American Journal of Computational Linguistics *Microfiche 83*

A LEXICON

FOR A COMPUTER QUESTION-ANSWERING SYSTEM

MARTHA W EVENS

Department of Computer Science
Illinois Institute of Technology

RAOUL N SMITH

Department of Linguistics
Northwestern University

This is an expanded version of a paper presented at the Fifteenth Annual Meeting of the Association for Computational Linguistics, March 16, 1977. Accepted for publication September 30, 1978.

SUMMARY OF

A LEXICON FOR A COMPUTER QUESTION-ANSWERING SYSTEM

An integral part of any natural language understanding system, but one which has received very little attention in application, is the lexicon. It is needed during the parsing of the input text for making inferences, and for generating language output or performing some action. This paper discusses the principal questions concerning the lexicon as it relates in particular to a question-answering system and proposes a specific type of lexicon to fulfill the needs of this system.

Rather than make a distinction between dictionary and encyclopedia, we have a single global data base which we call the lexicon. Homographs are differentiated and phrases with fixed meanings are treated as separate entries. All the information in this lexicon is encoded in the form of relations and words or word senses. These form a large network with the words as nodes and the relations as edges. In addition the relations define semantic fields and these are used to treat problems of ambiguity. Relations are used to encode both syntactic and semantic information. Axiom schemes are associated with each relation and these are used for inferencing. The lexical relations then are at the heart (or brain) of the system for representation, retrieval, and inferencing.

For each relation we describe its semantics and the axioms appropriate to it. In the positing of lexical relations our approach has

been influenced by the work of Apresyan, Mel cuk, and Zolkovsky. The lexical relations we have posited are the traditional synonymy and antonymy, taxonomy, part whole, grading and approximately forty others. The whole set, deliberately left open ended, is subdivided into nine subsets which include attribute relations, collocational relations and paradigmatic ones.

Each relation has its own lexical entry giving its properties and telling how to interpret lexical relationships in a first order predicate calculus form. For example, the information for the lexical entry *dog* includes the statement *dog* T *animal*, that is, that a *dog* is a kind of *animal*. The lexical entry for T, the taxonomic relation, in its turn includes information which allows the statement to be interpreted as

$$Holds(Ncom(\text{dog},X)) \rightarrow Holds(Ncom(\text{animal},X)).$$

The inventory of relations is expandable simply by adding lexical entries for new relations. In addition having both the lexical entries and the relations in the entries expressed in the same notational form as that of input sentences, namely in a first order predicate calculus notation, allows for a consistent, coherent, and easily modifiable system for analysis, inference, and synthesis.

TABLE OF CONTENTS

1. INTRODUCTION

The lexicon presented here is being developed as an integral part of a computer question-answering system which answers multiple-choice questions about simple children's stories. It thus must make information readily available for the parsing process, for building an internal model of the story being read, and for making inferences. Knowledge about words and knowledge about the world must both be stored in a compact but immediately accessible form.

Many decisions must be made, therefore, about the design of the lexicon. The first problem is to decide on an organizing structure. Should lexical and "encyclopedic" information be stored separately or together? Which items will have separate lexical entries? Which will be included in other entries? What about homonymy and polysemy? What connecting links between words and word senses will be recorded and how?

The next problem is to determine a characterization of word meanings. This leads to some deep theoretical questions What kind of lexical semantic representations are appropriate? What is the structure of these representations? What are the semantic primes, the elements of that structure? The design of the lexical entry is thus subject to theoretical biases as well as the practical constraints of space, retrieval efficiency, and effective support of inference-making.

The decision to use lexical relations as fundamental elements of the structure of the lexicon has strongly influenced our design. Relations are used to encode both semantic and syntactic information. Axiom schemes essential to inferencing are associated with each relation. Relational information makes up a significant part of the lexical entry.

Lexical relations offer significant advantages. They allow us to generalize familiar inference patterns into axiom schemes. They can encapsulate the defining formulae of the commercial dictionary. They have an intuitive appeal which we believe reflects a certain measure of psychological reality. On a practical level they allow us to express both syntactic and semantic information in a form that is compact and easy to retrieve. They can be used in many ways. For example, the following paragraph from a test administered to first and second graders by a local school system says:

(P1) Ted has a puppy. His name is Happy.
Ted and Happy like to play.

(Q1) The pet is a: dog boy toy

In order to answer this question we need to know what *pet* means. In our lexicon the lexical entry for *pet* contains a simple definition. a pet is an animal that is owned by a human. In order to answer this question we also need to know that a puppy is a young dog. This information in predicate calculus form would be part of the lexical entry for puppy. We would, of course, need axioms of the same form as well for the entries for kitten, lamb, etc. Instead of such a representation we express this information by using a lexical relation, CHILD. The lexical entry for *puppy* contains CHILD *dog*. Similarly, the lexical entry for *kitten* contains CHILD *cat*; while the lexical entry for CHILD contains the axiom scheme from which the relevant axioms are formed when needed.

We treat verbs in a similar way. Corresponding to each case relation there is a lexical relation which points to typical fillers of that case slot. The lexical entry for $bake_2$ includes TLOC *kitchen*. It also includes T *make* where T is the well-known taxonomy relation, so that if the story says that "Mother baked a cake " we can infer that she made one and CAUSE $bake_1$ so that we can deduce that the cake has baked The selection restrictions that help us tell instances of $bake_1$ and $bake_2$ apart can also be expressed compactly using the T relation. We also need to make deductions from main verbs in predicate complement constructions, deductions such as the speaker's view of the truth of the proposition stated in the complement as derived from the factivity of the verb. In order to answer several questions from the test cited above the reader must infer that everything that Mother says is true. Lexical entries for main verbs that take predicate complements contain pointers to the implication class. These relations can then be expanded to give the proper axioms.

The lexicon includes separate entries for each derived form unless the root can be identified by a simple suffix-chopping routine. Lexical relations are useful here too in saving space. The lexical entry for *man* contains PLURAL *men*. The lexical entry for *went* consists of PAST *go*. The lexical entry for *death* consists of NOMV *die*. There are, as well, lexical entries for some multiple word expressions such as *birthday party*, *ball game*, *piggy bank*, and *thank you*.

As to the form of presentation here, the next section presents some of the practical problems and theoretical convictions which determined our most critical design decisions. Then, after a brief description of some earlier developments in the theory of lexical relations, we explain the system of relations which structure our lexicon, discussing each group of relations in turn Finally we describe the actual form of our lexical entries.

2. DESIGN DECISIONS

The lexicon in this system must make information readily available for parsing, for building the story model, and for making inferences during question answering. Both knowledge about words and knowledge about the world must be stored in a compact but immediately accessible form. Therefore, many decisions must be made concerning the design of the lexicon. The problems involved include the organization of lexical and encyclopedic information, the choice of a lexical model and the determination of appropriate semantic primes, the representation of selection preferences, the recognition and storage of homographs, and the criteria for establishing separate entries for idioms and other fixed phrases. This paper attempts to develop some consistent solutions to these problems, solutions which determine the design decisions for the lexicon in this question-answering system.

a. *The dictionary and the encyclopedia - one data-base or two?*

Any question-answering system must use lexical information in at least two ways, in parsing and in making inferences. The first critical decision that must be made is whether two separate data-bases are needed to support these separate functions or whether a single unified global data-base is better. Traditionally human beings have used two separate stores of information, the dictionary and the encyclopedia. Some linguistic and computational models of language have also been based on the assumption that information about words should be stored in two separate collections.

In Chomsky's *Aspects* model (1965) there are two separate storage places for lexical information, one in the base component and another in the semantic component. Marx (1972) stored syntactic information in a 'dictionary" and semantic information in an 'encyclopedia" Winograd (1971) has two separate word lists, one used by the parser and one by the semantic routines, even though the parsing and the semantic routines are very closely interwoven in his BLOCKS system

Before deciding on whether to carry on this tradition one must ask whether there is really a clearcut distinction between these two kinds of lexical information Is there a simple algorithm for deciding which data should go where?

Bierman, Bierwisch, Kiefer, and the Semantic Function of the Lexicon

Both the dictionary and the encyclopedia are ways of recording information stored in human memory But human memory is probably not organized in the usual graphic form of an alphabetic word list, therefore alternative memory structures should be examined One such alternative has been presented by Bierman (1964) In his system lexical-semantic fields are primary, they define the basic organization of semantic information The function of the lexicon, if it has one in the semantic domain, is to index these fields, to store pointers to the location of a word in the various fields containing it An appropriate image for such a system is a very large single page dictionary with language specific nodes connected by semantic relations (See also Werner 1969)

Can the dictionary and the encyclopedia be distinguished in this context? Bierwisch and Kiefer (1970) assume that both kinds of information are contained in the same lexical entry The distinction between lexical and encyclopedic knowledge corresponds then to the difference between the core and the periphery of a lexical entry, where

> The *core* of a lexical reading comprises all and only those semantic specifications that determine, roughly speaking, its place within the system of dictionary entries, i.e. delimit it from other (non-synonymous) entries. The *periphery* consists of those semantic specifications which could be removed from its reading without changing its relation to other lexical readings within the same grammar (ibid: 69-70)

Unfortunately they do not specify whether the lexical-semantic relations which form the structure of the fields are part of the core or the periphery.

The major difficulty with this criterion is its instability. As new entries are added to the system, information sufficient to distinguish one entry from another may have to be shifted from the periphery to the core --and thus from the encyclopedia to the lexicon. For instance, suppose a new entry, "leopard--a large, wild cat" is to be added. The entire lexicon must be searched for entries which mention large wild cats. If one is found, say "lion--a large wild cat", then enough information must be added to both definitions to differentiate *leopard* and *lion* from each other.

Soviet Lexicography and the Lexical Universe.

Apresyan, Žolkovsky, and Mel'čuk run into the same difficulty of distinguishing dictionary and encyclopedic information in attempting to define the lexical universe of a word C_0.

> The main themes dealt with under the heading 'lexical universe' are: 1) the types of C_0; 2) the main parts or phases of C_0; 3) typical situations occurring before or after C_0 etc. Thus, the section lexical universe for the word *skis* consists of a list of the types of skis (*racing, mountain, jumping, hunting*), their main parts (*skis* proper and *bindings*), the main objects and actions necessary for the correct use (exploitation) of skis (*sticks, grease, to wax*), the main types of activities connected with skis (a ski-trip, a ski-race...) and so on. Even these scanty examples make it clear that the information about the lexical universe is, at least partially, of an encyclopaedic nature. We say "partially" because genuine encyclopaedic

> information about skis (their history, the way they are manufactured, etc.) is not supplied here: the sections contain only such words and phrases as are necessary for talking on the topic, and nothing else. (1970:19)

The problem here is that "what is needed for talking about the topic" depends very much on who is going to do the talking. The definition of *ski* in *Webster's New International* (2nd Edition) begins:

> One of a pair of narrow strips of wood, metal, or plastic, usually in combination, bound one on each foot and used for gliding over a snow-covered surface.

Apresyan, Žolkovsky, and Mel cuk do not provide for three of the items mentioned here: what skis are made of (wood, plastic, or metal), what shape they come in (long and narrow) and where they belong spatially (on the feet). Yet these items could be essential in understanding inferences in a story.

> It was snowing. Jim took out his skis.
> He waxed the wooden strips....

You could need this information in answering questions.

> Jim skied down the mountain....
> What was he wearing on his feet:
> slippers skis skates?

Although in English or Russian it is possible to refer to skis without knowing that they are long and narrow, it is not possible in Navajo where physical shapes determine verb forms. While the entry in Webster's goes on at length beyond the sentence given above, it does not include all the items which Apresyan, Žolkovsky, and Mel'čuk mention. This, however, is not surprising; the boundaries of the lexical universe are not well defined.

Difficulties in updating a system with separate dictionary and encyclopedia,

This lack of definition causes tremendous problems in a dynamic system. A "real" dictionary or encyclopedia, the one in a person's brain, is constantly changing. Information is added, corrected, and perhaps lost. A truly interesting memory model must be dynamic. The problems of updating this information are not easy to solve, the problem of distinguishing between dictionary and encyclopedic information in the updating process seems insuperable.

Recognizing definitions phrased in ordinary English is already difficult (Bierwisch and Kiefer 1970, Lawler 1972). Determining the reliability of such information is also a problem and the dichotomy of dictionary and encyclopedia increases this difficulty. Unfortunately information does not come neatly packaged and marked "for the dictionary" or "for the encyclopedia". And addition of information to one part of the entry may necessitate updating other parts of the entry. For example, if we learn that *record* is a verb as well as a noun we need to add morphological information, describe the relations between *record* and *write*, and we should probably describe recording materials. Mention must be made that *record* is a factive, i.e. if someone records that something happened, one can assume that from the standpoint of the speaker the something really did happen. Which of this information is dictionary information and which is encyclopedic? And once this decision is made, information added to that entry may require additions to other entries In the *record* example, the entries for *erase* and *write* would have to be updated. Also, a decision must be made on whether a new entry is needed and whether homography or polysemy exists for this new entry.

The work of Kiparsky and Kiparsky (1970), Lakoff (1971) and McCawley (1968) has shown that syntax and semantics cannot be separated into such neat compartments. But if syntax and semantics are interwoven then does it make sense to put syntactic information in one box and semantic information in another? The answer to this question given at least by generative semantics calls into question the traditional distinction between the dictionary and the encyclopedia.

We accept the generative semanticist arguments that syntax and semantics cannot be separated and thus do not separate syntactic and semantic information. Furthermore, as shown above there seem to be no practical criteria for distinguishing dictionary information from encyclopedic information. Thus our system has one single global data base. For brevity and since it is a kind of collection of words, it will be called "the lexicon".

b. *Lexical Models - Componential Feature Analysis vs. Relational Networks.*

A second critically important decision involves the choice of an appropriate lexical model, the determination of what semantic primes to use and how they should be combined in lexical semantic structures. Two important competing models are provided by componential feature analysis and by relational networks. In a componential analysis model the primes are semantic features and words are defined by bundles of features. This is a natural extension of the distinctive feature approach to phoneme description which has been used to explain many phonological phenomena. Certain practical problems arise. The number of words in any language is far larger than the number of phonemes. The number of distinctive features which serve to discriminate them must be larger too. The word-semantic feature matrix for a given language would be vastly larger than the phoneme-phonetic feature matrix. In addition, this matrix would be extremely sparse. Also, it is not clear whether all the entries in this matrix could be +/- as in a phoneme matrix. Are semantic features either definitely absent or definitely present or are some features present by degrees? The size of the componential analysis matrix would immediately introduce difficulties in a computerized model. Fortunately, both numerical analysis and document retrieval offer experience in handling immense matrices by machine. When a rix is extremely sparse it turns out to be sensible to store a list of entries with row and column numbers Here it would mean storing a list of features for each word. This. in fact, is close to Katz's proposal (1966).

In a relational network model, however, the primes are relations and words or word senses. Relations connect words together in a network in which the words are nodes and the relations are edges. In fact, words,are defined in terms of their relationships to other words.

These models differ radically in their approach to the critical lexical task of finding related words. In the componential analysis model related words share related features. Presumably, the more features two words share the more closely related they are. Thus, some kind of cluster analysis must be used to identify related words. In the relational network model the lexicon is formed from relationships between words. Thus related words are immediately available

Both models, componential and relational, require a search for semantic primes The componential analysis model requires the discovery of possibly thousands of semantic features. For a relational network model an inventory of lexical relations and their properties must be developed. This is apparently a significantly simpler task than the discovery of semantic features, for the number of relevant relations is probably quite small. Žolkovsky and Mel'čuk (1970) list about fifty in their paper.

Related to both of these models is the notion of semantic fields. Intuitively, semantic fields are collections of related words used to talk about a particular subject. Semantic fields seem to offer some help in coping with the problems of ambiguity and context. Many utterances, taken out of context, are ambiguous. But remarkably, people almost never perceive this ambiguity. They immediately choose the correct word sense and ignore the others. Apparently the topic of conversation determines a semantic field and the word sense chosen is the one which lies in this field. The semantic field somehow defines the verbal context. (Or as Fillmore 1977:59 phrases it "meanings are relativized to scenes".)

The componential analysis model makes it possible to define distinct semantic fields, but getting from one word in the field to the others may take a significant amount of processing time. Every set of semantic markers can be used to define a semantic field; the field consists of all the words with definitions containing the markers. The smaller the number of markers the larger the field obtained. It is possible to decide immediately whether a given word is in the field or not, just by checking its list of markers.

In the relational network model related words are easy to find, but the boundaries of semantic fields are extremely fuzzy and indistinct. A semantic field can be defined by starting at a particular node and going a given number of steps in any direction. The semantic fields obtained this way, however, have very arbitrary boundaries and overlap considerably.

Certain basic philosophical-psychological assumptions may create a strong bias for one of these models over the other. Someone who believes that semantic features exist as Platonic ideals or who accepts them as psychological realities may easily find componential analysis a most natural kind of description and regard the necessary search for features or sememes as highly relevant. Someone who feels that "There is no thought without words" would be much more likely to prefer a relational network description. A lexicon *is*, in an important sense, a memory model. Intuition about our own internal memory models must have a strong influence on the lexicon we shape.

We have chosen a relational network model for both intuitive and practical reasons. We find lexical-semantic relations theoretically interesting (see Evens et al. ms). Useful inventories of these relations are available, in a later section we describe some of these sources. As will be shown they provide a convenient way of storing axiom schemes for deductive inference.

As lexical semantic structures we use the same first-order predicate calculus notation in which semantic representations are written in the question-answering system - meanings of words and meanings of sentences must have the same underlying form As McCawley (1970) has argued "dentist" and "doctor who treats teeth" must contain the same units of meaning tied together in the same way

c *Selection Preferences*

A third important problem to be faced in constructing a lexicon which is to support a parser is the problem of selection restrictions.

Chomsky (1965) developed the theory of selection restrictions in order to block the generation of nonsense sentences in the syntactic component of his model. The lexical entry for *frighten*, for example, contains the information that it requires as object a noun with the featur [+animate], while *drink* requires an animate subject. If these conditions are not met, generation is blocked. Selectional restrictions seem much too restrictive. Trailer trucks drink diesel fuel and the earth drinks in the rain. In describing dreams we can invent perfectly appropriate sentences in which inanimate objects by the dozens get up, run around, and drink until frightened back to place. Still it is true that sentences like these are somehow more surprising than sentences in which cows drink from a brook and are frightened by lighthing. We need some method of recording the ordinary, everyday ways in which words combine without excluding the unusual, the poetic, the metaphoric uses. We will call them *selection preferences* instead of selection restrictions.

Some truly semantic means of identifying semantic anomalies are needed. Raphael mentions this question rather casually, almost as an aside in the SIR paper. He draws taxonomic trees, one for the nouns and one for the verbs, from the vocabulary of a first grade reader. Then he makes statements like this

> 1. Any noun below node 1 is a suitable subject for any verb below node 1'.
>
> 2. Only nouns below nodes 3 or 4 may be subjects for verbs below node 3'. (1968, p. 51)

He makes it clear that he is indeed trying to solve the selectional problem

> The complete model composed of tree structures and statements about their possible connections, is a representation for the class of all possible events. In other words, it represents the computer's knowledge of the world. We now have a mechanism for testing the 'coherence' or 'meaningfulness' of new samples of text. (1968, p. 51)

Werner (1972) has suggested a method for handling the selectional problem which uses noun taxonomies in very much the same way that Raphael does. His proposal includes an elegant way of storing selectional information within his memory model. In his network model, noun phrase arguments are connected to the verb by prepositions. The node representing the lexical entry for the verb has arcs connecting it to compound nodes, one for each preposition which can be used with the verb. The object of each preposition is a node in the noun taxonomy. This noun or any noun below it in the taxonomy may serve as an argument for the verb. Here is an oversimplified example of a network for *sell*.

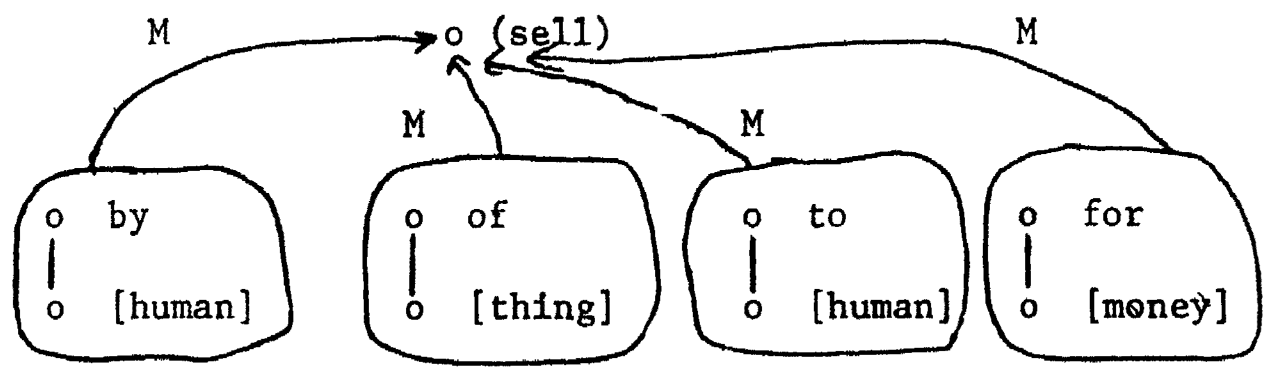


Figure 1: Werner's Answer to the Selection Problem

This network says that *sell* takes a human subject, a thing as object, the preposition *to* followed by a human, the preposition *for* followed by money. The square brackets around [human] indicate that this is just a pointer to the top noun in the taxonomy for human beings. Any node in this taxonomy below the node marked human, whether it is *Sam* or *a Navajo* or *Mother*, can be used as a subject for *sell*. He does not use the verb taxonomy as Raphael does. Each verb has its own set of selection indicators.

In his discussion of the goals of a semantic theory Winograd describes semantic markers and selection restrictions, quotes Katz and Fodor (1964) and indicates that he intends to embody this theory in his system. But in fact semantic markers in the BLOCKS program are derived from a marker tree (Winograd, 1971, Figure 59) which is organized taxonomically. In the implementation process Winograd seems to have moved from a strict Katz-Chomsky position to a position somewhat closer to Raphael and Werner.

The Raphael and Werner proposals are the guiding principles here, adapted to accommodate case-defined arguments. The lexical entry for *move*$_1$, the intransitive *move*, must tell us about selection as well as how to relate subject, object, and prepositional phrases to cases. The information is organized this way:

		grammatical function	*case frame*	*selection information*
	1.	subject	experiencer	thing
move$_1$ #	2.	*from*	source	thing, place
	3,	*to, into, onto*	goal	thing, place

The numbers 1, 2, 3, indicate argument positions for the predicate calculus representation. The next column lists the grammatical function. Next come case indications. Last comes the selection information, the top node in the relevant part of the taxonomy. For *move*$_1$ the subject is an experiencer. The source is usually marked by the preposition *from*. The goal is usually marked by a preposition like *to, into,* or *onto*. The selection information in column four is rather dull, since any argument can be a physical argument or *thing*, the source and goal can both be places. There is a rule that any physical goal can be replaced by a class of adverbs containing *back* and *there*, so these alternatives do not have to be listed.

An attempt is being made to use the verb taxonomy as Raphael suggested. In this lexicon *go* is marked as taxonomically related to *move*. The entry for *go* does not contain the information labelled # above. Instead, when this information is needed, the look-up routine climbs the taxonomic tree in the lexicon until it finds a verb which has this information and copies it from that entry. Thus it gets case-argument and selectional information for *go* from the entry for *move*. It is not clear yet whether this will really work with a sizable vocabulary.

This selectional information is treated as selection preference and not selection restriction. Each candidate word sense for a verb is checked for selectional preference. If no arrangement of the available noun phrase arguments is consistent with these preferences another word sense is examined But if all word senses have been rejected on the basis of selectional information, the sentence is not rejected Instead we look again at the candidate word senses and count for each one the number of steps up the taxonomic tree we have to make to resolve the conflict. The word-sense which requires the fewest steps is chosen. The hope is that the system will be able to "understand" simple metaphors this way. It would be interesting to try to create metaphors by picking noun phrase arguments close to but not under the nodes indicated by the selection information.

d. *The Homonymy-Polysemy Problem - Criteria for Separate Entries.*

Words with the same physical shape but different meanings constantly cause trouble in natural language processing. In designing a lexicon we must decide whether or not to create a separate entry for each variation in meaning and type of use. Quillian is particularly interested in words with multiple meanings and he experimented with several in his memory model. In Quillian (1968) the word *plant* is treated as a three-way homonym with three separate type nodes, each with a separate definition-plane:

PLANT	Living structure which is not an animal, frequently with leaves, getting its food from air, water, earth.
PLANT2	Apparatus used for any process in industry.
PLANT3	Put (seed, plant, etc.) in earth for growth.

The type node for the first forms a disjunctive set with token nodes pointing to the other two

The word *food* has a single definition with alternative formulations:

> That which living being has to take in to keep it living and for growth. Things forming meals, especially other than drink.

A polysemous word like this has a single type node and a single definition-plane, but the two alternative definitions are combined with an OR link.

Apresyan, Mel'čuk and Žolkovsky attack the homonymy-polysemy problem with vigor. Graphically coincident words are considered homonyms, given distinctive superscripts and listed as separate entries, if their definitions "have no common part" (Apresyan, Žolkovsky and Mel'čuk 1970:3) They do not define "a common part," but they do give an example. KOCA (scythe), KOCA^{2} (braid of hair), KOCA^{3} (spit). If two definitions have a single common part, the word is classified as polysemantic with a single entry divided into separate parts. They distinguish two types of polysemy. In one case the difference between two words is regular. The relation of a verb to its typical object is such a regular meaning change, e.g. record(v) - record(n), fish(v) - fish(n), and aid(v) - aid(n). These regular variations in meaning are numbered with Arabic numerals, while irregular variations are numbered with Roman numerals. Thus part 3 of the lexical entry for *bow*, the definition, might have the form:

bow I. 1. To bend the head in assent or reverence. (vi)
2. To submit or yield. (vi)
3. To cause to bend. (vt)
4. An inclination of the head. (n)
5. A bent implement used to propel an arrow or play a stringed instrument. (n)

II. 1. The forward part of a boat. (n)
2. One who rows in the bow of a boat. (n)

There seems to be some redundancy between definition-elements and the lexical functions. Shouldn't regular variations in meaning be captured by regular lexical functions? If so, then the distinction Apresyan, Žolkovsky and Mel'čuk make between regular and irregular meaning variations will be apparent from the form and need not be indicated by different notation, such as Arabic and Roman numerals.

For convenience in lexical lookup we have a single physical entry for each graphical form. Each word sense whether irregular or regular is numbered separately with Arabic numerals. Thus the adjective is *cool*$_1$, *cool*$_2$ is the verb to become *cool*$_1$, and *cool*$_3$ is the verb meaning to cause to become *cool*$_1$. Separate information about lexical relations, etc. is stored for each subentry.

e. *Idioms*

Idioms present a serious problem to the designer of an English lexicon Some criteria must be established for deciding which idioms deserve separate lexical entries and how multi-word phrases should be stored.

When does an idiom deserve to be treated as a separate lexical unit? Apresyan, Žolkovsky, and Mel'čuk (1970) and Kiparsky (1975) represent opposite poles of opinion here. In the explanatory-combinatory dictionary (ECD) of the Soviets word combinations which have a definition of their own or "a peculiar combinability pattern" have separate entries. Kiparsky (1975) considers an idiom as a separate lexical unit only if it involves syntactic patterns which are no longer productive. Thus "house beautiful" and "come hell or high water" are treated as units, but "make headway" is not. Instead *headway* is defined as "progress" and marked as appearing after *make* and *lose*. Kiparsky's proposal places a greater burden on the recognition program which would have to be able to retrieve and put together the pieces of the idiom using his lexicon The system described here follows Apresyan, Žolkovsky and Mel'čuk, and treats fixed phrases as units. In particular, all noun-noun combinations like *piggy bank* and *birthday cake* are separately defined, although this is certainly a productive part of English.

Judith Levi (1974, 1975) has proposed a theoretically elegant and intuitively attractive method of generating these forms. According to

Levi the underlying structure for "birthday boy" is "boy-have-birthday" and the underlying structure for "birthday cake" is "cake-for-birthday." Then under certain conditions *have*, *for*, etc. can be deleted to give us the noun adjunct expressions. Given these rules, she argues, it is not necessary to treat these expressions as separate lexical items. While her rules seem sufficient to allow us to synthesize these compounds correctly, difficulties arise when we try to use them for analysis. The question-answering system needs to be able to infer from "birthday boy" that the boy in question is having a birthday, but to avoid inferring from "birthday cake" that the cake is having a birthday. For correct recognition we need to be able to recover the unique underlying structure if one exists. (For a similar criticism see Downing 1977:814-15.) Levi's theory accounts for the generation of new noun-noun compounds. However, in order to account for the recognition process we need lexical entries for familiar fixed compounds and her theory to analyze new compounds. We have used Levi's structure as a basis for our representation of compound nouns.

Noun-noun compounds have separate entries. A *birthday cake* is treated as "a cake for a birthday." A *ball game* is represented as "a game that has a ball". A *piggy bank* is defined as "a bank that is a pig."

The system is told that Jim has a piggy bank and asked what the bank looks like. It could be argued that anyone with sufficient cultural knowledge ought to be able to answer this even if all the banks in his past were shaped like bee-hives, but we need a place to write down this cultural encyclopedic knowledge and a lexical entry for *piggy bank* seems like a good place to put it.

Becker in his work on "The Phrasal Lexicon" (1975) has produced evidence on the Soviet side of this argument. His data suggest that fixed phrases comprise approximately half of our spoken output and have an independent lexical existence. He includes in his lexicon euphemisms ("the oldest profession"), phrasal constraints ("by {pure / sheer} coincidence"), deictic locutions ("for that matter"), sentence builders ("(person A) gave (person B) a long song and dance about (a topic)"), situational utterances ("How can I ever repay you?"), and verbatim texts (proverbs, song titles, etc.). He claims that

> we speak mostly by stitching together swatches of text that we have heard before; productive processes have the secondary role of adapting the old phrases to the new situation....most utterances are produced in stereotyped social situations, where the communicative and ritualistic functions of language demand not novelty, but rather an appropriate combination of formulas, cliches, idioms, allusions, slogans... (1975:60)

He has collected 25,000 phrases for the phrasal lexicon.

Catherine Flournoy (1975) has found several hundred fixed phrases in a computer study of Father Coughlin's speeches. This is not a new idea to students of oral epic poetry. Homer constantly used fixed phrases to fit syntactic and metrical slots. Dawn is always "rosy-fingered"; Hector is constantly "tall Hector of the shining helm."

There is a serious space-time tradeoff here between parsing time and lexical storage space. It is probably true that people possess and constantly use a phrasal lexicon. Whether we should use storage space for items which we can parse/produce without ambiguity is another question.

Currently we provide separate entries for any phrase that we cannot parse and interpret correctly from the entries for individual words. Brief entries for these phrases seem absolutely necessary for any practical recognition scheme. These entries also seem to be the appropriate place for indexing pointers to the cultural information necessary for making inferences and answering questions about birthday cakes and birthday parties. There are theoretical arguments for such entries as well. We believe, as Becker does, in the phrasal lexicon, although we do not include entries for any phrases that can be parsed and interpreted correctly without a separate entry. Any complete system for language processing must also, of course, contain rules like Levi's to provide an ability to process novel forms.

f. *Preliminary Design Decisions for the Lexicon.*

The goal of this project is a lexicon sufficient for parsing, forming semantic representations, and making inferences, compact but still allowing rapid lexical lookup.

The lexicon is a global data-base for the question-answering system, a combination lexicon-encyclopedia. Syntactic and semantic information are combined in the same lexical entries. Lexical semantic representations are written in the same form as the semantic representations for sentences, in a many-sorted, first order predicate calculus. Homographs which vary in meaning or use are differentiated by Arabic numeral subscripts. Separate entries are included for phrases with fixed meaning.

The lexicon is organized in terms of lexical relations. Semantic fields defined by relations are used to handle problems of ambiguity and context.

The relations are used to express and retrieve many different kinds of information, from past participles to selection preferences to proper habitats for lions. Thus the system of lexical relations is crucial to representation, retrieval, and inference.

3. SOME THEORIES OF LEXICAL RELATIONS

While developing our lexical relations we examined a variety of relational theories in anthropology and linguistics and even collected folk definitions of our own (Evens 1975). We have been particularly influenced by the anthropological fieldwork of Casagrande and Hale (1967) by the memory models of Raphael (1968) and Werner (1974), and most of all by the ECD of Apresyan, Mel cuk, and Žolkovsky (1970). But we looked at each of these relational theories from the peculiar point of view of computer question-answering and the particular lexical environment of children's stories, adding and discarding relations to fit the problem.

Casagrande and Hale - Lexical Relations in Folk Definitions.

Casagrande and Hale (1967) collected 800 Papago folk-definitions and sorted them into groups on the basis of semantic and grammatical similarities. They produced the following list of thirteen lexical relations. (Table 1)

Table 1. The Relations of Casagrande and Hale

Relation	Word	English Gloss of Papago Definition
1. Attributive	*burrowing owl*	but they are small; and they act like mice; they live in holes.
2. Contingency	*to get angry*	When we do not like something we get angry.
3. Function	*tongue*	with which we speak

4	Spatial	*bucket*	in which we get water
5.	Operational	*bread*	which we eat
6.	Comparison	*wolf*	they are rather like coyotes, but they are big
7.	Exemplification	*sweet*	as sugar
8.	Class Inclusion	*crane*	a bird
9.	Synonymy	*amusing*	funny
10.	Antonymy	*low*	not high
11.	Provenience	*milk*	we get it from a cow
12	Grading	*Monday*	the one following Sunday
13.	Circularity	*near*	when something is sitting nearby we say *near*

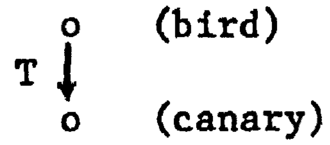
Casagrande and Hale make no claim that they have found all possible lexical relations. These definitions were collected as part of a study of dialect variation in Papago and Pima, The words to be defined were chosen because they might exhibit dialect differences and not to elicit all possible defining formulae. They suggest for intuitive reasons adding the part-whole relation to their list although they did not identify it in their data, They also provide an interesting discussion of word association data in which they give stimulus-response pairs from the *Minnesota Norms* of Russell and Jenkins (1954) exemplifying each of their lexical semantic relations (except for circularity). They cite some word association pairs which do not have exact analogues in the Papago definitions. These are "coordinate" pairs like "needle-thread" or "bread-butter", "clang" responses like "table-stable", or sequential responses "wish-bone" and "whistle-stop". They remark about the bread-butter pair that the relationship involved between "bread" and "butter" is similar to that discussed for contingency, except that in the Papago sample, the contingency relationship is not used if both X and Y are nominal concepts. *Webster's New Collegiate Dictionary* does not mention *butter* in the *bread* entry but it has a separate entry: "bread and butter. Bread spread with butter; hence, Colloq. livelihood ...." (p. 103) It does mention *thread* in the *needle* entry and *needle* in the *thread* entry. This kind of association belongs in every lexicon.

Werner's Lexical Relations.

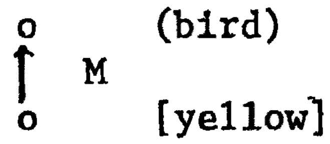
There are two ways to go from the study of folk definitions. One way is to find or invent lexical relations to fit all the folk definitions one can collect in a given language, and then look for more in the formulas of published dictionaries of that language. The other is to abstract a minimal set of language-universal lexical-semantic relations and then attempt to express other proposed lexical relations in terms of the minimal set. Werner has made substantial steps in this second direction (Werner and Topper 1976).

Werner's basic semantic relations are the taxonomic relation (T), the modification or attribution relation (M) and the temporal sequencing relation, queuing (Q). These he calls "the basic cement of the organization of cultural knowledge and memory." (1974:17).

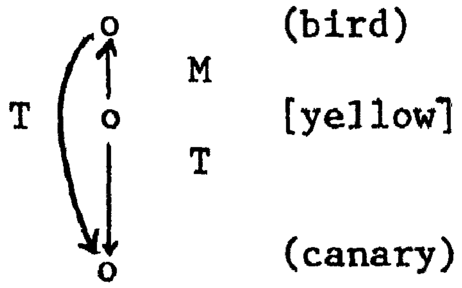
The relation of taxonomy (T), the one expressed in English by "a canary is a (kind of) bird , is written (bird) T (canary) and is represented in Werner's diagrams by a directed arc labelled T.



The relation of modification or attribution (M), the one expressed in English by "the yellow bird" or "the bird is yellow", is represented by a directed arc labelled M.

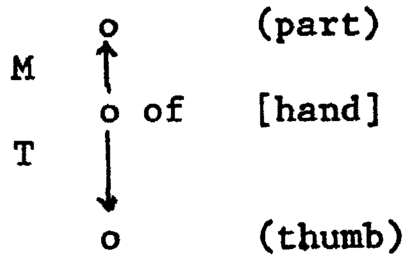


These last two diagrams can be combined to express the idea that a canary is a yellow bird.



The queuing relation Q represents the idea of order or sequence. For example, (Monday) Q (Tuesday). This relation is fundamental in the representation of plans in Werner's memory model. "*Knowing how*... requires the retention of temporal order. there are things to be done first, second, and so on and usually nonsense results if the order is changed (One can't drink the beer before the bottle cap is removed)." (ibid, p. 11)

Relations like 'consists of,' 'part of,' 'cause of,' 'like' are handled as complex relations and composed from the primitive relations M and T using the logical operators *not* ($\sim$), *and* ($\wedge$) *or* ($\vee$) and particular lexical items. For the 'part of' relation he gives the example "the thumb is a part of the hand" (ibid, pp. 50, 51)



This diagram essentially says "the thumb is a (kind of) hand-part."

This is an extremely elegant and general theory. Werner's claim of linguistic universality seems well-founded. His model is in many ways intuitively appealing although we are not convinced that our basic lexical relations and our basic semantic relations are the same.

Our decision to try to design a lexicon with a larger set of lexical relations is really an engineering decision, based on two probably temporary practical difficulties.

(i) We do not know how to prove theorems in Werner's model.

(ii) We believe that a variety of language specific lexical relations can produce a more compact lexicon with more efficient search routines.

Raphael's SIR model

Raphael's Semantic Information Retrieval program (1968) combined a semantic net representation with a relational calculus which makes inferences in this net. SIR inputs simple English sentences, translates them into node-relation-node form, uses a relational calculus to prove theorems, asks for more information, if needed, and answers questions using those inferences. The relations which Raphael used are:

$x \subset y$ (An x is a y, e.g. A boy is a person.)

$x \in y$ (x is a y, e.g. John is a person.)

equiv[x;y] (x and y are two names for same thing.)

owng[x;y] (Every y owns an x.)

own [x;y] (y owns an x.)

partg [x;y] (Some x is part of every y.)

part [x;y] (An x is part of y.)

right [x;y] (x is to the right of y.)

jright [x;y] (x is just to the right of y.) (ibid, p. 92)

Each relation R has an inverse $\bar{R}$. If aRb then the pair (R,b) is stored on the property list of a and ($\bar{R}$,a) is stored on the property list of b. For each relation there are axioms. Further axioms describe how different relations interact. For instance, the set inclusion relation has the following properties:

$\mathcal{T}[\subset]$ i.e., set inclusion is transitive

$$\text{equiv}[x,y] \rightarrow x \subset y$$

$$a \in x \wedge x \subset y \rightarrow a \in y$$

The interaction between set inclusion and partg is expressed by the axiom

$$\text{partg}[x;y] \wedge z \subset y \Rightarrow \text{partg}[x,z]$$

In other words, if an x is part of a y and a z is a y then an x is part of a z. For example, if you know that mammals have hair and that whales are mammals, then you know that whales have hair.

Some of Raphael's relations represent particular information, some represent generic information. It is the generic relations which correspond to the kind of lexical relations we are working with: set inclusion, equiv, partg. and owng.

Apresyan, Žolkovsky, and Mel'čuk.

The Explanatory-Combinatory Dictionary of Apresyan, Mel'čuk, and Žolkovsky (1970) contains a wide variety of lexical relations. Whenever they notice a lexical regularity, they invent a lexical relation to express it. Their paper contains about fifty relations and they outline ways of combining the given relations to get still more. Many of these relations appear in an earlier paper by Zolkovsky and Mel'čuk, which emphasizes the importance of specifying the grammatical transformations associated with each lexical pairing. Suppose a story says:

> The prince's gift of a magic apple to Zamiya dismayed his mother.

In order to represent this correctly or answer a question like "What did the prince give Zamiya?" the system needs to know not only the lexical relation between *give* and *gift* but also the transformation which carries one string into another. In this lexicon the accompanying transformation will be indicated in the lexical entry for the relation, not in the lexical entry for the particular words *give* and *gift*.

Most of the relations given in these two papers are as appropriate in English as in Russian. Some, although appropriate in English, embody more sophistication than seems necessary in this project. The Soviet collection of relations is open-ended. They expect to identify more in further lexicographical work and to discover further properties of the relations already identified. This seems highly intuitive. It is probably the case that people go on expanding their repertoires of lexical relations and learning their properties and that this learning continues to a much greater age than the acquisition of syntax. Lexical relations can be added to our lexicon just by adding a lexical entry for the relation. At this point the actual addition of entries can only be done by internal manipulation. Eventually it would be preferable for the system to "learn" such relations or at least accept them in English form. The authors refer to their relations as functions and the examples are written in functional notation: Figur(*passion*) = *flame*; Anti(beautiful) = *plain ugly*. Since these functions are definitely not single-valued, we have used the term lexical relation, in deference to the mathematical conventions.

4. THE SET OF LEXICAL RELATIONS

The research reviewed above and our own experience with children's stories has led us to posit nine major categories of relations. See Table 2. These categories do not have any internal structure as a set; however many of the relations themselves seemed to share some commonality usually semantic, and so it became natural to group them into sets of categories. Our category list begins with the more familiar and classical relations of synonymy and taxonomy, and presents an expanded sub-categorization within antonymy. The grading category includes a somewhat diverse collection of three relations. The attribute relations and the part-whole category seem firmly motivated. The next two categories consist of co-occurrence or collocational relations. The last two groups of relations are paradigmatic in nature.

The set of relations presented here is by no means complete. Indeed, it is deliberately open-ended. Whenever a new lexical regularity is seen in the data, a new relation is added. In order to make the system of relations extensible, therefore, a separate lexical entry has been constructed for each relation containing its special properties and associated axiom schemes. (Examples of this appear below, for example, in section d. In addition definitions of properties, such as transitivity, and a discussion of their use in this system can be found in Appendix II).

There are several arguments for this methodology. Primarily we are convinced that lexical relations do not constitute a fixed set of language-universal semantic primes. We also feel that we have not yet discovered the most appropriate collection or our own use. In addition we hope to

Table 2. TABLE OF LEXICAL SEMANTIC RELATIONS*

a.	*Classical Relations*				
1.	T	taxonomy	lion	T	animal
2.	S	synonymy	amusing	S	$funny_1$
b.	*Antonymy*				
1.	COMP	complementarity	single	COMP	married
2.	ANTI	antonymy	hot	ANTI	cold
3.	CONV	converseness	to buy	CONV	(3-2-1-4) to sell
4.	RECK	reciprocal kinship	husband	RECK	wife
c.	*Grading*				
1.	Q	queuing	Monday	Q	Tuesday
2.	SET	set-element	flock	SET	sheep
3.	STAGE	manifestation	ice	STAGE	water
d.	*Attributes*				
1.	MALE	male - unmarked term	drake	MALE	duck
2.	FEMALE	female - unmarked term	lioness	FEMALE	lion
3.	CHILD	juvenile - parent	calf	CHILD	cow
4.	HOME	habitat - object	Africa	HOME	lion
5	SON	characteristic sound - animal	bark	SON	dog
6.	MADEOF	substance	ski	MADEOF	wood
e.	*Parts and Wholes*				
1.	PART	part - whole	horn	PART	cow
2.	CAP	head - organization	chief	CAP	tribe
3.	EQUIP	personnel - object	crew	EQUIP	gun
4.	PIECE	count - mass	lump	PIECE	sugar
5.	COMESFROM	provenience	milk	COMES-FROM	cow

The numbering matches that in the text.

f. Typical Case Relations

1	TAGENT	typical agent	conqueror	TAGENT	to conquer
2.	TOBJECT	typical object	dinner	TOBJECT	to dine
3.	TRESULT	typical result	hole	TRESULT	to dig
4.	TCAGENT	typical counter agent	loser	TCAGENT	to beat$_2$
5.	TINST	typical instrument	needle	TINST	to sew
6.	TSOURCE	typical source	earth	TSOURCE	to sprout
7.	TEXPER	typical experiencer	lover	TEXPER	to love
8.	TLOC	typical location	kitchen	TLOC	to bake$_2$

g Other Collocation Relations

1	COPUL	special copula verb	to fall	COPUL	victim
2	LIQU	destroying verb	to correct	LIQU	mistake
3.	PREPAR	verb which means prepare	to lay	PREPAR	table
4	DEGRAD	verb to deteriorate	to decay	DEGRAD	teeth
5.	{ INC	increase verb	to mount	INC	tension
	{ DEC	decrease verb	to shrink	DEC	cloth
6	PREPOS	preposition - object	on	PREPOS	list

h. Paradigmatic Relations

1.	CAUSE	cause - thing or action effected	to send	CAUSE	to go
2	BECOME	become + adj	to redden	BECOME	red
			to clean$_2$	CAUSE* BECOME	clean$_1$
3	BE	be + predicate	to neighbor	BE	near
4.	NOMV	process noun - verb	death	NOMV	to die

5.	ADJN	adjective - noun	solar	ADJN	sun
6.	ABLE	used in combination with case relations only	combustible	EXPER*ABLE	to $burn_1$
7.	IMPER	irregular imperative -	go ahead!	IMPER	to talk

Inflectional Relations

1.	PAST	past tense - infinitive	went	PAST	to go
2.	PP	past participle - infinitive	gone	PP	to go
3.	PLURAL	plural - singular	men	PLURAL	man

model the acquisition of relations at some later point. And finally, we are attempting to introduce some modularity of design into a difficult programming project.

a. The Classical Relations: Taxonomy and Synonymy.

Aristotle demanded that every definition begin with the statement of the genus to which the term belonged. The *genus* now is called a superordinate taxon and the relation between the term and its genus is labelled as the taxonomy relation. Even today commercial lexicographers following the classical tradition use taxonomy along with synonymy as the fundamental relations. These relations also have played an essential part in attempts at question-answering. In Raphael's (1968) system they appear as set inclusion and equivalence. In Simmons' (1973) system they are called IMPLY and EQ. The inference-making scheme in Marx's question-answering system is based on these two relations. For example, one of his test paragraphs says that a dog is brown and the question asks, "Is the animal brown?" (Marx 1972:224). A dictionary lookup of *dog* finds the taxonomic relationship between *dog* and *animal*. *Animal* is substituted for *dog* and the two sides match. Marx uses synonymy in the same way. Suppose the text says "John wants money" and a question asks "Does John desire money?" (1972:229). A dictionary lookup finds that *desire* is a synonym of *want*. The substitution of one for the other results in a successful pattern match.

1.) *Taxonomy*. The taxonomy relation T is expressed in many ways in English; perhaps "is a kind of" is the most typical:

A dog is a kind of animal.

A dog is an animal.

Dogs are animals.

The notation *dog* T *animal* is used to state this relationship. In the lexicon it is represented by an edge from the *dog* entry to the *animal* entry labelled T.

Werner's work on the taxonomy relation in memory models has shown that this relation plays a crucial role in lexical theory as well as in practical question answering. He has discussed the theoretical aspects of the taxonomy relation at length (Werner 1969, 1972, 1973, Perchonock and Werner 1969, Werner and Fenton 1970) and has used it in several studies (Werner and Begishe 1969, 1970).

Casagrande and Hale (1967) and Raphael (1968) use the name *inclusion* for this relation. It is certainly related to set inclusion. If A T B then the set of objects named by A, the extension of A, is a subset of the set of objects named by B, the extension of B. The set of dogs is a subset of the set of animals. If we look instead at the intensions of A and B, the sets of attributes implied by the terms, we again find a set inclusion relationship but in the other direction. If A T B then the intension of A includes the intension of B. The characteristics that let us identify an object as a dog include the characteristics that make it an animal. Because of the possible confusion about the direction of the inclusion relation, it seemed like a good idea to use another name. The term *taxonomy* is the natural choice since it is now well-known in anthropology.

2.) *Synonymy*. The synonymy relation poses some difficult philosophical problems. Do two words ever have the same meaning, or are there always differences? What criteria can be used to decide whether two words are synonymous? Apresyan, Žolkovsky and Mel'čuk (1970:5) have attempted to state a precise criterion: the two words should be semantically substitutable for each other, the meaning of one should be expressible through the other in any context. But this criterion substitutes one problem for another. How can one tell whether such a substitution is successful, whether the resulting sentences have the same meaning? It can be argued that different sentence forms exist precisely in order to allow the expression of differences in meaning. However impossible it may be to define synonymy precisely, this concept is used daily in ordinary discourse. Dictionary writers use it constantly. To simplify matters it is assumed here that the synonymy relation holds between two words whenever any of the dictionaries in the bibliography defines one as the other. This should be read as "rough synonymy" or "approximate synonymy."

b. Antonymy.

Antonymy has long been recognized as a lexical relation. *Webster's New Collegiate Dictionary*, for example, regularly lists antonyms. Its definition of *cold* includes "Ant. Hot" (1951:161). (The definition of *hot*, although it mentions *cold*, does not include "Ant. cold") The same dictionary defines *antonym* as "A word so opposed in meaning to another word that it negates or nullifies every single one of its implications." It is true that antonymy indicates some important facts about implications, and these need to be captured, but it is not true that antonymy involves negating every proposition in sight. The problem is that there are many kinds of oppositeness of meaning.

We have found four separate lexical relations which correspond to separate subcategories of antonymy: complementarity, antonymy proper, converseness, and reciprocal kinship.

1) *Complementarity,* isolated by Lyons (1968), is the kind of oppositeness that holds between *single* and *married* or *male* and *female.* The denial of one implies the assertion of the other: the assertion of one implies the denial of the other.

If John is married, then John is not single.

If John is not married, then John is single

If John is single, then John is not married.

If John is not single, then John is married.

This kind of relation seems to hold primarily between two adjectives or two adverbs belonging to the same primitive concept. If we set up a lexical relation COMP, then the appropriate axiom schemes seem to be, for the case where Adj_1 COMP Adj_2, if Z_2, looked at along dimension Z_1, has property Adj_1, then it also has the property $Not(Adj_2)$ and vice versa. In the notation used for the semantic representations in the question-answering system this is stated:

$$\mathit{Holds}\ (P(Z_1,Z_2,Adj_1)) \rightarrow \mathit{Holds}\ (P(Z_1,Z_2,\mathit{Not}(Adj_2)))$$

if, on the other hand, it has the property $Not(Adj_1)$ then it also has the property Adj_2 and vice versa.

$$\mathit{Holds}(P(Z1,Z_2,Not(Adj_1))) \rightarrow \mathit{Holds}(P(Z_1,Z_2,Adj_2))$$

(and similarly for adverbs). COMP is a symmetric relation. If A COMP B, then B COMP A. In other words it is its own inverse. In this lexicon if A is marked COMP B, then B is marked COMP A and so inferences are available in both directions. Anything marriageable is either married or single, not both; if one term applies, the other must not.

2.) *Antonymy*. Lyons restricts the term *antonymy* to the situation where the assertion of one implies the denial of the other, but the denial of one does not imply the assertion of the other. *Red* and *green* are antonyms in this sense. If X is red, it is not green. On the other hand, if X is not red it does not have to be green. It could be blue or yellow instead. *Hot* or *cold* behave in the same way. If X is hot then it is not cold, but if X is not hot we do not know for sure that it is cold; it may just be lukewarm. We set up a lexical relation ANTI to express this kind of antonymy. Again it applies particularly to adjectives and adverbs belonging to the same primitive concept. The lexical entry for ANTI gives an appropriate axiom scheme for the case in which Adj_1 ANTI Adj_2: If Z_2 is Adj_1 then it is not Adj_2.

$$\mathit{Holds}\ (P(Z_1,Z_2,Adj_1)) \qquad \mathit{Holds}(P(Z_1,Z_2,\mathit{Not}(Adj_2)))$$

(Similarly for adverbs.)

Verbs may be included in this kind of antonymy. Consider the pairs *love-hate* and *open-shut*. For a child, at least, "X loves Y" may imply "~X hates Y." The appropriate axiom scheme for $verb_1$ ANTI $verb_2$ would be: if a simple sentence containing $verb_1$ is true, then the negation is true when $verb_2$ is substituted for $verb_1$

$$\mathit{Holds}(R(verb_1,Z_1,Z_2,Z_3,Z_4))$$

$$\sim\mathit{Holds}(R(verb_2,Z_1,Z_2,Z_3,Z_4))$$

Since such verb pairs do not appear in our examples such problematical inferences have been avoided.

There are some important semantic realities here which are not being captured. There is a set of incompatible color terms: *red, orange, yellow, green, blue, purple, brown, black, white.* One can describe any small area of a physical object in one of these terms if it is forbidden to use hedges like *turquoise* and *pink. Hot* and *cold,* like *big* and *small,* are opposite ends of a scale. Between *hot* and *cold, warm* and *cool* can be placed somewhere. Binary lexical relations are not adequate here. Perhaps developments in the theory of fuzzy sets will eventually provide a better description.

There are logical problems here too. If the story says the toy is red, then we want to answer "no" to the question "Is the toy green?" But toys can be both red and green in spots, patches, or stripes. If the story says that the toy is red and green, we do not want to get lost in a self-contradiction.

Adjectives which imply grading (cf section *c* below) involve potential self-contradictions of a slightly different kind. Lyons discusses the sentence reference "A small elephant is a large animal." The current representation for that sentence in our system would be:

$$N_1 = Ncom(\text{elephant}, X_1) \qquad P_1 = P(\text{size}, X_1, \text{small})$$

$$N_2 = Ncom(\text{animal}, X_1) \qquad P_2 = P(\text{size}, X_1, \text{large})$$

For more details see the section on semantic representations. But *small* ANTI *large* so we must conclude from P_1 that $P(\text{size}, X_1, Not(\text{large}))$. The problem is that when we call something a small elephant we imply a comparison with some norm for elephants. However, this comparison does not appear in our representation. (This difficulty has also been discussed by Bierwisch 1969 and Simmons 1973.)

3.) *Converseness*. This is Lyons' name for a third kind of antonymy. As examples he gives the pairs *buy-sell* and *husband-wife*. This kind of oppositeness does not seem to involve negation at all. Rather it involves some kind of permutation of the associated individuals. Dale calls this relation reciprocity and explains it this way:

> *Buy* and *sell* are reciprocals, as are *give* and *receive*. What distinguishes these from antonyms (which they are, in a sense) is that whenever a sentence using one of them is appropriate, there is another appropriate sentence using the other member of the pair. For example, *John buys books from Bill* has the same meaning as *Bill sells books to John*. *He gave flowers to her* has the same meaning as *She received flowers from him*. This is a sort of "semantic passive"--like the passive transformation in syntax, it presents the same meaning from a different point of view. (1972: 144)

Whether Dale's sentences have exactly the same meaning or not is debatable, but anyone would agree that one implies the other. What is needed is some compact way to indicate what these other appropriate sentences are and to derive them when they are needed. Žolkovsky and Mel'čuk (1970) have a clever way of doing this for verbs. They use a notation of the form:

Buy CONV (3 2 1 4) *Sell*

to indicate that

X_1 buys X_2 from X_3 for X_4

becomes

X_3 sells X_2 to X_1 for X_4

We have borrowed this notation, applying it to cases rather than subjects and objects. It is interesting that the Soviets include regular syntactic passives in their discussion of this relation. Since in this system inferences are made on the basis of the fully formed semantic representations from which passives have been eliminated, they need not be included here.

4.) *Reciprocal Kinship.* If we had followed Reichenbach (1966) in treating kinship relations as functions of several arguments then we could have used CONV for pairs like *husband-wife* also. Since kinship and social relationships like *teacher-student* are expressed in terms of *have*, however, it makes sense to posit a new relation RECK for RECiprocal Kinship and other social terms. Husband and wife relationships are represented this way:

Len is Martha's husband. $R(\text{have}, X_1, X_2, \text{husband})$

Martha is Len's wife $R(\text{have}, X_2, X_1, \text{wife})$

We want to be able to derive one of these sentences from the other, using the lexical information husband RECK wife, i.e. if X_1 has X_2 as husband then X_2 has X_1 as wife. The axiom scheme for A RECK B says that if X_1 has X_2 as A then X_2 has X_1 as B.

$$Holds(R(\text{have}, X_1, X_2, A)) \rightarrow Holds(R(\text{have}, X_2, X_1, B))$$

Other kinds of converseness or reciprocity have not occurred often enough to warrant a separate relation and a separate axiom scheme. They are entered as individual inferences in each entry.

Antonymy seems to be a highly diverse lexical concept. With further study it may spawn still more lexical relations.

c. *Grading.*

Grading relations like antonymy relations involve alternatives of some kind. Graded alternatives appear to be organized in lists or other kinds of formal structures. Our collection of grading relations is in a state of flux, many aspects of grading are still not properly defined.

1.) *Queuing*. The notation Q is borrowed from Werner but used in a very restricted sense to connect adjacent items on lists, as in *Monday* Q *Tuesday*. It could be read "is immediately followed by."

2.) *Set-element*. SET relates the name for the set to the name of the elements, e.g. *flock* SET *sheep*. This is the relation which the Soviets call Mult. This relation seems to be particularly well-founded psychologically, for English has many special words of this type pride of lions, bevy of maidens, gaggle of geese, and it is certainly a source of word-play

3.) *Manifestation*. By contrast the STAGE relation, as in *ice* STAGE *water*, seems very shaky. The axiom schemes are not satisfactory and some of the territory is covered by the CHILD relation described in the section on attribute relations.

There seems to be a gap in our collection here. We have no parallel to the comparison relation of Casagrande and Hale (1967). Of course in the most common type of examples where the items related are taxonomic brothers, or cohyponyms as they are sometimes called, the comparison relation can be expressed by a combination of T and T. Recent work by Litowitz (1977) suggests that comparisons are an important component of the defining strategy of children. The boundary between the grading relations and the attribute relations described in the next section is also uncomfortably arbitrary.

d. Attribute Relations.

According to Casagrande and Hale (1967 168) whenever "X is defined with respect to one or more distinctive or characteristic attributes Y".

a definition is "attributive". Given this all-inclusive description it is not surprising that the attributive category was the largest in their sample. They propose several subcategories including stimulus properties like size and color, distinctive markers, habitat, behavior, sex, generation, and line of descent. But in order to facilitate inference we need to associate axiom schemes with each relation. Thus we have broken these subcategories into still more precise relations.

1.) *Male.* The relation MALE as in *drake* MALE *duck* relates the masculine to the unmarked term. We want to be able to infer that if something is a drake, then it is a duck and it is male, i.e.

$$Ncom(\text{drake},Z_1) \rightarrow Ncom(\text{duck},Z_1) \wedge P(\text{sex},Z_1,\text{male})$$

This axiom can be derived when needed from an axiom scheme in the lexical entry for *male* which says that whenever ZN_1 MALE ZN_2 then a ZN_1 is also a ZN_2 and it is male; i.e.,

$$Ncom(ZN_1,Z_1) \rightarrow Ncom(ZN_2,Z_1) \wedge P(\text{sex},Z_1,\text{male})$$

2.) *Female.* Similarly, FEMALE, as in *lioness* FEMALE *lion*, relates the name of the female to the unmarked term.

3.) *Terms for juveniles.* The most common attribute relation in our vocabulary is CHILD, which relates the term for the offspring to the term for its parent, as in *puppy* CHILD *dog*, *kitten* CHILD *cat*, *lamb* CHILD *sheep*. The lexical entry for CHILD contains the axiom scheme

$$Ncom(ZN_1,Z_1) \rightarrow Ncom(ZN_2,Z_1) \wedge P(\text{age},Z_1,\text{young})$$

When *puppy* and *dog* have been substituted for ZN_1 and ZN_2 respectively we get an axiom that tells us that if Z_1 is a puppy then Z_1 is a dog and Z_1 is young.

4.) *Habitat.* The habitat relation we have called HOME, so that *Africa* HOME *lion.*

5.) *Characteristic Sound* The relation SON was borrowed from the Soviets. SON relates an object and the verb expressing the kind of sound it produces.

to bark SON *dog*

to roar SON *lion*

to meow SON *cat*

to choo choo SON *train*

This relation seems to underlie a crucial part of the vocabulary of young children. Why is such a tremendous amount of time spent teaching children words like *meow?* Was this information once life-preserving or is it a way of teaching how sound is structured into words, the phonology of the language? For whatever reason, children who never see a farm are carefully taught to associate the sound *moo* with cows.

6.) *Substance.* The relation we call MADEOF as in

ski MADEOF *wood*

relates an object to the substance of which it is made. Casagrande and Hale classify as provenience both *batea:* "which is made out of mesquite" and *milk:* "we get it from a cow" (1967:184). Since in English these relationships are expressed in different ways, for example, the ski is made of wood - wooden ski, but milk comes from a cow - cow's milk, and since the appropriate inferences are different (the milk was once in the cow but the ski was not in the wood), we chose to classify them separately.

As the vocabulary expands we expect the list of attribute relations to expand. Litowitz (1977) is currently collecting definitions from children and isolating further relations. Smith and Maxwell (1977) have identified certain attribute relations which occur repeatedly in defining formulae in *Webster's Seventh:* COLOR, TIME, LOCATION, SIZE, and QUALITY. These relations, among others, will be added eventually to our lexicon.

e. Parts and Wholes.

1.) *Part-Whole.* The relation which links *finger* to *hand* and *carburetor* to *car* we call PART:

finger PART *hand*

carburetor PART *car*

The PART relation seems to be crucial in the definition of many every day objects. While it is clearly important in computer models of memory, it seems hard to isolate from natural English sentences. Raphael's (1968) SIR model used some subtle heuristics to determine whether a particular instance of the verb *have* should be represented by the *part* relation or the *own* relation. Sometimes dialog with a human is necessary to resolve the ambiguity. Simmons (1973) recognizes a three-way ambiguity in *have* which is represented variously as HASPART, POSSess, and ASSOC (1973:76)

Mary has long fingers	HASPART
Mary has money	POSSess
Mary has fun in the park	ASSOC

Apparently the part-whole relation is hard to identify in Papago also. Casagrande and Hale do not find it in their Papago sample. They classify as exemplification definitions which are translated into English as "cows have horns" and "horses have tails." However, on the basis of intuition and the word-association data of Russell and Jenkins (1954) they posit a fourteenth relation (1967:191):

Constituent: X is defined as being a constituent or part of Y.

The example given is *cheek-face.*

Apresyan, Mel čuk and Žolkovsky do not have an explicit part-whole relation but they do include two relations in this same area. We have borrowed

2.) *Head-Organization.* CAP relates the head to the organization.

chief CAP *tribe*

3.) *Personnel-Object.* EQUIP relates the associated staff to the organization or object they serve.

crew EQUIP *ship*

4.) *Count-Mass.* The relation PIECE which carves a countable chunk out of a mass also belongs to the part-whole family. For example,

lump PIECE *sugar*

item PIECE *news*

Jespersen was intrigued by this mechanism which he named *individualization* (1933:209); he discovered and listed many such examples. This seems to be the relation which the ECD calls SING (Apresyan et al., 1970:11).

5.) *Provenience.* We include here also the relation COMESFROM, as in *milk* COMESFROM *cow.* This is one aspect of the relation which Casagrande and Hale (1967) call provenience. (It should possibly be listed as an attribute relation along with its close cousin MADEOF.).

Our current lexicon contains only two axioms for the part-whole relation. One is transitivity: if X PART Y and Y PART Z, then X PART Z. The other, borrowed from Raphael, connects PART and Taxonomy. Essentially it says that if all X's are Y's and all Y's have Z's as parts, then all X's also have Z's as parts. There is an extensive philosophical literature involving this relation. Martin (1971) presents a system of axioms for part-whole and a review of work by Lesniewski, Woodger, and Tarski.

f. Typical-Case Relations.

Casagrande and Hale discovered that certain familiar objects, body parts, foods, tools, and other objects of material culture were most often defined not by the relations discussed above but rather by their use in daily life, by common activities associated with them. For example, under the "function" relation they classify examples in which "X is defined as the means of effecting Y" such as

eye: "...with which we see things"

money: "...we buy things with it" (1967:175)

The "operational" class includes examples in which X is defined as "the characteristic goal or recipient" of action Y

bridle: "...which they put on horses" (1967:178)

What they call the "spatial" relation also seem to be of this same type.

grindstone: "...on which a knife is sharpened" (1967:177)

Folk definitions collected from speakers of English often are of this variety, sometimes combined with taxonomy, e.g. "a house is a building in which people reside" (Evens 1975:340). Children in particular seem to prefer functional definitions (cf. Ruth Krauss' collection of children's definitions, *A Hole is to Dig*, 1952).

Apresyan, Mel'čuk, and Žolkovsky's system includes a family of functions S_1, S_2, S_3, S_4 which relate nouns and verbs or adjectives.

Their semantic structures are based on grammatical relations. For verbs these are a subject relation, a direct object relation, and two kinds of indirect object relations. The functions S_1, S_2, S_3, and S_4 correspond to these grammatical relations. S_1 relates the verb to its generic subject. S_2 relates the verb to its generic direct object, etc. For example (1970:10)

S_1(to sell)=seller

S_2(to sell)=goods (that which is sold)

S_3(to sell)=buyer, client, customer (the one to whom the goods are sold)

S_4(to sell)=price (that for which the goods are sold)

The ECD also contains four other substantive relations (1970:11). The values are nouns. The arguments can apparently be verbs, adjectives or nouns. First is S_{mod} which gives the noun denoting the mode of action; S_{mod} *(to write)=handwriting*. S_{loc} gives the noun denoting the place of the argument; S_{loc}*(action)=scene*. S_{instr} gives the noun denoting the instrument; S_{instr}*(communication)=means*, S_{instr}*(to think)=brain*. S_{res} gives the noun denoting the result; S_{res}*(to hunt)=bag*.

Since the semantic representations in the question-answering system are structured in terms of cases rather than grammatical relations we have set up a group of "typical-case" relations, one for each case relation in our case system. The typical-case relation relates the verb to typical fillers of that case argument slot. Thus, corresponding to the semantic relation AGENT we have a lexical relation TAGENT. The fact that someone who bakes can be called a baker is expressed in our lexicon

baker TAGENT *to bake*

The stuff that you eat is usually called food; *food* TOBJECT *to eat*. The result of digging is usually a hole; *hole* TRESULT *to dig*. When the Cubs beat the Cardinals the Cardinals are the losers; *loser* TCAGENT *to* $beat_2$. The thing you sew with is called a needle; *needle* TINST *to sew*. (This is the Casagrande and Hale operational relation.) Most plants sprout from earth; *earth* TSOURCE *to sprout*. One who loves is called a lover; *lover* TEXPR *to love*. People usually bake cakes in a kitchen; *kitchen* TLOC *to* $bake_2$. It should be noticed that the relation TLOC bears a close resemblance to HOME which gives the typical habitat for an animal or other object. The Soviet relation S_{loc} seems to include both. It is not clear that semantic theory can justify using two relations here. We have made a distinction because our system of semantic representations treats nouns and verbs differently, so that the associated axiom schemes for TLOC and HOME are formally different. It would be possible to use only one relation and test the argument for part of speech before choosing an axiom scheme Perhaps the real problem is in the system of semantic representations.

This particular choice of lexical relations is based on the particular case system being used. We claim, however, that the same basic scheme would be effective for a lexicon functioning with a *different* system of semantic representations based on *any other* set of case or grammatical relations. This is so since in this scheme corresponding to each semantic relation in the semantic representation there is a lexical relation in the lexicon relating verbs and typical fillers of argument slots.

g. Other Collocation Relations.

The relations in this group, like the typical case relations examined in the preceding section, are basically cooccurrence relations. They connect words which cooccur constantly and point to words which have special meanings in particular contexts. This is an important part of the lexical knowledge of the native speaker often neglected in dictionaries. Most of our relations in this group are borrowed from the Soviet lexicographers: COPUL, LIQU, PREPAR, DEGRAD.

1.) *Special Copula Verb.* The COPUL relation indicates the correct copula verb for nouns where be/become is not appropriate. For example, *to fall* is the special copula verb for victim, *to fall* COPUL *victim,* as in "Constance fell victim to Louis' charm."

2.) *Destroying Verb.* LIQU relates a noun and the verb which means to liquidate or destroy it. This seems to be useful in English as well and some examples belong to a child's vocabulary.

to erase LIQU *mistake*

to wipe out LIQU *traces*

3.) *Prepare for use.* The relation PREPAR relates a noun and the verb which means to prepare the object, to make it ready for use. This is particularly useful in making deductions about why people are doing things.

to lay PREPAR *table*

to make PREPAR *bed*

to load PREPAR *gun*

4.) *Verb to deteriorate.* The relation DEGRAD connects nouns and the appropriate verbs meaning to deteriorate, to go bad.

to decay DEGRAD *teeth*

to wear out DEGRAD *clothes*

5.) *Increase and decrease in activity.* The pair of relations INCrease and DECrease connect nouns and special-purpose verbs for increase and decrease.

to grow INC *child*

to shrink DEC *cloth*

(In terms of the Soviet relations $\overline{\text{INC}}$(x)=Incep(Plus(x)) and $\overline{\text{DEC}}$(x)=Incep (Minus(x)).)

6.) *Preposition - Object.* PREPOS behaves much like the relation which the Soviets call LOC. It links suitable prepositions to particular nouns. In English things go *on* lists, not *in* them. The fact that *on* is the appropriate preposition for *list* is recorded as *on* PREPOS *list.*

These are all collocational relations that we have observed in our data. Mel'čuk's ECD contains even more collocation relations but we have not included them because they seem too literary or too sophisticated for the vocabulary of childrens stories. For example, Bon (Apresyan Mel'čuk, and Žolkovsky 1970:13) points to attributes meaning "good":

Bon(*conditions*) = *favorable*

Bon(*aims*) = *lofty*

Both the typical-case relations and the other collocation relations which we have described are syntagmatic relations. They connect words with other words which coocur frequently in natural language sentences, sometimes with special meanings. We turn now to a group of paradigmatic relations which connect words which express aspects of the same core of meaning as it appears in various contexts or in different parts of speech.

h. Paradigmatic Relations.

The relations which we have grouped together as paradigmatic relations are highly disparate in kind and importance. CAUSE, BECOME, and NOMV are, we believe, essential to the structure of the English lexicon; ABLE and ADJN seem potentially quite useful. There seem to be very few examples of BE. All except BECOME were influenced by the inventory of Apresyan, Žolkovsky, and Mel'čuk.

1.) *Cause.* Traditional dictionaries use *cause* constantly to describe relationships between verbs. Dennison (1972) defines *to send* as "to cause to go". *Webster's New Collegiate* (1951) defines *to boil* as "to cause to bubble...." (p.96). Schank (1975) treats *cause* as the most important relation. McCawley (1975c) in discussing *to open* argues for two lexical entries, $open_1$ for "intransitive" uses: "the door opened" and $open_2$ for "transitive" uses: "John opened the door." $Open_1$ and $open_2$ are related by *cause: to* $open_2$ is *to cause to* $open_1$. McCawley's formulation will be followed here.

The first and longest entry in *Webster's New Collegiate Dictionary* for *open* belongs to the adjective. The definition of the intransitive verb begins "to become open". This suggests a renumbering:

$open_1$ - adjective - "the door is open"

$open_2$ - to become $open_1$ - verb intransitive - "the door opens"

$open_3$-to cause to $open_2$ - verb transitive - "John opens the door"

Open is only one of hundreds of verb-adjective homographs in English. *Cool* behaves like *open*. We start with the adjective $cool_1$, "the jello was cool". The intransitive verb $cool_2$ means "to become cool ", "the jello cooled in the refrigerator." The transitive verb $cool_3$ means "to cause to become $cool_1$". "Jane cooled the jello in the refrigerator." Other verb-adjective homographs like *clean* show a different pattern, the intransitive verb is missing.

$clean_1$ - adjective - The room was clean.

*$clean_?$ - to become $clean_1$. - The room cleaned.

$clean_2$ - to cause to become $clean_1$ - Jane cleaned the room.

Not all verb-adjective pairs are homographs. Modern English retains traces of an old suffix -en which turns adjectives into verbs. *To redden* is to make or become red. Sometimes the verb and the adjective are etymologically distant: *to* age_1 is to become *old*.

We need a lexical relation CAUSE relating *send* and *go*, $open_3$ and $open_2$.

send CAUSE *go*

$open_3$ CAUSE $open_2$

The appropriate axiom scheme for the case $verb_1$ CAUSE $verb_2$ tells us that: if the sentence containing $verb_1$ holds, then so does the sentence containing $verb_2$. Formally,

$$Holds(R(verb_1,Z_1,Z_2,Z_3,Z_4)) \rightarrow Holds(R(verb_2,Z_2,Z_3,Z_4,Z_5))$$

2.) *Become Adjective*. We also need a lexical relation BECOME relating age_1 and *old*, $open_2$ and $open_1$.

age_1 BECOME *old*

redden BECOME *red*

$open_2$ BECOME *$open_1$*

If $verb_1$ BECOME adj_1; then if the sentence containing $verb_1$ holds, then the object that did the *becoming* must now have the property expressed by the adjective, i.e.

$$Holds(R(verb_1, Z_1, Z_2, Z_3, Z_4)) \rightarrow Holds(\hat{P}(ZC_1, Z_1, adj_1))$$

where ZC_1 is the primitive concept corresponding to adj_1. (This axiom may conceivably react in uncomfortable ways with tense.) For the moment the relation between *$clean_2$* and *$clean_1$* the "cause to become" relation will be compounded from CAUSE and BECOME. It will probably occur often enough to deserve a name of its own, perhaps MAKE.

3.) *Be*. The relation BE parallels BECOME very closely. While BECOME relates the verb of becoming and the predicate adjective, BE relates the verb of being and the predicate adjective. For example, to *neighbor* is the verb which means to be *near*.

to neighbor BE *near*

This is the inverse of the relation which the Soviets call PRED. For some reason it seems to be much less common than BECOME.

4.) *Process Noun and Verb*. NOMV relates a process noun and its verb. *Death* is the nominalization of the verb *to die; death* NOMV *to die*. (This is the Soviet relation V_0 and the inverse of the relation S_0.)

5.) *Adjective and noun*. The relation ADJN, the inverse of the Soviet A_0, relates adjectives and nouns, as in *solar* ADJN *sun*. This relation may have to be split into two or more pieces. Magnus Ljung (1970) suggests that adjectives formed from nouns by adding -y, e.g. *sunny* as opposed to *solar*, mean "having more than a normal amount of" whatever the noun denotes. Adjectives in -al and -ful may present certain other semantic regularities.

6.) *Able.* The relation ABLE is used in combination with case relations only.

understandable	OBJECT*ABLE	*to understand*
literate	AGENT*ABLE	*to read*
legible	OBJECT*ABLE	*to read*

The Soviet version of this relation has different subcategories - Able_1, Able_2, Able_3, Able_4 - to indicate grammatical arguments of the verb. Able_1 *(to burn) = combustible* things are precisely those which can be subjects of the verb *to burn.* On the other hand, Able_2*(to eat) =edible,* since edible things are those which can be objects of the verb *to eat.* Since the semantic representation system in the question-answerer uses cases to connect verbs and arguments, we handle different kinds of ABLEness by combining ABLE with a case.

7.) *Irregular imperative.* The relation IMPER comes directly from the Soviet inventory. It relates colloquial imperative expressions to the appropriate main verb.

fire!	IMPER	*to shoot*
go ahead!	IMPER	*to talk*

This relation essentially involves very irregular imperatives; and this brings us to the inflectional relations.

i. Inflectional Relations.

Inflectional relations are dull but useful. Regular noun plurals and verb forms are handled by a suffix-chopping algorithm but words like *men* and *sang* defeat it completely. We get around this difficulty in essentially the same way as some commercial dictionaries do. A separate entry is included for these words. The lexical entry for *men* consists of PLURAL *man*. The entry for *sang* is PAST *to sing;* for *sung* we have PP *to sing*. The axiom-generator for PLURAL changes the number associated with the object if necessary and moves to the main entry to pick up other axiom schemes there.

The inflectional relations are, of course, paradigmatic relations, but are grouped separately because of their strong family resemblance and particularly uninteresting nature.

5. THE ORGANIZATION OF THE LEXICON AND THE SEMANTIC REPRESENTATIONS

The lexicon is a large network in which the nodes are lexical entries and the arcs are lexical relations; all the arcs are double-ended.

To represent the network in the data base, each entry contains a list of attribute-value pairs. Each pair consists of an arc (i.e. a relation name) and the name of the entry at the other end of the arc. Each lexical relation, L has an inverse $\bar{L}$. If $entry_1$ contains the attribute-value pair $L\text{-}entry_2$, then $entry_2$ contains $\bar{L}\text{-}entry_1$.

Each relation also has a lexical entry which gives its properties and also tells how to interpret lexical relationships in the predicate calculus. For example, the entry for *dog* includes the information *dog* T *animal (dog* is taxonomically related to *animal)*. The system uses the information in the lexical entry for T to interpret this as·

$$Holds(Ncom(\text{dox},X)) \rightarrow Holds(Ncom(\text{animal},X))$$

The lexical entry for T also tells us that T is transitive. The inventory of relations is expandable. To add a relation we need only add a lexical entry.

When the meaning of the word cannot be expressed solely in terms of lexical relations, a definition is added to the lexical entry, phrased in the same form as the semantic representations and using the same depth lexis. These lexical semantic relations are written in the same form as semantic representations for sentences. The lexical entry for *pet* includes the information that a pet is an animal which is owned by a human

$$Ncom(\text{animal},Z_1) \wedge Ncom(\text{human},Z_2) \wedge R(\text{own}, Z_2,Z_1)$$

This becomes the axiom

$$Holds(Ncom(\text{pet},Z_1)) \rightarrow Holds(Ncom(\text{animal},Z_1)) \wedge Holds(Ncom(\text{human},Z_2)) \wedge Holds(R(\text{own},Z_2,Z_1))$$

If an individual Z_1 is a pet, then Z_1 is an animal and is owned by a human Z_2.

Thus this lexicon is a relational network model with words and lexical relations as semantic primes. Definitions are written using lexical relations and first order predicate calculus formulas.

The design of this lexicon is independent of a particular representation scheme and the lexical relations we propose can be equally useful in another context. Nevertheless an overview of the semantic representations is included here in order to enable the reader to understand the notation in the examples of lexical entries in the next section. Anyone who does not find notational problems attractive should skip these paragraphs; with the exception of a few lines of formal details the rest of this paper will make sense without it.

An Overview of the System of Semantic Representations.

The question-answering system of which this lexicon is a fundamental part uses a first order predicate calculus system of semantic representations. As it reads a paragraph, the system makes an internal model of the story, identifying objects and events and the relationships between them. The representations are written in a first order predicate calculus so that they can be used in an existing theorem prover (Henschen, Overbeek, and Wos 1974). In a first order predicate calculus we are allowed predicates, functions, and quantifiers like "there exists" and "for all" but predicates are not allowed to be arguments of other predicates. This particular calculus is many-sorted; that is, there are many different classes of objects in the system.

Suppose a story begins:

> Peter heard a meow. Mother said, "The kitten is hungry." She sent Peter to the store. He bought milk and a big, red lollipop.

As we process this story we need first of all to recognize the different entities in the story. Here we have seven individual objects:

X_1 - Peter	X_5 - store
X_2 - meow	X_6 - milk
X_3 - Mother	X_7 - lollipop
X_4 - kitten	

We can write *Ncom*(lollipop,X_7) to signify that X_7 is a lollipop since *lollipop* is the common noun that names X_7. The story mentions two properties of the lollipop; it is big and it is red. The lexicon tells us that *red* is an adjective of color, so we represent this property using a functional notation

P(color,X_7,red)

Similarly,

P(size,X_7,big)

records the fact that the lollipop is big. These properties are numbered and put on a list for convenient retrieval. We may write

P_1 = P(color,X_7,red) P_2 = P(size,X_7,big)

This story also tells us some relations between entities. "He bought a lollipop," can be expressed as

R_1 = R(buy,X_1,X_7)

since he refers to X_1, Peter, and X_7 is the lollipop.The third sentence in the story:

She sent Peter to the store

contains a relation R_3

R_3 = R(send,X_3,X_1)

and a property of that relation

P_4 = P(direction,R_3,X_5)

The predicate *Holds* is used to make assertions. To assert the third sentence we write

Holds(P_4)

The connection between the milk and the lollipop in the last sentence is described by an interrelation *I*, $I(\text{and},X_6,X_7)$ so that the whole sentence becomes

$$Holds(R(\text{buy},X_1,I(\text{and},X_6,X_7)))$$

(There is a rule to rewrite this later as

$$\simeq Holds(R(\text{buy},X_1,X_6)) \wedge Holds(R(\text{buy},X_1,X_7))$$

but it is applied only if some kind of inference is required from this sentence, e.g., if a question asks, "Did Peter buy some milk?")

To obtain these representations we, of course, need a great deal of information from the lexicon (like the information mentioned above that *red* is an adjective of color and that *big* is an adjective of size). Lexical information is also used in setting up representations for questions like

1. What color is the lollipop?

   $P(\text{color},X_7,?)$

The answer to this question can be found by a simple matching process because the story representation already contains this kind of lexical information. A question such as

2. Did Peter buy some candy?

requires further lexical lookup since the word *candy* does not appear in the story. The answer is found using the lexical relation T (taxonomy or class inclusion) between *lollipop* and *candy* - the entry for lollipop includes T *candy*. Similarly, the entry for *candy* includes $\bar{T}$ *lollipop*, where $\bar{T}$ is the inverse or converse relation of T, which relates the same pairs of objects in the opposite order. Likewise a multiple choice question such as

3. Where does milk come from: cats cows trees cars

can be answered correctly using the provenience relation COMESFROM listed in the entries for *milk* and *cow*. The question

4. Where did Peter go?

is represented

$P(\text{direction},R(\text{go},X_1),?)$

The lexicon is then used to look for connections between *go* and *send*. The lexical entry for *send* includes the information CAUSE *go*. The entry for the lexical relation CAUSE contains several axiom schemes. With *send* and *go* substituted in the correct positions we get the axioms

$Holds(R(\text{send},Z_1,Z_2)) \rightarrow Holds(R(\text{cause},Z_1,(R(\text{go},Z_2)))$

and

$Holds(R(\text{cause},Z_1,ZR_1)) \rightarrow Holds(ZR_1)$

In order to answer the question

5. How old is the cat?

we must first identify the cat in the story, that is, recognize that a kitten is a cat and then realize that it is a young one. The lexical relation CHILD is essential to this task. The definition of *kitten* consists of CHILD *cat*. The lexical entry for *cat* contains $\overline{\text{CHILD}}$ *kitten*. The lexical entry for the relation CHILD contains axiom schemes which, when *kitten* and *cat* are filled in in the proper places, tell us that if X is a kitten then it is a cat and it is young. That is, if *Ncom*(kitten,X) then *Ncom*(cat,X) and *P*(age,X,young).

In addition some questions force us to look at the interaction between two or more lexical relations. To answer the question

6. What animal did Peter hear?

we need to know that a meow is a typical cat sound, which is expressed by the lexical relation SON, *meow* SON *cat*. We also need to know that a cat is an animal, *cat* T *animal*, and that a kitten is a young cat, as above *kitten* CHILD *cat*.

This has been an extremely brief introduction to the semantic system used in the question-answering scheme of which this lexicon is a part. For those who are interested in the representations themselves Appendix I contains a brief formal presentation. A more complete description is in preparation. (M. Evens and G. Krulee, "Semantic Representations for Question-Answering Systems.")

Lexical relations, we are convinced, are an extremely useful addition to any lexicon, whatever the underlying semantic system. The axioms which are associated with each relation, of course, have to be expressed in the semantic representations of the system in which the lexicon is being used.

6. THE FORM OF THE LEXICAL ENTRY

The most crucial step for the lexicographer is the design of the lexical entry. Somehow all the different kinds of lexical information previously decided upon must be neatly packaged into a compact, consistent, and accessible package. The lexicon is a large network in which the nodes are lexical entries and the arcs are lexical relations. Lexical entries can be found from an alphabetic list, so that the network may be entered at any point. There is a subnetwork containing lexical relations and their logical properties.

Each entry begins with the letter string which names it. Homographs are numbered 1,2,3,... to prevent confusion. Thus, $clear_1$ is the adjective, $clear_2$ is the verb 'to become $clear_1$', and $clear_3$ is the verb 'to cause to $clear_2$' or 'to cause to become $clear_1$'. Entries contain

(i) Category - Part of speech, sort, lexical relation, etc.

(ii) Irregular inflectional morphology.

This latter is stated in terms of a special set of lexical relations-- PAST, PP (past participle), and PLUR(al) are the only ones needed for our simple data-base. The lexical entry for *make* includes $\overline{\text{PAST}}$ - *made*, $\overline{\text{PP}}$ - *made*. *Made* has a separate lexical entry but a very short one

made	PAST -	*make*
	PP -	*make*

The lexical entry for *child* includes PLUR - *children*, the lexical entry for *children* consists of:

children PLUR - *child*

(iii) Lexical relations and pointers to their values in the form of attribute-value pairs. The lexical entry for *puppy* contains CHILD - *dog*. The lexical entry for *dog* contains $\overline{\text{CHILD}}$ - *puppy*. The lexical entry for the lexical relation CHILD tells us how to interpret these. It contains an axiom scheme which when filled in tells us that X is a puppy if and only if X is a dog and X is young *Ncom*(puppy,X) means that *Ncom*(dog,X) and also *P*(age,X,young). Information often classed as derivational morphology will be included here, the lexical entry for *soap*, for example, contains $\overline{\text{ADJN}}$- *soapy* Some of this derivational information could be stated instead in general rules and probably should be in any larger data base.

(iv) Parameters appropriate to particular categories.

(v) Definitions. These are in the form of logical inferences that may be drawn when a given word is used, and which are idiosyncratic enough not to be coded in terms of lexical relations. Only a few words have definitions. *Puppy*, for example, does not because the information that a puppy is a young dog is indicated by the lexical relation CHILD - *dog*. *Pet*, on the other hand, has a definition

$$Ncom(\text{pet},Z_1) \quad Ncom(\text{animal},Z_1) \wedge Ncom(\text{human},Z_2) \wedge R(\text{own},Z_2,Z_1)$$

When this definition is retrieved it is transformed into the axiom

$$Holds(Ncom(\text{pet},Z_1)) \rightarrow Holds(Ncom(\text{animal},Z_1)) \wedge$$
$$Holds(Ncom(\text{human},Z_2)) \wedge Holds(R(\text{own},Z_2,Z_1))$$

In other words, if some individual Z_1 is a pet, then Z_1 is an animal owned by some human Z_2.

Omitted from this lexicon are the examples which are an important and valuable part of other dictionaries. This system does not have the generalizing power to use examples effectively and, in addition, they occupy a great deal of space. The most natural way of handling examples in such a model might be to accumulate them from semantic representations of sentences which the system parses. The task of organizing, pruning, and generalizing from examples is too formidable to tackle here.

Nouns: Taxonomy seems to be the most important lexical relation for nouns, but many others appear in the texts as well.

dog T *animal*	A dog is an animal.
cent T *money*	A cent is a kind of money.
puppy CHILD *dog*	A puppy is a young dog.
soil S *earth*	Soil is the same thing as earth
cake TRESULT *bake*	The typical bring-into-being verb for *cake* is *bake*.
bubble TRESULT *blow*	The typical bring-into-being verb for *bubble* is *blow*.

The syntactico-semantic features are used in noun entries only.

Feature Names	*Feature Values*		
Gender	Male	Female	Neuter
Animateness	Human	Animate Not human	Inanimate
Number	Singular	Plural	
Count/Mass	Count	Mass	

Originally, following Winograd, the number and count features were combined into a single feature with three values: singular, plural, and mass. But McCawley has recently (1975a) given examples of plural mass nouns: *clothes guts, brains*, etc. It is impossible to argue with counterexamples from everyday language. The feature information can be expressed compactly as a vector of 1's and 0's.

	Gender	Animateness	Number	Count
	M F N	H A I	S P	C M
Ted	1 0 0	1 0 0	1 0	1 0
puppy	1 1 1	1 1 0	1 0	1 0
sugar	0 0 1	0 0 1	1 0	0 1
fish	1 1 1	0 1 0	1 1	1 1
boat	0 1 1	0 1 1	1 0	1 0
clothes	0 0 1	0 0 1	0 1	0 1

These features are used to determine pronoun choice, for example, not to provide semantic information. *Puppy* is marked as having the feature *human* so that the system can parse "the puppy who barks" and "the cat who walks alone."

Definitions for nouns begin with the specification of the function, *Ncom* or *Nprop*:

BANK *Ncom*(bank,Z_1): *P*(location,*R*(save,Z_2,Z_3),*Prep*(in,Z_1))

(A bank is a place where things are saved.)

Smith and Maxwell (1977) include here commonly understood metaphorical extensions, metaphorical cliches (e.g. *pitch=hell*). These also can be expressed by lexical relations (cf. the Soviet Figur function which gives figurative forms; presumably *pitch* Figur *hell)*. No obvious ones occur in this data base, so that this item is not currently included.

Sample entry for puppy:

Category: common noun

Relations: S *pup*

CHILD *dog*

Parameters: 111 110 10 10

The relationship *puppy* T *animal* is not included. It can be inferred from *puppy* CHILD *dog* and *dog* T *animal* This omission of relationships which can be easily inferred saves space but costs time. It is probably the case that people actually store these relationships directly. The fact that most, if not all, puppies in the child's world are pets is not stored either. This is open to question.

Sample entry for pet

Category: common noun

Relations: T *animal*

RECK *owner*

Parameters: III 110 10 10

Definition: $Ncom(\text{pet},Z_1)$: $Ncom(\text{animal},Z_1) \wedge Ncom(\text{human},Z_2)$

$$\wedge\ R(\text{own},Z_2,Z_1)$$

The word *owner* definitely belongs in the lexical universe of *pet*. We can recover it from the presence of *own* in the definition and the fact that *owner* TEXPER *own*. In a child's world, though, the pet-owner relationship seems to be a reciprocal kinship relationship like daughter-mother.

Non-Copula Verbs Every non-copula verb entry includes case information, in the form of a list of one or more arguments. For each argument we need four pieces of information:

(i) How it is realized syntactically: subject, object, or a list of prepositions.

(ii) The case(s) involved.

(iii) Whether the case must be explicitly specified (OBL), whether it is optional and unnecessary (OPT), or whether when absent it must be understood (ELLiptical).

(iv) Selection preferences: the top node of the taxonomy subtree.

(The classification names in (iii) are borrowed from the SPEECHLIS project, Nash-Webber, 1974). The elliptical cases belong to verbs which Chomsky (1965) marked [+object-deletion], which allow the object deletion transformation. Such verbs are *eat* and *read* where the object is easily understood. But this phenomenon also occurs with other associated noun phrases, not just the object. The sentence

John and Mary gave an alarm clock.

begs for a dative-experiencer in isolation, but sounds perfectly appropriate in answer to the question

What did John and Mary give the Andersons for a wedding present?

For *give* both the object and the experiencer may be deleted. A sign on the door saying "We gave" is acceptable because everybody understands that it means "We gave money to the United Fund " For *buy* the arguments are

	i	ii	iii	iv
1	Subject	agent, source	OBL	human, organization.
2.	Object	objective	OBL	thing

(In the Wall Street Journal dialect this argument is ELLiptical.)

3.	*From*	source	OPT	human, organization.
4.	*For*	instrument	OPT	money

For *give*, they are

1.	Subject	agent, source	OBL	human, organization.
2.	Object	objective	ELL	thing
3.	*To*, Object	experiencer	ELL	human, organization.

The next item tells us whether a verb is an action verb or not (±ACTION). Action verbs and adjectives can appear in imperative sentences but non-action verbs and adjectives cannot.

throw the ball!

* own the house!

be sensible!

* be tall!

These can also appear in embedded sentences dependent on imperative performatives like *order* and *tell*.

Sally told Sam to throw the ball.

* Sally told Sam to own the house.

Sally told Sam to be sensible.

* Sally told Sam to be tall.

And they can take the progressive aspect, while non-action verbs and adjectives cannot. This feature is important in calculating duration.

Sam is throwing the ball.

* Sam is owning the house.

Sam is being sensible.

* Sam is being tall.

The next item tells us whether the verb allows a regular passive or not. Only those which do not allow a passive are marked. Aprésyan, Mel'čuk, and Žolkovsky treat this also using lexical relations. Eventually this will probably be computable from other information in the entry.

Some important items apply only to verbs that take sentential complements. This includes the complementizer(s) the verb takes and whether or not it allows not-transportation. The possible complementizers are:

THAT Mother said that Mike should move.

FORTO Mother told Mike to move.

ING Mother did not like Mike's sitting there.

FROM Mother prevented Mike from going.

Verbs like *think* which give us roughly synonymous sentences whether *not* is in the main clause or the subordinate clause are said to permit *not*-transportation.

John didn't think Mary had gone.

John thought Mary hadn't gone.

Many verbs do not permit *not*-transportation, of course. These sentences are not synonymous:

John didn't say that Mary had gone.

John said that Mary hadn't gone.

This complementizer information is coded by adding to the entry: THAT, FORTO, ING, FROM, or NOT, as appropriate.

The next item is the implicational structure of the verb. There are seven such verb-classes and an eighth wastebasket class from which no inferences can be made (Joshi and Weischedel 1973; Karttunen 1970; Kiparsky and Kiparsky 1970); see table 3. In this system factives are the unmarked case since we always assume that we can assert arguments unless we are explicitly told not to. The lexical entry for each verb which can take a predicate complement and which is not a factive is not a factive is marked with its class name. Each class name appears in the lexicon with its appropriate inference pattern. For a negative-if verb, for example, this is:

If $R(V,Z_1,S)$ can be asserted then S can be denied.

If $R(Not\ (V), Z_1,S)$ can be asserted then S is in limbo.

Class	*Implicational Structure*	*Examples*
Factive	$R(S) \supset S$ $\sim R(S) \supset S$	R: realize S: Meg baked the cake Jerry *realized* that Meg baked the cake.
Implicative	$R(S) \supset S$ $\sim R(S) \supset \sim S$	*We managed* to finish the job.
Only-if	$\sim R(S) \supset \sim S$	They *allowed* Jim / Jim *had an opportunity* to visit China.
If	$R(S) \supset S$	Larry *persuaded* / *forced* Bill to accept the job.
Negative-if	$R(S) \supset \sim S$	Larry *prevented* Bill from winning the game.
Negative implicative	$R(S) \supset \sim S$ $\sim R(S) \supset S$	John *failed* to go. Hugh *refrained* from smoking.
Counter-factive	$R(S) \supset \sim S$ $\sim R(S) \supset \sim S$	Mary *pretended* that Ben went home.
Dull	No implications	Jerry *wanted* Meg to elope with him.

Table 3 Classification of Main Verbs in Predicate Complement Constructions (adapted from Joshi and Weischedel 1973)

The next to last item is the performative classification. The classification used is that proposed by McCawley (1975b) as an extension to the work of Austin and Vendler: Verdictive, Operative, Advisory, Imperative Commissive, Behabitive, Expositive (1-7). This is really a luxury in a recognition-only system for children's paragraphs. The only speech-act verbs involved in our data are *say* and *tell*. Performative classification does interact with syntax (especially modals), particularly in use with "would like to", "would", "will", and "let me".

The last item tells whether a verb takes indirect question (IQ). It is probably the case that when factivity and performative structure are understood, this item will be predictable. The IO verbs are apparently all expositives, but not all expositives are IQ's and the IQ classification seems to cut across McCawley's subclassification of the expositives.

Presuppositions are included in the definition, at present, rather than as a separate item.

Sample entry for $bake_1$: (The cake baked in the oven.) (The rock baked in the sun.)

Category: noncopula verb

Relations: $\overline{\text{CAUSE}}$ *bake*$_2$
$\overline{\text{TLOC}}$ *pan, oven*

Parameters: Args - 1. Subject - result, experiencer - OBL - *thing*.

Action - Yes

Sample entry for $bake_2$. (Mother baked a cake.)

Category: noncopula verb

Relations: T *make*

CAUSE *bake*$_1$

$\overline{\text{TAGENT}}$ *baker*

$\overline{\text{TLOC}}$ *kitchen*

Parameters: Arguments - 1. Subject - agent - OBL - *human*.
2. Object - result - ELL - *food, pottery*.
3. *For* - experiencer - OPT - *human, event*.

Action - Yes

This entry does not include *bakery*. A large lexicon could use a new lexical relation, STORE.

Sample entry for tell$_1$

Category noncopula verb

Relations T *speak*

S *say*

$\overleftarrow{\text{TAGENT}}$ *narrator*

Parameters Arguments - 1 Subject - agent - OBL - *human*.

2 Object - objective - ELL - *story*.

3. To, object - experiencer - OBL - *human*

Action - Yes

Comp - THAT

PERF (performative) - Expositive

IO - Yes

Sample entry for tell$_2$

Category noncopula verb

Relations T *speak*

S *command*

Parameters Arguments 1 Subject - agent - OBL - human.

2 Object - experiencer - OBL - animal human.

3 Object - objective - OBL - Sentence.

Action - Yes

Comp - TO

IS (implicative structure) - Dull

PERF - Imperative

IQ - No

Definition: $R(\text{tell}_2, Z_1, R(Z_2, Z_3), Z_3)$

$$R(\text{say}, Z_1, R(\text{order}, Z_1, R(Z_2, Z_3), Z_3))$$

(that is, if someone tells somebody to perform an action then he is saying that he orders that person to perform the action.)

Copula Verbs: These are marked as verbs of perception or verbs of motion as appropriate if they are not of the 'be-become-seem' variety. Verbs of perception are marked with the perceptual sphere. This helps to construct appropriate semantic representations. There is a close relation between the following sentences and we need to make inferences from one to another.

Sally listened to the trumpets. (active)

Sally heard the trumpets. (cognitive)

The trumpets sounded beautiful to Sally. (flipped)

The third sentence is called flipped because its arguments are switched from those in the first two. *Sound* is the flip perception verb for *hear* (cf. Rogers 1972). Thus, the entry for the copula verb *Sound* is marked:

type - *perception*

sphere - *aural*

flip - *hear*

Adjectives: The first special item for an adjective is the primitive concept. For *red* it is *color;* for *big* and *small* it is *size.*

The second item is the selection preference. For *red* it is *thing;* for *big* it is *thing, thought.* The selection preference could probably be stated once in the entry for the primitive concept and not repeated. Since it is useful to have it readily available in parsing, it is included separately in every adjective entry.

With adjectives as with verbs we often have causally related homographs. The *warm* in "warm coat" has a different meaning from the *warm* in "warm pie." A warm pie has a temperature greater than room temperature, but a warm coat makes you warm. These are called $warm_1$ and $warm_2$ and are connected by $warm_2$ CAUSE $warm_1$. How does one recognize which is which? If the head noun is *clothing* or one of the 'furnace-stove-oven' family or indeed anything else which has function *heat*, $warm_2$ is assumed.

Adjectives, like verbs, are marked 'Action - Yes' or 'Action - No'

Lexical entries for adverbs are very much like those for adjectives.

The main strategy followed in the design of the lexical entry has been to make it as compact as possible. It seems likely that more information will have to be added later.

7. SUMMARY

This lexicon is designed to serve as the global data base for a computer question answering system. It is therefore an integrated lexicon-encyclopedia, storing information needed for parsing, for development of an internal model, and for making inferences. Syntactic and semantic information are integrated into each entry.

Lexical entries are provided for all words which appear in the text except for those derived forms whose roots can be recovered by a trivial computation. Thus there are entries for *went* and *gone* but not for *goes* and *going*, for *unwanted* but not for *wanted*. Entries are also provided for some word combinations, such as *birthday cake* and *thank you*. Lexical entries are tied together by lexical-semantic relations which provide the internal structure of the lexicon.

Relations present both practical advantages and theoretical charms. The most immediate practical advantages appear in the mechanisms for saving space. Relations allow us to abbreviate entries, to state axiom schemes once and produce particular axioms only when they are needed, to include selection preferences in a compact form. They are in one sense a generalization of defining formulae already present in commercial dictionaries. Thus there is a possibility that we can extract some relation values automatically from existing dictionaries (cf. Smith and Maxwell 1977). From a theoretical standpoint relations provide a model of lexical memory with some modicum of psychological reality. Lexical-semantic relations and the theory of semantic fields suggest a tentative approach to the problem of identifying the context, of finding the right frame or script.

Appendix I. The Semantic Representations.

This is a somewhat more formal description of the system of semantic representations described informally in Section 5. (more details may be found in M. Evens and G. Krulee, "Semantic Representations for Question-Answering Systems," in preparation).

The representations are currently written in a many-sorted first order predicate calculus with individual constants and variables, function constants, and a predicate constant.

(1) Individual constants of each sort.

The object constants are written X_1, X_2 ....

Each corresponds to a unique object in the story.

(2) An infinite list of variables of each sort.

The object variables are written ZX_1, ZX_2, ...

When we do not wish to specify the sort of a variable it is labelled Z_1, Z_2, ...

(3) Function constants *Ncom* and *Nprop*.

These are used to name objects.

Ncom: Common nouns X objects → names

Ncom(boy,X_1)   *The boy* went home.

Nprop: proper nouns X objects → names

Nprop(Anne,X_2)   *Anne* went home.

(4) A. Function constant *R*   a five-place function

R is used to represent clauses with noncopula verbs.

R: noncopula verbs	X objects relations properties inter- relations names	X objects relations properties inter- relations names	X objects relations properties inter- relations names	
	X objects relations properties interrelations names	→ relations		

R(hit,X_1,X_2,Z_1,Z_2) ;   *The boy hit Anne.*

Most of the time this example will be abbreviated "*R*(hit,X_1,X_2)".

Unspecified arguments are often omitted for convenience in writing. Within the system they are represented by variables and thus will match anything. If the story says, "Donna sang a song", the internal representation is *R*(sing,X_3,X_4,Z_1,Z_2) with *Nprop*(Donna,X_3) and *Ncom*(song,X_4).

The question: "Did Donna sing?" becomes $R(\text{sing}, X_3, Z_3, Z_4, Z_5)$ which matches the statement from the story.

B. Function constant *P* a three-place function

P is used to represent expressions with adjectives and adverbs.

P: primitive concepts X objects X modifiers → properties
relations
properties
interrelations
names

$P(\text{manner}, X_1, \text{kind})$ *The boy is kind.*

The kind boy.

C. Function constant *I* a three-place function

I is used to represent expressions with conjunctions and conjunctive adverbs.

I: interrelational operators	X	objects relations properties interrelations, names modifiers	X	objects relations properties interrelations names modifiers → interrelations

$I(\text{and}, X_2, X_1)$ *Anne and the boy.*

$I(\text{because}, R(\text{smile}, X_2), R(\text{howl}, X_1))$

Anne smiled because the boy howled.

D. Function constant *Prep* a two-place function

Prep is used to present prepositional phrases.

Prep: preposition X objects → modifiers
relations
properties
interrelations

$Prep(\text{to}, X_2)$ *to Anne*

E. Function constants + * X ÷ two-place functions

These functions represent arithmetic operations.

+: objects X objects → objects

*

÷

+(2,3) *2+3*

F. Function constant *Not* a one-place function

Not: noncopula verbs → noncopula verbs

modifiers → modifiers

interrelational operators → interrelational operators

$R(Not(\text{sing}),X_2)$	*Anne did not sing.*
$P(\text{manner},X_1,Not(\text{kind}))$	*The boy is not kind.*
$R(\text{sing},I(Not(\text{or}),X_2,X_1))$	*Neither Anne nor the boy sang.*

(5) Predicate constant *Holds*.

This applies to every sentence at the top level.

It represents the underlying performative in the narrative paragraph.

$Holds(R_1)$ where $\begin{cases} R_1 = R(\text{see},X_2,R_2) \\ R_2 = R(\text{eat},X_1) \end{cases}$ asserts: *Anne sees the boy eat.*

Relations often appear as arguments of properties.

$P(\text{manner},R_2,\text{hungrily})$ *The boy eats hungrily*

So do properties

$P_1 = P(\text{manner},X_1,\text{kind})$	*The boy is kind.*
$P_2 = P(\text{quantity},P_1,\text{very})$	*The boy is very kind.*

The notation "$P_1 = P$(manner ,X_1,kind)" merely indicates that P(manner,X ,kind) is the first property formed in the representation of a particular story.

Both noun and verb phrase complements are represented by writing the subordinate relation or property as an argument of the formula which represents the main clause

Relative clauses are represented as interrelations. Clauses introduced by relative pronouns are ordinarily treated as conjoined main clauses. For adverbial clauses the conjunctive adverb serves as interrelational operator I(when,S_i,S_j) or I(because,S_i,S_j). Generic relatives and other types of generic expressions are treated as conditions.

(6) Lists and the Model of the Story World

As the system reads the story it forms a model of the world the story describes. The representations developed here are organized in five separate lists

Lists of Individuals

Lists of Names

Lists of Relations

Lists of Properties

Lists of Interrelations

We have defined the following sorts. For each sort we have some constants and an infinite supply of variables.

Sort	*Constant Symbols*	*Variable Symbols*
objects	$X_1, X_2, \ldots$	$ZX_1, ZX_2, \ldots$
names	$N_1, N_2, \ldots$	$ZN_1, ZN_2, \ldots$
relations	$R_1, R_2, \ldots$	$ZR_1, ZR_2, \ldots$
properties	$P_1, P_2, \ldots$	$ZP_1, ZP_2, \ldots$
interrelations	$I_1, I_2, \ldots$	$ZI_1, ZI_2, \ldots$
common nouns (house, dog, ...)	$A_1, A_2, \ldots$	$ZA_1, ZA_2, \ldots$
proper nouns (Anne, Sam, ...)	$B_1, B_2, \ldots$	$ZB_1, ZB_2, \ldots$
noncopula verbs (go, sing, ...)	$V_1, V_2, \ldots$	$ZV_1, ZV_2, \ldots$
primitive concepts (color, time, ...)	$C_1, C_2, \ldots$	$ZC_1, ZC_2, \ldots$
modifiers (red, on Tuesday, ...)	$M_1, M_2, \ldots$	$ZM_1, ZM_2, \ldots$
interrelational operators	$J_1, J_2, \ldots$	$ZJ_1, ZJ_2, \ldots$
prepositions (to, in, ...)	$T_1, T_2, \ldots$	$ZT_1, ZT_2, \ldots$.

When we do not wish to specify the sort of a variable, we call it $Z_1, Z_2 \ldots$.

Assume further the standard machinery of the first order predicate calculus:

The logical operators		$\sim$ (not)
		$\wedge$ (and)
		$\vee$ (or)
		$\rightarrow$ (if...then)
The quantifiers	$(\forall X)$	(for all X)
	$(\exists X)$	(there exists X).

References

Apresyan, Y D., I. A. Mel'čuk, and A. K Žolkovsky, 1970 Semantics and Lexicography. Toward a New Type of Unilingual Dictionary. In Kiefer, 1970, 1-33.

Bach, Emmon and Robert T. Harms, eds. 1968 *Universals in Linguistic Theory* Holt Rinehart and Winston, New York

Becker, Joseph D. 1975. The Phrasal Lexicon. In Schank and Nash-Webber, 1975.

Bierman, A. K. 1964. *Logic, a Dialogue.* Holden-Day, San Francisco

Bierwisch, M. 1969. On Certain Problems of Semantic Representations *Foundations of Language* 5, 153-184.

Bierwisch, Manfred and Ferenc Kiefer. 1970. Remarks on Definitions in Natural Language. In Kiefer, 1970.

Casagrande, J. B. and K. L. Hale. 1967. Semantic Relations in Papago Folk Definitions. In Hymes and Bittle, 1967, 165-196.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax* MIT Press, Cambridge.

Cole, P. and J. Sadock. 1977. *Syntax and Semantics 8 Grammatical Relations* Academic Press, New York

Dale, Philip S 1972. *Language Development, Structure and Function* Dryden Press, Hinsdale, Illinois.

Dennison, 1972. *Webster's Notebook Dictionary*, The Dennison Reference Library, Dennison Manufacturing Company, Framingham, Massachusetts.

Downing, Pamela. 1977. On the creation and use of English compound nouns, *Language* 53.4: 810-842.

Edmundson, H. P. and M. N. Epstein. 1972. *Research on Synonymy and Antonymy, A Model and its Representation.* University of Maryland Computer Science Center, College Park, Maryland, TR-185.

Evens, Martha W. 1975. *Semantic Representations for Question-Answering Systems.* Unpublished Ph.D. Dissertation, Department of Computer Sciences, Northwestern University, Evanston, Illinois.

Evens, Martha W. and Gilbert K. Krulee. In Preparation. Semantic Representations for Question Answering Systems.

Evens, Martha, Bonnie Litowitz, Judith Markowitz, Raoul Smith, and Oswald Werner. ms. *Lexical-Semantic Relations: A Comparative Survey.*

Fillmore, Charles. 1977. The case for case reopened, in Cole and Sadock, 1977, 59-81.

Flournoy, Catherine. 1975. Distribution and Collocation. Unpublished paper, English Department, University of Chicago

Fodor, Jerry and Jerrold Katz. 1964. *Structure of Language, Readings in the Philosophy of Language.* Prentice-Hall, Englewood Cliffs, New Jersey.

Garvin, Paul, ed. 1970 *Cognition: a Multiple View.* Spartan, New York.

Henschen, Lawrence, Ross Overbeek, and Larry Wos, 1974. A Theorem-Proving Language for Experimentation. *CACM* 17, 3, 303-314.

Hymes, D. and W. E. Bittle, eds. 1967. *Studies in Southwestern Ethnolinguistics.* Mouton, The Hague.

Jespersen, Otto 1933. *Essentials of English Grammar.* Allen and Unwin, London.

Joshi, Aravind K. and Ralph M. Weischedel. 1973. Some Frills for the Modal Tic-Tac-Toe of Davies and Isard Semantics of Predicate Complement Constructions. Third International Joint Conference on Artificial Intelligence, Stanford, August, 1973.

Karttunen, L. 1970. On the Semantics of Complement Sentences. In *Papers from the Sixth Regional Meeting of the Chicago Linguistics Society*.

Katz, Jerrold J. 1966. Semantic Theory Reprinted in Steinberg and Jakobovits. 1971. 297-307.

Katz, Jerrold and Jerry Fodor. 1964. In Fodor and Katz, 1964.

Kiefer, Ferenc, ed. 1970. *Studies in Syntax and Semantics*. Reidel, Dordrecht-Holland.

Kiparsky, Paul. 1975. On the Treatment of Fixed Phraseology in Generative Grammar. Paper delivered at the Eleventh Regional Meeting of the Chicago Linguistics Society, April, 1975.

Kiparsky, P. and C. Kiparsky. 1970. Fact. In Bierwisch and Heidolph, 1970.

Krauss, Ruth, 1952. *A Hole is to Dig, A First Book of Definitions*. Harper and Row, Evanston.

Lagaly, Michael, Robert A. Fox, and Anthony Bruck. 1974. *Papers from the Tenth Regional Meeting of the Chicago Linguistics Society*. Linguistics Department, University of Chicago.

Lakoff, George. 1971. On Generative Semantics. In Steinberg and Jakobovits, 1971.

Lawler, John M. 1972. Generic to a Fault. In Peranteau et al., 1972.

Levi, Judith. 1974. On the Alleged Idiosyncrasy of Non-Predicate NPs. In Lagaly et al., 1974.

Levi, Judith. 1975. *The Syntax and Semantics of Non-Predicating Adjectives in English.* Unpublished Ph.D. Thesis, Linguistics Department, University of Chicago.

Litowitz, Bonnie. 1977. Learning to Make Definitions. *Journal of Child Language*, Vol. 4, 1-16.

Ljung, Magnus. 1970. English Denominal Adjectives. *Gothenburg Studies in English* Vol. 21. Studentlitteratur, Lund.

Lyons, John. 1968. *Introduction to Theoretical Linguistics.* Cambridge University Press London.

Martin, Richard. 1971. *Logic, Language and Metaphysics.* New York University Press, New York.

Marx, Stephen M. 1972. *Deductive Question Answering with Natural Language Inputs.* Unpublished Ph.D. Thesis, Computer Science Department, Northwestern University.

McCawley, James D. 1968. The Role of Semantics in a Grammar. In Bach and Harms 1968, 124-169.

McCawley, James D. 1970. Semantic Representation. In Garvin, 1970, 227-247.

McCawley, James D. 1975a, Lexicography and the Count-Mass Distinction. First Annual Conference of the Berkeley Linguistics Society, 1975.

McCawley, James D. 1975b. Remarks on the Lexicography of Performative Verbs. Preprint.

McCawley, James D. 1975c. Conversational Implicature and the Lexicon. Paper Delivered at Northwestern University 5/19/75.

Miller, George and Philip Johnson-Laird. 1976. *Language and Perception*. Harvard University Press, Cambridge, Massachusetts.

Minsky, Marvin, ed. 1968. *Semantic Information Processing*. MIT Press, Cambridge.

Naroll, R. and R. Cohen, eds. 1970. *Handbook of Anthropology*. Natural History Press, Garden City, New York.

Nash-Webber, Bonnie. 1974. Semantics and Speech-Understanding. BBN Report No. 2896, AI Report No. 19, Bolt Beranek and Newman, Cambridge, Mass.

Palmer, F. R. 1976. *Semantics, A New Outline*. Cambridge University Press, London.

Peranteau, Paul, Judith Levi, Gloria Phares, eds. 1972. *Papers from the Eighth Regional Meeting of the Chicago Linguistica Societu*. Linguistics Department, University of Chicago.

Perchonock, Norma and Oswald Werner. 1969. Navaho Systems of Classification, Some Implications for Ethnoscience. *Ethnology*, Vol. 8, No. 3.

Quillian, M. Ross. 1968. Semantic Memory. In Minsky, 1968.

Raphael, Bertram. 1968. Sir, A Computer Program for Semantic Information Retrieval. In Minsky, 1968.

Rameh, Clea, ed. 1976. Semantics, Theory and Application. Proceedings of the Georgetown University Round Table on Language and Linguistics.

Reichenbach, Hans. 1966. *Elements of Symbolic Logic*. The Free Press, New York.

Rogers, Andy. 1972. Another Look at Flip Perception Verbs. In Peranteau et al., 1972.

Russell, Wallace A. and James H. Jenkins. 1954. The Complete Minnesota Norms for Responses to 100 Words from the Kent-Rosanoff Word Association Test. Studies on the Role of Language in Behavior. Technical Report 11. Department of Psychology, University of Minnesota.

Schank, Roger C. and Kenneth M. Colby, eds. 1973. *Computer Models of Thought and Language*. W. H. Freeman and Company, San Francisco.

Schank, Roger C. and Bonnie Nash-Webber, 1975. Theoretical Issues in Natural Language Processing, ACL Workshop, Cambridge, June, 1975.

Schank, Roger C. and the Yale A. I. Project. 1975. Sam - A Story Understander, Research Report 43, Yale University, New Haven.

Simmons, Robert F. 1970. Natural Language Question Answering, CACM 13,15-30.

Simmons, Robert F. 1973. Semantic Networks, Their Computation and Use for Understanding English Sentences. In Schank and Colby, 1973.

Smith, Raoul and Edward Maxwell. 1977. An English Dictionary for Computerized Syntactic and Semantic Processing Systems. In Zampolli and Calzolari, 1977.

Steinberg, Danny D. and Leon A. Jakobovits, eds. 1971. *Semantics, An Inter-Disciplinary Reader*. Cambridge University Press, Cambridge.

Webster's New Collegiate Dictionary. 1951. Based on Webster's New International. G and C Merriam Company, Springfield, Massachusetts.

Webster's New International Dictionary. 1958. Second Edition. G and C Merriam, Springfield, Massachusetts.

Werner, Oswald. 1969. On the Universality of Lexical/Semantic Relations. American Anthropological Association Annual Meeting, November, 1969.

Werner, Oswald. 1972. On the Structure of Ethnoscience. Conference on Methods of Structural Analysis, University of North Carolina, April, 1972.

Werner, Oswald. 1973. The Synthetic Informant Model. Proceedings of the 9th International Conference of Anthropological and Ethnological Sciences. Mouton, The Hague.

Werner, Oswald. 1974. Intermediate Memory, A Central Concept in Ethnoscience. Preprint, Anthropology Department, Northwestern University

Werner, Oswald and Kenneth Y. Begishe. 1969. *The Anatomical Atlas of the Navaho.* Mimeo, Anthropology Department, Northwestern University.

Werner, Oswald and Kenneth Begishe. 1970. The Taxonomic Aspect of the Navaho Universe. Proceedings of the 39th International Congress of Americanists, Lima, Peru, August, 1970.

Werner, Oswald and Joann Fenton. 1970. Method and Theory in Ethnoscience and Ethnoepistemology. In Naroll and Cohen, 1970.

Werner, Oswald and Martin Topper. 1976. On the Theoretical Unity of Ethnoscience Lexicography and Ethnoscience Ethnographies. In Rameh, 1976, 111-143.

Winograd. Terry. 1971. *Procedures as a Representation of Data in a Computer Program for Understanding Natural Language.* MAC TR-84. MIT Artificial Intelligence Laboratory. Ph.D. Thesis, MIT.

Zampolli, A. and N. Calzolari, eds. Computational and Mathematical Linguistics Proceedings of the International Conference on Computational Linguistics in Pisa, 1973. Olschki, Firenze. 1977.

Žolkovsky, A. K. and Mel'čuk, I. A. 1970. Semantic Synthesis. *Systems Theory Research,* 19, 170-243.