance with a generation grammar. In this paper a limited set of examples of these operations will be presented; additional computer printouts of examples have been collected as a special supplement that is available on request from the authors.

Syntactic and Semantic Analysis: The grammar reproduced in Figure 3 has proved sufficient to account for the analysis of the sentences in the following paragraph about physiological psychology of the eye.

> The eye is the organ of sight. The retina ia the light sensitive surface of the eye. Cones and rods are the special sensors in the retina. Cones and rods react to light. When we see anything, we see light reflected from the objects we look at. Reflected light passes through the lens and falls on the retina of the eye. Seeing an object actually means seeing the reaction of our retina.

The sentences comprising this paragraph were selected to represent a range of fairly difficult structures including various kinds of embeddings.

```
(((QPP QBEAJ) [BA BB (BC AB AC)] S)
 ((DVBE ADJ) (AC MOD B) QBEAJ)
 ((VBEQ NP) [BA BB BC] NP)
 ((NNP DVBE) [[AA EQUIV BC] AB [AC EQUIV BC]] S)
 ((VDO S) [BA BB BC] S)
 ((VDO SI) [BA BB BC] S)
 ((QNDO S) [BA BB AA] S)
 ((QNDO SI) [BA BB AA] S)
 ((QN DVBE) [BC EQUIV A] S)
 ((QN VDO) (A B D) QNDO)
 ((QAVDO S) [BA (BB MOD AA) BC] S)
 ((QAV VDO) (A B D) QAVDO)
 ((ADJ PP) (A BB BC) ADJ)
 ((ADJ NP) (B MOD A) NP)
 ((VCOMP PRTVCOMP) (A BB BC) CONJVCOMP)
 ((CONJ NP) (D A B) NCOMP)
 ((CONJ VCOMP) (D A B) PRTVCOMP)
 ((PREP QN) (D A B) QPP)
 ((PREP DNP) (D A B) PP)
 ((PREP NP) (D A B) PP)
 ((RFLAVB CONJS) [BA A BC] S)
 ((S CONJSPRT) [A BB BC] CONJS)
 ((COM SI) (D A B) CONJSPRT)
 ((COM S) (D A B) CONJSPRT)
 ((VED NP) (B SMOD [*OBJECT A B]) NP)
 ((VED PP) (A BB BC) VCOMPED)
 ((ART NP) (B TMOD A) NP)
 ((NP NCOMP) (A BB BC) NNP)
 ((NP V) [A B ****] S)
 ((NP VCOMP) [A B ****] S)
 ((NP PRED) [A BB BC] S)
 ((NP VPREP) [A B ****] SPREP)
 ((NP SPREP) (A SMOD [BA (BBA BBB A) BC]) NP)
 ((NP CONJVCOMP) [[A BA ****] BB [A BC ****]] CONJS)
 ((NP PP) (A BB BC) NP)
 ((NP VCOMPED) (A SMOD [*OBJECT B A]) NP)
 ((NP CONJSPRT) [A BB BC] CONJS)
 ((V PP) (A BB BC) VCOMP)
 ((V DNP) (D A B) PRED)
 ((V NP) (D A B) PRED)
 ((V PREP) (A B D) VPREP)
 ((DNP DVBE) [A EQUIV BC] S)
 ((DNP PRED) [A BB BC] S)
 ((VBE RELPR) (D A B) VBEQ)
 ((VBE DNP) (D A B) DVBE)
 ((VBE NP) (D A B) PRED)
 ((DART NP) (B TMOD A) DNP)
 ((SEQUIV **) (D A B) B*SEQ)
 ((** B*SEQ) [A BB BC] S))
NIL
```

Figure 3. Recognition Grammar Rules

Example 1, below, shows the manner of inputting the sentence and dictionary data into the system.

```
       ANALYSIS MODE //

 READY--
THE EYE IS THE ORGAN OF SIGHT . . BODYPRT FUNCTION VISION .
 ORGAN EQUIV BODYPRT .
 (SUPS- THE EYE IS THE ORGAN OF SIGHT)
 DDET ORGAN EQUIV DDET BODYPRT FUNCTION VISION
 (SUPS- VISION FUNCTION BODYPRT DDET ORGAN DDET)
 COGACT ASSOC OBJECT *TOP BODYPRT *TOP
 (SUPS- BODYPRT OBJECT ASSOC COGACT)
 OBJECT *TOP *TOP ACT
 (SUPS- ACT OBJECT)
 DO *TOP
 (SUPS- DO)
 *TOP
 (WCS- (SIGHT . VISION)
   (OF . FUNCTION)
   (ORGAN . BODYPRT)
   (THE . DDET) (IS . EQUIV) (EYE . ORGAN) (THE . DDET))
 NP PREP NP DART VBE NP DART
 1
P
 [((EYE . ORGAN) TMOD (THE . DDET))
  (EQUIV . PRIMIT)
  (((ORGAN . BODYPRT) (OF . FUNCTION) (SIGHT . VISION))
   TMOD (THE . DDET))]
```

Example 1.

A sentence is typed in followed by a period. Optionally a set of supersets for each word of the sentence can then be input followed by a period. Following this second period, SEF triples can be given to the system as was done in Example 1. The third period--i.e., the one following the SEFs--is taken by the system to mean completion of input. At that point the system looks up each word in the dictionary to obtain superset classes and syntactic word classes. If it does find these, it requests SUPS for each word that is not in the dictionary. It reiterates this request for semantic word classes until each word has developed a SUP-chain that terminates with the symbol, *TOP. It then asks for syntactic word-classes with the request "WCS-" identifying the word sense by using the dot pair (word . superclass). When all these data are present--either already in the dictionary or having been input with the sentence--the system computes its semantic analysis using the grammar and SEFs available to it. In Example 1, the bracketed structure shows the syntactic analysis and the selection of word senses for each word in the sentence. This example shows that the system correctly transformed "is" in the context "The NP is the NP" into the well-defined relation "EQUIV." The relation "TMOD" is used by the system to alert it to the presence of an article. If the article is definite it refers to a particular or already existing token of data; if the article is indefinite or absent it is understood to represent any token or instance of its concept.

The most complex sentence of the paragraph is presented as Example 2, below. The analysis of this sentence shows four embedded sentences each of which is surrounded by square brackets. The first of these, "We see light" is in an IMPLY relation to the remainder. The expression "...light reflected from..." gives rise to a noun phrase that is modified by the sentence "*object reflected ... light," where "*object" stands for "something." The phrase "...from objects we look at " gives rise to the structure (object SMOD we (look at objects) ****), a noun modified by an intransitive sentence that uses that noun as the object of a preposition. By following the syntactic word class pairs through the grammar of Figure 3, the interested reader can observe the application of relatively simple transformations to compute these structures.

```
 READY--

WHEN WE SEE ANYTHING , WE SEE LIGHT REFLECTED FROM THE OBJECTS
 WE LOOK AT . . COGACT TOWARD OBJECT , PERSON COGACT ****,
 THROWBACK SOURCE OBJECT, *OBJECT THROWBACK RADIATION,
 PERSON COGACT OBJECT , PERSON COGACT RADIATION"
PERSON COGACT OBJECT , PERSON COGACT ENERGY, SEN COM SEN,
 SEN IMPLY SEN .
 (SUPS- WHEN WE SEE ANYTHING , WE SEE LIGHT REFLECTED FROM OBJECTS WE
  LOOK AT)
 IMPLY PERSON COGACT OBJECT COM PERSON COGACT RADIATION THROWBACK
 SOURCE OBJECT PERSON COGACT TOWARD
 (SUPS- TOWARD PERSON SOURCE THROWBACK RADIATION PERSON COM PERSON
  IMPLY)
 DIRECTION ANIMAL LOC ACT ENERGY ANIMAL *TOP ANIMAL *TOP
 (SUPS- ANIMAL ANIMAL ENERGY LOC ANIMAL DIRECTION)
 OBJECT OBJECT *TOP *TOP OBJECT LOC
 (SUPS- LOC)
 *TOP
 (WCS- (AT . TOWARD)
   (LOOK . COGACT)
   (WE . PERSON)
   (OBJECTS . OBJECT)
   (FROM . SOURCE)
   (REFLECTED . THROWBACK)
   (LIGHT . RADIATION)
   (SEE . COGACT)
   (WE . PERSON)
   (, . COM)
   (ANYTHING . OBJECT) (SEE . COGACT) (WE . PERSON) (WHEN . IMPLY))
 PREP V NP NP PREP VED NP V NP COM NP V NP RELAVB
 1

PRINT
 [[(WE . PERSON) (SEE . COGACT) (ANYTHING . OBJECT)]
  (WHEN . IMPLY)
  [(WE . PERSON)
   (SEE . COGACT)
   ((LIGHT . RADIATION)
    SMOD [(*OBJECT . PRIMIT)
     ((REFLECTED . THROWBACK)
      (FROM . SOURCE)
      (((OBJECTS . OBJECT)
        SMOD [(WE . PERSON)
         ((LOOK . COGACT) (AT . TOWARD) (OBJECTS . OBJECT))
         (**** . PRIMIT)])
       IMOD (THE . DDET)))
     (LIGHT . RADIATION)])]]
 1
```

Example 2.

Examples of the analysis of question structures are shown in Example 3, below. In some cases the question word is deleted while in others the question is transformed to declarative structure.

```
  READY--
 DO WE SEE ANYTHING ?
  1
 P
  [(WE . PERSON) (SEE . COGACT) (ANYTHING . OBJ)]
  1

  READY--
 WHAT DO WE SEE ?
  1
 P
  [(WE . PERSON) (SEE . COGACT) (WHAT . OBJ)]
  1
 EXIT

WHEN WE SEE ANYTHING , WHAT DO WE SEE ?
 (SIPS- WHEN ANYTHING ,)
 RELAVB OBJ
 COMMA
 (SIPS- COMMA RELAVB)
 DO DO
 (NCS- (, . (COMMA) (ANYTHING . OBJ) (WHEN . RELAVB))
 COM NP RELAVB
 1

P
 [[(WE . PERSON) (SEE . COGACT) (ANYTHING . OBJ)]
  (WHEN . RELAVB)
  [(WE . PERSON) (SEE . COGACT) (WHAT . OBJ)]]
 1
 READY--
HOW DOES THE OBJECT REFLECT LIGHT ?
 (SIPS- HOW DOES REFLECT)
 MANNER DO THRO"
 MANNER DO THROWBACK
 (SIPS- MANNER)
 VJ JAI
 (NCS- (REFLECT . THROWBACK) (OBJECT . OBJ) (DOES . DO) (HOW . MANNER))
 V NP MOD NAV
 1
P
 [(OBJECT . OBJ) TMOD (THE . DET))
  ((REFLECT . THROWBACK) MOD (HOW . MANNER))
  (LIGHT . RADIATION)]
 1
```

Example 3.

Disambiguation: An example deriving from Katz [1967] was chosen to illustrate the system's ability to select correct word senses from a potentially ambiguous situation. The example frame, "The man hit the colorful ball" is varied by substituting "gave," "attended" and "hit" in the verb slot. The relevant dictionary, grammar and SEF entries are presented in Figure 4. Since the dictionary provides two senses each for "colorful," "ball" and "gave," in the worst case the frame using "gave" might provide eight interpretations (as it in fact did without SEF checking). With the use of the relevant SEFs, the system provided the interpretations shown in Example 4. The two interpretations for "...gave a colorful ball" are expected in that SEFs are allowed for "person present object" and "person present event." In the remaining cases of "hit and "attend" only one interpretation was obtained.

Answering Questions: Our approach to answering questions in this system is described briefly in Simmons and Silberman [1967]. A more detailed description of the question-answering system and experiments with it is in preparation [Schwarcz et al., 1968]. Briefly, the system attempts a direct match with the concept structure of each triple resulting from the semantic analysis of a question. Failing to find a direct match, it generalizes each element of a question triple to include all of its equivalences and subclass elements. Thus a question triple with the element "bird" would generalize to include "condors, robins, bluebirds, etc." This approach failing, the system uses more complicated inferences based on combinations of relations into compound and complex relational products.

```
ANALYSIS

  **ANALYSIS MODE**

THE MAN GAVE A COLORFUL BALL.

 2 INTERPRETATIONS

  1
[(MAN TMOD THE) <GAVE - PRESENT>
 ((<BALL - SPHERE> MOD <COLORFUL - MULTICOLORED>) TMOD A)]

NEXT/FINISHED/RULES/SEF'S?  NEXT

  2
[(MAN TMOD THE) <GAVE - PRESENT>
 ((<BALL - DANCE> MOD <COLORFUL - GAY>) TMOD A)]

NEXT/FINISHED/RULES/SEF'S?  FINISHED

THE MAN ATTENDED A COLORFUL BALL.

 1 INTERPRETATION

  1
[(MAN TMOD THE) ATTENDED
 ((<BALL - DANCE> MOD <COLORFUL - GAY>) TMOD A)]

NEXT/FINISHED/RULES/SEF'S?  FINISHED

THE MAN HIT A COLORFUL BALL.

 1 INTERPRETATION

  1
[(MAN TMOD THE) HIT
 ((<BALL - SPHERE> MOD <COLORFUL - MULTICOLORED>) TMOD A)]
```

Example 4.

Disambiguation Example

```
LOOKUP THE MAN GAVE ATTENDED HIT A COLORFUL BALL.
THE: FUNCTION WORD, 1 SENSE:
  1 ART          DET
MAN: FUNCTION WORD, 1 SENSE:
  1 NP          PERSON, ANIMAL, OBJ
GAVE: FUNCTION WORD, 2 SENSES:
 1 V          PRESENT, OFFER, ACT
 2 V          TRANSFER, MOVE, ACT
ATTENDED: FUNCTION WORD, 1 SENSE:
 1 V          GOTO, MOVE, ACT
HIT: FUNCTION WORD, 1 SENSE:
  1 V          CONTACT, ACT
A: FUNCTION WORD, 1 SENSE:
  1 ART          DET
COLORFUL: FUNCTION WORD, 2 SENSES:
  1 ADJ          MULTICOLORED, BRIGHT, QUAL
  2 ADJ          GAY, LIVELY, QUAL
\\\BALL: FUNCTION WORD, 2 SENSES:
  1 NP          SPHERE, OBJ
  2 NP          DANCE, EVENT


PRINT RULES.

  ADJ  NP  (B MOD A)  NP
  ART  NP  (B TMOD A)  NP
  V  NP  (D A B)  PRED
  NP  PRED  (A BB BC)  S

OK

SEFS.

  (PERSON PRESENT EVENT)
  (PERSON PRESENT OBJ)
  (PERSON GOTO EVENT)
  (PERSON CONTACT OBJ)
  (EVENT MOD GAY)
  (OBJ MOD MULTICOLORED)
```

Figure 4. Dictionary, Grammar, and SEF Entries for Disambiguation Example

We believe the approach is very general and approximately equivalent to that taken in the General Problem Solver [Newell *et al.*, 1963]. A top-down generator is used to transform question triples according to possible inference rules and various heuristics are followed to minimize tree search among the resulting vast set of possibilities.

For examples we asked questions of the following sentences:

> The eye is the organ of sight. The retina is the light-sensitive surface of the eye. Cones and rods are the special sensors in the retina. Cones and rods react to light.

These sentences were analyzed with the grammar of Figure 3 and appropriate SEFs to limit the interpretations to one per sentence. Questions were asked and the system answered as in the following examples:

Example 5. What is the eye?

Eye is organ of sight.
Organ of sight.

Example 6. What is the function of the eye?

Sight.

Example 7. What is the surface of the eye?

Surface of eye be light-sensitive surface.
Surface of eye be retina.
Light-sensitive surface of eye be retina.

Example 8. To what is the retina sensitive?

Light-sensitive to light surface.

Example 9. What are the sensors in the eye?

Sensors in eye be sensors.
Sensors in eye be cones.
Sensors in eye be rods.

Example 10. Are there rods in the eye?

Sensors in eye.

Example 11. Does the eye contain cones?

Retina inverse-of-in sensors.
Sensors in eye be sensors.
Sensors in eye be cones.
Sensors in eye be rods.

Example 12. How does the eye react to light?

Cones and rods in retina in eye react to light.
Rods react to light.
Cones and rods react to light.

Example 5 is a result of direct lookup. The correspondence of "function" to "sight" in Example 6 results from the SUP-chain "sight-cogact-function" showing that sight or any other cognitive act is a kind of a function. In this example also, the structure (eye EQUIV (organ ASSOC sight)) implies (eye ASSOC sight) by right-collapsibility* of the "EQUIV" relation. The relation "ASSOC" is defined as symmetric and thus the question transforms to (eye ASSOC function) which is answered by (eye ASSOC sight).

Example 7 is essentially a direct lookup that is successful because of the symmetric property of EQUIV that allows the reversal of the clauses. The answer to Example 8 depends on an additional fact given to the system, "light-sensitive means sensitive to light." With this added information the question which was analyzed to the following structure:

(retina MOD (sensitive TO what))

is directly answered by the structure:

(retina EQUIV (surface MOD (sensitive TO light))) .

---

*The property right-collapsible is defined for R1 as follows:
(X1 R1 (X2 R2 X3)) IMPLIES (X1 R2 X3).

Example 9 and Example 10 require a chain of inference depending on the property transitive attached to (in . contained-in). Thus, "sensors contained-in retina," and "retina contained-in eye" imply "sensors contained-in eye." In Example 11 a similar logic applies with the addition of the information that "contained-in inverse contained."

Example 12 shows one method for treating simple "how" questions. By analyzing the question into the statement "eye reacts to light" the system naturally returns relevant material, which is one main requirement of such questions. It should be noticed, incidentally, that the transitivity of the contained-in relation is used again in this example.

The question-answering system has two important weaknesses. First, we do not yet formally distinguish between the requested operation (i.e., count, list, name, etc.) and the data-identifying portions of the question. This lack partially accounts for the second weakness--a certain degree of vagueness in the generated answers, as can be seen in Examples 8 and 10, where appropriate answers would have been "light" and "yes" respectively. Syntactic and semantic inadequacies in the generation of answers will be discussed in the following section.

Syntactic Generation and Lexical Paraphrase: Our primary emphases in developing a theory of verbal understanding have been to account for the recognition of verbal meanings as communicated by sentences and to demonstrate understanding by the model's ability to answer English questions. Other measures of understanding include the capabilities for syntactic and lexical paraphrase and for the generation of new sentences that are in controllable relations to the data that have been stored as a consequence of understanding meanings from sentences that have been analyzed. It was our hypothesis that the structure for recognition and question answering would prove largely sufficient for generation and paraphrase. In the main this hypothesis was supported and we added the generation grammar and the special machinery for paraphrase to the model in short order. However, it is apparent that generation and paraphrase from deep conceptual structures require more theoretical explanation than we are prepared to deal with in this paper. Particularly required is an outline of correspondences with and contradictions of current generative linguistic theories. At this point, having only scratched the surface in experimenting with the generation area, we will present a brief discussion of our method and save detailed treatment for a later paper.

Generating English phrases or sentences from the conceptual structure is accomplished by the use of transformational phrase structure rules similar to those used in the analysis phase. Since the structure is composed of nested triples, these rules have the form of a three-element left half which is transformed to a structure or a string as a right half. Example rules for generating "the angry pitcher struck the careless batter" are shown below:

```
(NP MOD ADJ) (B A) NP
(NP TMOD ART) (B A) NP
(NP V NP) (A B C) S
```

The notation conventions are identical to those used in recognition rules (see p.18). The generation algorithm, given a complex triple of concepts, first discovers for each concept its mapping onto word-sense and associated syntactic class and associates with each triple of concepts a triple of word-classes. Then, beginning with the most deeply nested triple and working outward, it looks up each triple of syntactic word-classes in the set of generation rules, and if found rewrites that triple by the phrase structure name and applies the transformation to the associated concept triple. This process is iterated until all elements in the nested structure have been accounted for (whether or not they result in the terminal symbol S). If a given triple can be rewritten in more than one way, the algorithm applies all rules, generating several syntactic paraphrases of the same structure. Thus, for the triple (NP MOD ADJ) the two strings "NP that is ADJ" and "ADJ NP" might result. At the end of the process the concepts are transformed into print images.

Lexical paraphrase is accomplished by allowing the free substitution of concepts that are in an equivalence relation. These concepts may map onto words or phrases. Thus in the examples presented below, "eye" is conceptually equivalent to "eyeball," and "organ of sight" is equivalent to "sensor for vision." An equivalence class is established by the statement "X SEQUIV Y" which is semantically analyzed like any other sentence except that SEQUIV is an operator used to construct an equivalence between concepts in the system.

The following examples illustrate both syntactic and lexical paraphrases accomplished by the system:

Example 13. The eye is the organ of sight.

The eye is the sensor for vision.
The eyeball is the sensor for vision.
The eyeball is the organ of sight.

Example 14. The retina is the light-sensitive surface in the eye.

The retina is the surface that is light-sensitive in the eyeball.
The retina is the sensitive to light surface in the eye.
The retina is the light-sensitive surface in the eyeball.

Example 15. Light falls on sensitive retina.

Radiance falls on retina that is sensitive.
Radiance falls on sensitive retina.
Light falls on surface of eye that is retinal.
Light falls on retinal surface of eye.
Radiance falls on retinal surface of cranial orb.

No stylistic controls have so far been established to select either the generation transformation or the lexical item where several choices have been offered, and in expressing answers to questions no method has yet been developed for selecting a "best" answer. Such controls offer an entire field of study such as that currently in progress by Klein.* Our generation procedure is also undeveloped with respect to choice of articles and the various forms of agreement in tense, number, etc. In respect to such syntactic features, the system makes provision for recording them, but we have not yet used them in any of our recognition, question answering or generation experiments.

---

*Personal communication, S. Klein, University of Wisconsin, Computer Sciences Department.

V. Discussion and Conclusions

In the preceding sections we have described a theory and a model of verbal understanding that is based on a formalization of conceptual structures sufficient to represent a wide range of the verbal meanings that are expressed in English sentences. The model includes a linguistic component that is composed of a lexicon, and syntactic and semantic systems which are together sufficient to translate from a wide range of English sentences into the formal conceptual structure. The formal conceptual structure includes inference rules, a limited quantificational capability and a logical structure of relations that are definable by properties for use in inference procedures. These features of the model support a range of question answering and verbal problem solving capabilities.

The model has so far been limited to representing single sentence meanings, although the conceptual structure naturally embeds fragmentary meanings in their most relevant contexts. We do not believe, however, that a theory of sentence meanings is broad enough to encompass the communications mediated by natural languages. Related work in our laboratory by Olney [1967] has investigated anaphoric and discourse analysis to a degree that is sufficient to show us that complete understanding of a sentence can only be modeled in the context of its discourse structure. This line of research has also provided several workable approaches to finding antecedents for pronouns and other kinds of anaphoric structures. An important next step in the development of the model will be to incorporate this line of thought and experimentation and so expand from a model of sentence understanding to one that represents understanding of larger discourse structures.

In Section IV we have briefly shown and discussed examples of the system's capability to produce syntactic and semantic analyses of sentences and questions; to select appropriate word senses according to context; to answer questions; to generate English sentences and to produce meaning-preserving paraphrases. We believe (but have not shown) that only minor modifications are required for the system to deal with a wide range of verbal analogy problems and to accomplish sentence-for-sentence translation. We claim that these results support the theory of verbal understanding outlined in Section II and demonstrate that this theory is adequate as a first approximation to account for how natural language sentences can communicate verbal meanings from one person (or system) to another.

Despite this strong claim for our model, we believe that it is only one example of a family of models that are sufficient to support a theory of verbal understanding. What appears to be common among members of this family is a capability to represent textual information as a structure of unambiguous concepts and well-defined relations. An ideal model would contain only well-defined relations to connect concepts, each of which might in its turn reduce to a structure of perceptual features which themselves were well-defined primitives. Such an ideal may never be attainable, but the closer we approach it, the more satisfactory will be our theories of verbal understanding and the more powerful the language processors that can be constructed.

A number of weaknesses in the model have become apparent as we have struggled with it. Our present treatment of the conceptual structure leaves us too tightly bound to the subject-verb-object order of English sentences and to the ordering of modifying phrases. By shifting to structures such as those outlined on p.13, the model can be simplified and the number of incompletely defined relations can be reduced. The segregation of semantic event forms into a separate neighborhood from the lexicon is also intuitively disturbing and probably economically unsupportable. Temporary weaknesses in the implementation of the model include incomplete developments in the areas of translating from English to logical quantification, the treatment of inflections and agreements, and at least minimal stylistic control of the sentence generation process. These are considered temporary weaknesses because in each case we have designed and are currently implementing improvements.

On the positive side we are very pleased with the model's capabilities for analyzing exceptionally complicated English sentences and obtaining one or more interpretations consistent with the grammar and the semantic system. We were excited and pleased to find that after text and questions have been semantically analyzed and so represented in a formal structure, a question-answering system is essentially identical with a general problem solver. The relative ease with which syntactic and lexical paraphrases can be generated from the deep conceptual structures supports our belief that the Chomskyan generation model is unnecessarily complicated in its treatment of the interface between semantic interpretations and syntactic base structures.

The model's implementation as a LISP 1.5 program leaves much to be desired. It is slow and cumbersome in its operation and sharply limited in storage capability, having in its final version 11,000 words of free space.

Yet only in LISP could we have tried so many variations of our original ideas until we were able to formulate them in terms of consistent workable programs. So on the one hand the system owes its existence to the facility with which complex ideas can be expressed in LISP while on the other, since sequential computers so poorly fit the requirements of large associative networks that LISP is well-suited to handle, the system is core-bound and painfully slow.

"Slow" means concretely that a typical sentence requires 90 seconds of compute time to analyze while an equivalent question requiring no great amount of inference may compute for three to four minutes. When these compute times are translated to wait-times on the time-shared system, analyzing and answering a question may take from fifteen to thirty minutes. Experimenting with such a system is obviously only tolerable to the most devoted believers in the eventual value of computer language processing.

In consequence a JOVIAL version of Protosynthex III, also for the Q-32 time-shared system, has been designed and already partly programmed.* So far the semantic analysis and generation systems are operating. This version has access to eight million words of disc storage. Its computing time for sentence analysis is gratifyingly reduced to tenths of seconds and its wait times on the time-shared system are typically within the turnaround time of 5-15 seconds. It is our current estimation that question answering with relatively short chains of inference will be vastly shortened with respect to the LISP version.

---

*Detailed design and programming by William J. Schoene.

Conclusions: We believe that the present system, Protosynthex III, demonstrates beyond question that sophisticated natural language processing by computers is a realistic goal and one that has been partly achieved here--although so far only on a sentence-by-sentence basis. We believe we have shown that with an appropriate lexicon, syntactic and semantic systems, that a wide range of English sentences can be translated with relative ease into formal structures that support logical operations of deduction and inference. It is further apparent to us that when a question and an answering text have been translated into the formal concept structure, question answering fits into the theorem proving and general problem solving models.

These conclusions mean to us that there is little mystery attached to the problem of language processing by computers--only a great deal of work. Before were were able to build Protosynthex III as a demonstration of a first general-purpose language processor, dozens of language processors and hundreds of man-years of research had to be accomplished throughout this and other countries. Many more years are required to move forward from this point to models that can deal with discourse structures, with large bodies of text and with the subtleties of meaning expressed by metaphor, by stylistic control, etc.

What can be accomplished today is to construct limited systems that deal in a limited manner with limited bodies of text. Such systems, programmed, require syntactic and semantic information in the form of dictionary entries, recognition and generation grammars, semantic event forms, and properties and rules of inference to define relations. All of these materials must be generated by skilled human users of the language and the system. For a single article such as this one, thousands of lexical entries and hundreds of syntactic and semantic rules would have to be produced. In itself, this linguistic effort would require months for the first article and years before any significant subset of a language had been so formally described.

One such limited system is under construction for application to a three-hour computer-aided instruction lesson and the linguistic effort required for computer modeling of the lesson has been begun. The outcomes of this study [Simmons and Silberman, 1967] should teach us much with regard to the eventual practicality and economic feasibility of language processing. In the closely related area of answering questions from data bases such as census reports or airline guides, the CONVERSE system by Kellogg [1967 a,b] has led to the similar conclusion that answering English questions from a data base is an eminently feasible--though possibly expensive--operation requiring significant linguistic effort in defining and formalizing the subset of English to be used.

We believe that this and other papers have demonstrated that natural language processing by computers is rapidly approaching a developmental phase in which the application of significant amounts of time and money can lead to eminently practical results. Significant improvements in automated translation, data base query systems, information and text retrieval, stylistic and content analysis can all be expected in the near future providing support is forthcoming for these efforts. This support will not only be required for the computer programming costs but also in equal or greater measure for the ancillary linguistic effort to formalize appropriate subsets of natural language.

REFERENCES

Abelson, R. P. and Canoll, J. D. Computer simulation of individual belief systems. The American Behavioral Scientist, May 1965, 9, 24-30.

Allport, F. H. Theories of perception and a concept of structure. New York: John Wiley & Sons, 1955.

Bendix, E. Componential analysis of general vocabulary: The semantic structure of a set of verbs in English, Hindi, and Japanese. The Hague: Mouton & Co., 1966.

Bobrow, D. G., Fraser, J. B., and Quillian, M. R. Automated language processing. In Cuadra, C. A. (ed.), Annual Review of Information Science and Technology, Vol. II. New York: Interscience, 1967.

Bohnert, H. G., and Becker, P. O. Automated English-to-logic translation in a simplified model. IBM, Thomas J. Watson Research Center, Yorktown Heights, New York, March 1966. AFOSR 66-1727 (AD-637 227). 117 pp.

Colby, K. M. Computer simulation of change in personal belief systems. Behavioral Sciences, 1967, in press.

Chomsky, N. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press, 1965, 251 pp. (Massachusetts Institute of Technology, Research Lab. of Electronics, Special Technical Report No. 11).

Deese, J. *The structure of associations in language and thought*. Baltimore: The Johns Hopkins Press, 1965.

Deese, J. Some contributions of psycholinguistic studies to content analysis. Paper presented to National Conference on Content Analysis, Philadelphia, November 16-18, 1967.

Fillmore, C. J. A proposal concerning English prepositions. In Dineen, F. P. (ed.), *Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies*, pp. 19-34. Washington, D.C.: Georgetown University Press, 1966.

Fillmore, C. J. The case for case. Prepublication copy, The Ohio State University, Columbus, Ohio, 1967.

Green, C. C., and Raphael, B. Research on intelligent question-answering system. Stanford Research Institute Scientific Report No. 1, Menlo Park, Calif., May 1967.

Katz, J. J. Recent issues in semantic theory. *Foundations of Language*, 1967, 3, 124-194.

Kay, M. The tabular parser: A parsing program for phrase structure and dependency. The RAND Corporation Memorandum RM-4933-PR, Santa Monica, Calif. July 1966.

Kay, M. Experiments with a powerful parser. The RAND Corporation Memorandum RM-[illegible]2-PR, Santa Monica, Calif., October 1967.

Kellogg, C. H. On-line translation of natural language questions into artificial language queries. SDC document SP- 827. April 28, 1967. 47 pp. (a)

Kellogg, C. H. CONVERSE-- a system for the on-line description and retrieval of structured data using natural language. SDC document SP-2635. May 26, 1967. 15 pp. Presented at IFIP/FID Conference on Mechanized Information Storage, Retrieval and Dissemination, Rome, Italy, June 1967. (b)

Kimball, J. P. Predicates definable over transformational derivations by intersection with regular languages. Information and Control, VII (1 and 2), July-August 1967, pp. 177-195.

Kuno, S. Computer analysis of natural languages. Presented at Symposium on Mathematical Aspects of Computer Science, American Mathematical Society, New York, April 5-7, 1966.

Miller, G. A., Galanter, E., and Pribram, K. H. Plans and the structure of behavior. New York: Holt, 1960.

Newell, A. and Simon, H. A. GPS, a program that simulates human thought. In Feigenbaum, E. and Feldman, J. (eds.), Computers and Thought. New York: McGraw Hill, 1963.

Olney, J. C and Londe, D. L. An analysis of English discourse structure, with particular attention to anaphoric relationships. SDC document SP-2769. November 10, 1967. 10 pp.

Quillian, R. Word concepts: a theory and simulation of some basic semantic capabilities. Paper 79, CIT, Pittsburgh, Pa., April 5, 1965.

Quillian, R. *Semantic memory*. Ph.D. Thesis, Carnegie Institute of Technology, Pittsburgh, Pa., February 1966.

Reitman, W. R. *Cognition and thought: an information processing approach*. New York: John Wiley & Sons, 1965.

Salton, G. Automated language processing. In Cuadra, C. A. (ed.) Annual *Review of Information Science and Technology*, Vol. III. Chicago, Ill.: Encyclopaedia Britannica, 1968.

Schwarcz, R. M. Steps toward a model of linguistic performance: A preliminary sketch. The RAND Corporation Memorandum RM-5214-PR, Santa Monica, Calif., January 1967.

Simmons, R. F. Answering English questions by computer: A survey. *Communications of the ACM*, 1965, *8*(1), 53-70.

Simmons, R. F. Automated language processing. In Cuadra, C. A. (ed) *Annual Review of Information Science and Technology*, Vol. I. New York: Interscience, 1966.

Simmons, R. F., Burger, J. F., Long, R. E. An approach toward answering English questions from text. *Proceedings of the 1966 Fall Joint Computer Conference*, 357-363.

Simmons, R. F. and Silberman, H. F. A plan for research toward computer-aided instruction with natural English. SDC document TM-3623. August 21, 1967. 40 pp.

Simmons, R. F. and Burger, J. F. A semantic analyzer for English sentences. SDC document SP-2987. January 10, 1968. 45 pp.

Wilks, Y. Computable semantic derivations. SDC document SP-3017. January 15, 1968. 160 pp.

Woods, W. A., Jr. Semantic interpretation of English questions on a structured data base. Mathematical Linguistics and Automatic Translation, Report NSF-17, The Computation Lab., Harvard University, Cambridge, Mass., 1966, 39 pp.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

ORIGINATING ACTIVITY *(Corporate author)* System Development Corporation Santa Monica, California	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b GROUP

REPORT TITLE

A Computational Model of Verbal Understanding

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

AUTHOR(S) *(First name, middle initial, last name)*

Robert F. Simmons Robert M. Schwarcz
John F. Burger

REPORT DATE 30 April 1968	7a. TOTAL NO OF PAGES 53	7b NO OF REFS 34
CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S) SP-3132	
PROJECT NO	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*	

DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY

ABSTRACT

This paper presents a theory of verbal understanding based on a formal model of conceptual structures that represent verbal meanings expressed in English sentences. Verbal understanding is defined as the capability for disambiguation, paraphrase, question answering, translation, etc., with regard to natural language sentences. The model has been implemented as Protosynthex III in LISP on the Q-32 time-shared system. Experimental results from the system include examples of the analysis of complex sentences, disambiguation of multisense words via sentence context, question answering via logical inference, and meaning-preserving paraphrase generation. The authors conclude that sophisticated natural language processing by computers is a realistic goal that has been partly achieved. The rate of progress toward complete achievement is seen to be proportional to the amount of developmental support available.

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Verbal understanding						
Natural language sentences						
Protosynthex III						
LISP						
Q-32 computer						
Computational linguistics						
Time-shared systems						
Disambiguation						
Question answering						
Analysis of complex sentences						

Unclassified
Security Classification