

REQUIREMENTS OF TEXT PROCESSING LEXICONS¹

Kenneth C. Litkowski
16729 Shea Lane, Gaithersburg, Md. 20760

As text processing systems expand in scope, they will require ever larger lexicons along with a parsing capability for discriminating among many senses of a word. Existing systems do not incorporate such subtleties in meaning for their lexicons. Ordinary dictionaries contain such information, but are largely untapped. When the contents of dictionaries are scrutinized, they reveal many requirements that must be satisfied in representing meaning and in developing semantic parsers. These requirements were identified in research designed to find primitive verb concepts. The requirements are outlined and general procedures for satisfying them through the use of ordinary dictionaries are described.

1. INTRODUCTION

Five years ago, Bolinger (1975, pp. 220-224), in discussing the increasing incorporation of meaning into linguistics, noted that these efforts had not yet made use of the insights of lexicography. The few substantial efforts, such as those spearheaded by Olney (1968, 1972), Mel'cuk (1978), Smith (1972), Simmons (1975), and Lehmann (1976) made some progress, but never came to fruition. Today, lexicography and its products, the dictionaries, remain an untapped resource of uncertain value. Indeed, many who have analyzed the contents of a dictionary have concluded that it is of little value to linguistics or artificial intelligence. Because of the size and complexity of a dictionary, perhaps such a conclusion is inevitable, but I believe it is wrong. To view the real potential of this resource, it is first necessary to develop a comprehensive model within which a dictionary's detail can be tied together. When this is done, the examination of definitions makes it possible to identify some requirements for semantic representation of lexical entries and for semantic parsers to be used in natural language processing systems. I describe herein what I have learned from this type of effort.

The principal purpose of this paper is only to show that a dictionary can provide very useful insights about what should go into a parser. I have not attempted to identify all the types of procedures that should go into a parser nor have I described a comprehensive and self-contained system that demonstrates how everything can be tied neatly together in a computer implementation. The specifications for such a system would be quite complex, as befits the complexity of a dictionary. However, I have adopted many of the notions in a manual system used to search for primitive verb concepts. In this system, I have developed and used a 400 page inverse dictionary, a thousand page coded dictionary, an elaborate index card filing system, and complex listings of rules I have followed. Many of these rules are based on the general schema presented in this paper. At this time, these notions have not only proved successful, but also have produced a wellspring of ideas yet to be fully articulated. I hope that I can portray some inkling of the fascination that has led me ever deeper into the recesses of a dictionary.

¹Paper presented at the 18th Annual Meeting of the Association for Computational Linguistics, Philadelphia, PA 1980 (Extended abstract at pp. 153-4 of the Proceedings).

It should be noted at the outset that the definitions in a dictionary were not developed to fit a grand scheme of semantic representation, possessing self-contained logical consistency (if such exists). The definitions have many flaws and the procedures I am following are uncovering many such flaws. In part, I expect that the rigorous approach necessary to extract the meaning content of definitions can assist in bringing about greater self-consistency within the overall structure of a dictionary.

2. GENERAL DESCRIPTION OF MY RESEARCH

I began my research (see Litkowski (1978) for a more complete description) with the objective of identifying primitive verb concepts by following definitional paths within *Webster's Third New International Dictionary* (W3). (All definitions quoted in this paper are taken from W3.) To search for primitives, it was first necessary to develop a comprehensive framework within which definitions could be analyzed. The theory of labeled directed graphs (digraphs) provided such a framework. Using digraphs, I developed several increasingly detailed models of the semantic structure of a dictionary. In these models, a point or node represents one or more definitions or concepts and a line or arc between these points represents a derivational relationship between definitions. Using such models, theorems of digraph theory were used to predict the existence and form of primitives within a dictionary. This justified continued effort to attempt to find such primitives. However, it immediately became clear that this would entail the development of semantic representations for definitions and the development of a semantic parser to transform definitions into these representations.

The models showed that the big problem to be overcome in trying to find the primitives is the apparent rampant circularity of defining relationships. To eliminate these apparent vicious circles, it is necessary to make a precise identification of derivational relationships, specifically, to find the specific definition that provides the sense in which its definiendum is used in defining another word. When this is done, the spurious cycles are broken and precise derivational relationships are identified. Although this can be done manually, the sheer bulk of a dictionary requires that it be done with well-defined procedures, i.e. with syntactic and semantic parsing. It is in the attempt to lay out the elements of such a parser that the requirements for semantic representations have emerged. The knowledge thus gained, developed incrementally and embodied in reduction rules, is then incorporated into procedures for the continued search for primitives. Thus far, these procedures have been used to reduce the initial set of 20,000 verbs in W3 to fewer than 4,000, with the prospect of much further reduction as the parsing principles are adopted. The search for primitives and the development of a semantic parser are proceeding hand-in-hand.

3. GENERAL REQUIREMENTS OF A PARSER

A lexical entry in a text processing system must contain information (1) to permit a parser to identify which sense of the word has been used in an utterance and (2) which characterizes the meaning of that sense so that it can be used for constructing the internal representation of the utterance. A syntactic parser, accessing lexical entries which contain only syntactical information, will take us only part of the way. Elaborate and complex semantic representational

and parsing procedures are necessary for a full representation of an utterance.

3.1 Syntactic Requirements

A parser must first be capable of dealing with the syntactic complexity of an utterance. In this paper, it is assumed that this can be accomplished by an ATN-type parser and that semantic parsing principles can be integrated with the syntactic procedures. However, it is recognized that the syntactic capabilities of the parser will continue to evolve; use of dictionary definitions as a large corpus upon which to test syntactic parsing principles may assist in this evolution. Moreover, by subjecting definitions to such a parser, it may be possible to identify and eliminate syntactic flaws overlooked by the lexicographers. On the other hand, the incorporation of semantic parsing into a syntactic parser may necessitate greater efficiencies for a parser. It is suggested below that what is included in a semantic representation of an entry in the lexicon can be used to determine what parsing paths should be pursued.

3.2 Semantic Requirements

The distinguishing characteristic of a semantic parser is that it must be capable of identifying which sense of a word is being used in a particular utterance. Rieger (1977, 1979) and Small (1979) have argued that this can be accomplished with sense discrimination nets, but feel that such nets must be developed by the AI community. In what follows, I suggest that nets can be developed through the analysis of definitions in an ordinary dictionary.

Rieger (1977) says that, to capture the meaning of a word, we need to identify all possible constructions in which it can participate. He then develops sense selection networks for determining which of the possible constructions is utilized in a particular instance. Such a network is a strategy for selecting an intended sense out of the mass of senses that a word might have. Small (1979) says that text processing requires complex interactions centered around word experts as the unit of linguistic knowledge. The central parsing process involves understanding the sense or role of a word in a particular context. A word expert is the same as Rieger's sense selection network and is cognizant of all possible contextual interpretations of the word it represents. Each expert should be capable of sufficient context-probing to determine successfully its functional or semantic role. A word expert can suspend its execution, stating conditions upon which it should be resumed (e.g. an adjective should wait for its accompanying noun, which would provide attributes which will help search through the adjective SSN). The parsing process results in a number of "concepts" (picture, event, and setting). Any word expert may make reference to a central tableau of control state descriptions during the disambiguation process.

A sense discrimination net consists of an ordered set of questions (the nodes of the network) and for each one the set of possible answers to the question (the branches emanating from each node), with terminal nodes being the semantic representations for a distinct sense. A word expert asks questions designed to identify its sense, but it can also develop other questions which the parser must then attempt to answer. Some of these questions do both at the same time. (e.g. to determine if a word is an adjective or a noun, the parser might need to inquire whether

the word to the right is a noun; to go further in analyzing the adjective, the parser would have to ask more detailed information about the sense of the noun.)

Rieger says that the questions will fall into the following classes about: (1) adjacent words, (2) the syntax or semantics of adjacent word senses, (3) invariant world knowledge, and (4) dynamic expectancies in the model. (In attempting to build SSN's from dictionary definitions, I want to make clear that I do not want to put general world knowledge or dynamic expectancies of an inference model into the semantic representations of the lexicon. However, it appears that definitions already do contain significant amounts of world knowledge.)

In this system, an SSN for a verb is described as a case framework, which is a specification of the syntax and/or semantics of the concepts that can be associated with that verb. Each verb SSN will have to make reference to the entities which the verb senses are capable of governing. At the bottom of the sense selection network will be the meaning case framework which must reflect a semantically accurate labelling of all concepts that it binds. Cases in meaning case frameworks are mandatory since it will otherwise be impossible to discriminate that sense.

Word experts affect sense discrimination (both of the instant word and of other words) and augment the conceptual information that constitutes the result of a parse. Small implies that complete disambiguation should take place by the time that a period is reached, I do not think that this necessarily follows. This will be discussed further in section 4.5 (dealing with ambiguity) and section 5 (dealing with multisentence processing).

Small (1979) notes that the augmentation of conceptual information that results from a parser is the cutting edge of his research because, although we may be able to complete some disambiguation, e.g. of a noun phrase, we may have to characterize that noun phrase further. As will be shown below, this may entail characterizing a noun phrase as a form, an appearance, a position, a quantity, a stage, a custom, a method, a tendency, or a property, thus going beyond the usual notions of linguistic cases and yet being necessary to an ultimate disambiguation in an SSN. This is part of what I am asserting can be done based on what is in the dictionary, using preposition definitions and perhaps also using constituents of the definitions themselves to help make such characterizations.

Rieger feels that sense selection networks can be developed by looking at the constructions in a single story, modeling them into an SSN, going to the next story, augmenting the SSN as necessary, and building the vocabulary in this way. Although this can be done, it will take a long time to build the lexicon in this manner. According to my thesis, this should be facilitated somewhat by using the definitions of a word as found in a dictionary.

Based on the foregoing, I believe that the following requirements must be satisfied by a semantic parser implemented through sense selection nets. Diagnostic or differentiating components are needed for each definition. Each definition must have a different semantic representation, even though there may be a core meaning for all the definitions of a word. Since the ability to traverse a net successfully depends on the context in which a word is used, each

definition, i.e. each semantic representation, must include slots perhaps with accompanying selectional restrictions) to be filled by that context. The slots will provide a unique context for each sense of a word. Context is what permits disambiguation. Since the search through a net is inherently complex, a definition must drive the parser in the search for context which will fill its slots. These notions are consistent with Rieger's; however, they were identified independently based on my analysis of dictionary definitions. Their viability depends on the ability to describe procedures for developing sense selection networks with the desired representations. This is discussed in section 4.

3.3 Representational Formalism

Although I use a specific formalism for semantic representation in the examples discussed in the next section, it was not developed for other than illustrative purposes. Any of a number of formalisms, such as those developed or used by Rieger (1977), Bobrow and Winograd (1977), or Norman et al. (1975), may satisfy the needs described in the next section. However, there are some basic requirements that any formalism must satisfy. For the most part, I follow Rieger in representing a verb definition in the form of n-tuples, with the predicate first, followed by other information, which may include obligatory cases that must be present in surrounding context, the use of sublists for providing selectional restrictions, and certain conditions that must be satisfied. Rieger also permits multiple-component case frames where a definition involves several meaning assertions as the terminal nodes of a sense selection network.

Another concept is that of a “descriptor”, which Rieger says is used to refer to a concept obliquely by describing it instead of naming it, for use when it comes time to identify the concept at SSN application time or a specific candidate whenever an actual model concept is required. He uses the following formalism:

[*D* (var) (feature1) ... (featureN)]

where (var) is an arbitrary reference name that satisfies the features. This formalism can be used to represent the unknown subject X of a verb with the features identifying any characteristics that the subject must satisfy. (See Rieger (1977, pp. 18-21) for further details.) In this schema, sublists within the features may be used to indicate further selectional restrictions that the context must satisfy.

4. DEVELOPMENT OF PARSER COMPONENTS

General procedures for developing sense selection networks are described using the intransitive senses of the verb “change” as examples. For this purpose, it is necessary to consider the intransitive definitions of “change” (shown in Table 1) and those definitions of other verbs where “change” is used intransitively as the main verb (shown in Table 2).

In discussing these general procedures, it should be noted that this has not been the principal purpose of my research, so I have not tried to develop them systematically, although that could be done. Rather, I have developed such procedures insofar as they have helped me

move closer to the primitives. Since I have not at this time identified the primitives from which all else is supposed to be derived (at least in theory), it goes without saying that any structures which I elaborate for the definitions of “change” will not be complete or accurate. On the other hand, it should be noted that the analysis which I have made shows further elements that have not been previously associated with this verb, and yet it has been accorded primitive status by some. This is one reason why I would argue that many nuances not yet been captured in assertions about meaning representation can be discovered from an analysis of dictionary definitions.

Table 1
Intransitive Definitions of “change”

<u>Sense</u>	<u>Definition</u>
1	become different in one or more respects without becoming something else
1a	lose or acquire some characteristic, property, or tendency ALTER
1b(1)	pass from one form, appearance, position, state, or stage to another SHIFT
1b(2)	<u>obs</u> pale or blush
1c	increase or decrease
1d	adopt different customs, methods, or attitudes <u>specif</u> experience a religious conversion
1e	<u>of the moon</u> pass from one phase to another <u>of the moon</u> <u>specif</u> pass through the phase of new moon
1f	<u>chiefly dial</u> turn sour <u>chiefly dial</u> become tainted
1g	shift one 's means of conveyance TRANSFER
1h	<u>of the voice</u> shift to lower register <u>of the voice</u> BREAK
1i	<u>Brit</u> shift gears
2	turn into or become something materially different from before
2a	undergo transformation or conversion - used with <u>into</u>
2b	pass over from one character or state - used with <u>to</u> undergo transition - used with <u>to</u>
2c	undergo substantial substitution or replacement or be wholly replaced
3	disrobe and rerearray oneself more suitably disrobe and rerearray oneself more suitably in clothes suitable for a social or formal occasion
4a	<u>obs</u> accept something else in return
4b	<u>obs</u> give up what one has in exchange - used with <u>for</u>
4c	engage in giving something and receiving something in return EXCHANGE

Table 2
Intransitive uses of “change”

<u>Use</u>	<u>Definition</u>
assibilate (vi)	change by introducing a sibilant sound
become (vi 2a)	change into being through taking on a new character or characteristic
break (vi 5c)	change sharply in purport, mood, or attitude
break (vi 6b)	change abruptly in line or set often with suggestion of opening
caramelize (vi)	change to caramel or a caramellike substance or color
chop (vi 3b)	change with or as if with the wind
chop and change (vi 2)	change esp. pointlessly or capriciously
coalify (vb)	change into coal by the process of coalification
come over (vi 1a)	change from one side (as of a controversy) to the other
come round (vi 2)	change in direction or opinion
curdle (vi 1)	change into curd
cut (vi 3g)	change in direction
deform (vi)	change in shape
devitrify (vi)	change from a vitreous to a crystalline condition usu. with loss of transparency and luster
differ (vi 1b)	change from time to time or from one instance or occasion to another
diphthongize (vi)	<u>of a simple vowel</u> change into a diphthong
effloresce (vi 2a)	<u>chem</u> change on the surface or throughout to a whitish mealy or crystalline powder from the loss of water of crystallization on exposure to the air
fade (vi 6a)	change gradually in loudness or visibility - used of a motion-picture image or of an electronics signal or image and usu. with <u>out</u> to specify change from loud to soft or bright to dark and with <u>in</u> to specify change from soft to loud or dark to bright
flash (vi 8)	<u>of a liquid</u> change suddenly or violently into vapor
flop (vi 3)	change suddenly (as from one course to another)
follow (vt 4b)	change in constant relation to
gel (vi)	change into a gel
gelatinize (vi)	change into a jelly
graduate (vi 2)	change gradually
hold (vi 1b(1))	not change
melt (vi 1a)	change from a solid to a liquid state usu. by the action of heat
push (vi 5b)	change in quantity or extent
quarter (vi 4)	change from one quarter to another - used of the moon
range (vi 6)	change within limits
reform (vi)	change for the better
resinify (vi 1)	change into a resin
rote (vi)	change by rotation
run (vi 11b)	change to a liquid state
run into (vt 1a)	change into
solate (vi)	change to a sol
specialize (vi 3)	change adaptively

transfer (vi 2)	change from one vehicle or transportation line to another
transform (vi)	CHANGE
transship (vi)	change from one ship or conveyance to another
turn (vi 3b(1))	change from ebb to flow or flow to ebb
turn (vi 4c(1))	change from submission or friendliness to resistance or opposition - usu. used with <u>against</u>
turn (vi 6b(1))	CHANGE - used with <u>into</u> or <u>to</u>
turn (vi 6b(2))	change to
turn off (vi 2b)	change to a specified state
waver (vi 1b)	change between objects, conditions, uses, or otherwise
weaken (vi 2)	change from a complex to a simple sound (as from a diphthong to a long vowel)
	change from a strong to a weak sound
	change from an open to a close vowel
whiffle (vi 1b(2))	change from one course or opinion to another as if blown by the wind

4.1 Syntactic Rules and Usage Notes

The first requirement for a sense selection network is that it should contain all the meanings of each set of homographs. For example, the SSN for “change” should contain all its noun and verb definitions. (Since a dictionary may contain several homographs in the same part of speech, e.g. “bore” has three distinct verb entries, all the definitions of each homograph would have to be combined into one SSN.) The first task of a parser, then, is to identify the correct part of speech for each word encountered in an utterance; in so doing, of course, the parser may have to deal with inflected forms of a word.

To some extent, syntactic parsing may permit further discrimination in the SSN. In fact, it may eventually be possible to group many definitions of a word according to the patterns of the syntactic context in which they can occur. This notion was previously explored with some success (see Earl (1973) for details and other references) under the rubric of “word government.” The extent to which this notion can be used for sense discrimination can be determined only after each SSN is elaborated, i.e. only after determining how much sense discrimination must rely on semantic considerations. Clearly, if syntactic parsing can do the job, a computer system will be much more efficient.

Certainly, in the case of transitive and intransitive verbs or verbs which use particles, syntactic parsing will be very useful in traversing the SSN. Many verbs have both transitive and intransitive definitions; for such verbs, answering the question, through syntactic parsing, whether the verb has an object can provide one branching node in its SSN. There is also a large number of transitive verbs in the dictionary with definitions which specify the object that must be present for them to be the applicable sense. For example, “bail” has one sense, viz., “to clear (water) from a boat by dipping and throwing over the side,” which requires the object to be the word “water.” In many cases, the object is specified generically, e.g. two senses of “abandon” specify the object as “oneself,” indicating that the object must be a reflexive pronoun; another

sense of “bail” specifies the object as “personal property,” indicating that the object must satisfy this selectional restriction. (The nature and treatment of selectional restrictions is discussed in section 4.4.) In these cases, questions about the object during syntactic parsing can provide additional branching nodes in the SSN.

Another significant class of definitions that may be recognized syntactically arises from verbs which take an adjective complement. The applicable definitions of such verbs always end with the phrase “to be;” e.g. one sense of “feel” is defined by “perceive oneself to be.” Thus, for verbs with this type of definition (or for verbs defined by a verb which takes an adjective complement), a syntactic question regarding the presence of an adjective complement can provide a branching node.

In W3, many verb definitions have accompanying usage notes, which provide information about the use of the verb being defined, usually in the form of a comment on idiom, syntax, semantic relationship, status, or various other matters. Of interest here are those usage notes which identify a particular idiom in which the particular sense of the verb is used, an accompanying particle (such as “up” or “out”), or an accompanying prepositional phrase. For example, 520 of the 788 senses of the verb “take,” all of which would be included in a complete SSN for this verb, involve some peculiarity of usage identified in the dictionary. Four senses of “change,” labeled 2a, 2b, and 4b in Table 1, have usage notes; three definitions in which “change” is used (as shown in Table 2: fade, turn (vi 4c(1)), and turn (vi 6b(1))) also have usage notes. The comments made in these usage notes can be used to formulate branching questions for an SSN, although not as directly as perhaps would be desired. These usage conditions do not specify that the presence of the idiom, particle, or preposition indicates the applicability of the definition, but only that the absence of the condition indicates the nonapplicability of the definition.

4.2 Preposition Definitions

Before continuing with the description of how to build a sense selection network, it is necessary to digress into a discussion of prepositional definitions, since they will play a crucial role in attempting to develop semantic representations of definitions. A preposition is defined as “a linguistic form that combines with a noun, pronoun, or noun equivalent to form a phrase that typically has an adverbial, adjectival, or substantival relation to some other word.” Prepositions are few in number (I have identified 126 in W3, half of which are phrases), but rich in significance for text processing, where they are typically used to identify conceptual cases. However, from my examination of preposition definitions, I do not believe their significance has been fully exploited.

Bennett (1975) asserted that spatial and temporal prepositions (a high percentage of all prepositions) lead to 23 primitive conceptual cases, even though in W3 the number of their definitions is at least two orders of magnitude higher. The difference seems to lie in the “apparent polysemy” which, as Bennett says, arises from the inclusion in prepositional definitions of “redundant features already determined by the environment.” In other words, many preposition definitions contain information about the context surrounding the preposition. I believe such

“redundancy” can be exploited in developing a semantic parser which will have a much greater facility for the type of conceptual case resolution that Small is concerned with.

Like verbs, prepositions appear to form a closed system in which they are defined in terms of other prepositions. However, unlike verbs, their primitives appear to be more easily identified. Of approximately 1400 definitions in W3, 70 percent are defined in terms of other prepositions, 20 percent are defined only by usage notes, and 10 percent are defined by verb forms. The usage note definitions, which have the appearance of primitives, uniformly begin with the phrase “used as a function word to indicate.” It is what follows the word “indicate” that can be used in developing the parser.

As mentioned above, in its definition, a preposition forms a relation between its object and some other word. The nature of this relation is what follows the word “indicate” in the usage notes. What I have found is that such relations follow certain patterns which can be articulated in formal recognition rules (1) for inclusion in a semantic parser, (2) for developing a semantic representation of verb definitions, and (3) for determining how to drive the parser. Usage note definitions of prepositions may specify:

- a condition (i.e. selectional restriction) that the object of the preposition must satisfy (e.g. that it is an age, a time, a state, or a group);
- a condition that must be satisfied by the context surrounding the prepositional phrase (e.g. the presence of an action verb, a specific type of action, something that is enveloped or covered);
- the assignment of a characterization to the prepositional object (e.g. that it is a location, an instrument, a purpose or goal, or a result (i.e. the traditional cases) or that the object is a thing observed as a spectator or a quantity of movement); and
- the assignment of a characterization to some element of the context outside the prepositional phrase (e.g. the fact of being present, of having parts or elements, or of being insertable into something else).

Every such definition does not contain all these specifications; at this time, I have not attempted an analysis of these definitions into their components. However, examples of how I have used these notions are described in subsequent sections.

At this point, I will only make some general observations about what these definitions imply with respect to parsing and semantic representation. In the first place, it appears that many words can be typecast with particular prepositional definitions, e.g. some verbs can be characterized as governing the patterns embodied in certain prepositional definitions. Such patterns are discernible from the definitions of such verbs.

Furthermore, it should be clear from the four types of specifications mentioned above how these definitions can be integrated into a general parser both for performing the parse and

for building a semantic representation of what is being parsed. By extension, these same considerations mean that it is possible to use specific prepositional definitions in parsing verb definitions and in creating the semantic representations of those definitions either explicitly or in terms of frames with slots accompanied by selectional restrictions that must be satisfied when a word is used in a particular utterance. These issues are dealt with in more detail below, along with specific examples.

4.3 Predicates, Slots, and Selectional Restrictions

The syntactic considerations described in section 4.2 clearly will not suffice for the construction of complete SSNs. Their further elaboration requires the development of semantic questions for use at the branching nodes. To develop such questions, it is necessary to know the semantic representations of the senses as they will appear at the terminal nodes of the SSN. (This is a circular statement, since it is necessary to know how to discriminate among the senses before distinct representations can be rendered. Therefore, the full elaboration of an SSN involves a process of iterative refinement of the discriminatory and representational components. Moreover, accurate representations must eventually be given in terms of primitive units; any intermediate representations must therefore be considered in this light.) In the following discussion, only the development of representations for verbs will be considered, although, as will be seen, the representation of other parts of speech is inextricably involved.

The representation of a verb definition essentially involves the assignment or identification of (1) an appropriate predicate, (2) the appropriate arguments or slots, and (3) selectional restrictions (if any) for each slot. The predicate and arguments would be arrayed as an n-tuple, with the selectional restrictions placed in the appropriate slots. The representation of particular definitions may involve a logical combination of more than one such n-tuple.

In Rieger's system, a predicate is considered a label for the accompanying argument configuration, but it has no intrinsic meaning. This is a convenient starting point for assigning a predicate, but in the case of analytic definitions (which consist of a genus and differentiae), the predicate should be the ultimate generic term. For example, the definitions shown in Table 2 can be assigned the predicate "change." This is more than just a label, but rather can be used to indicate that the basic argument configuration and selectional restrictions for the particular definition come from the definitions for "change." This is discussed in more detail below. (However, for the verb "change," the predicate "become different" will be used.)

The argument configuration must be developed from an analysis of the definitions and usually requires an examination of the definitions of the constituent words. To illustrate this process, definitions 1 and 2 of "change" will be used. For both definitions, the first argument or slot will be used to indicate the subject of the verb; since the subject may be in the PAT or AGT case (to be determined by the context), the corresponding slot for SUBJ will indicate that (PAT v AGT) is to be assigned to SUBJ. The words "become different" in both definitions imply the presence of four slots: FROM-STATE, TO-STATE, TIME1 and TIME2. However, since "different" is modified in two ways, some additional complexity is introduced. In definition 1, there is the notion that only an accidental attribute of the (PAT v AGT) "becomes different,"

while, in definition 2, there is the notion that some essential attribute “becomes different,” with the result that the (PAT \vee AGT) no longer exists. The net effect of this distinction is that for definition 1, there must be a “FROM-STATE,” a “TO-STATE,” and a “RESPECT” in which the change occurs, while for definition 2, there must be a “FROM-STATE” (which in this case is the SUBJ of “change”) and a “TO-STATE” which is the “RESULT” of the change. Possible semantic representations of these two definitions are shown in Figure 1.

Figure 1
Basic Frames for Definitions 1 and 2 of “change”

Definition 1: become different in one or more respects without becoming something else

[BECOME DIFFERENT (FROM-STATE NE. TO-STATE)
 ((SUBJ) (PAT v ACT)
 (* D * (PAT v ACT)
 (ESSENTIAL ATTRIBUTES ...)
 (ACCIDENTAL ATTRIBUTES ...)
 (“RESPECT” (TIME1) (FROM-STATE ...)
 (TIME2) (TO-STATE ...)))]

Definition 2: become something materially different from before

[BECOME DIFFERENT (FROM-STATE NE. TO-STATE)
 ((SUBJ) (PAT v ACT) (TIME1) (FROM-STATE)
 (* D * (PAT v ACT) ...))
 ((“RESULT”) (TIME2) (TO-STATE)
 (* D * (“RESULT” ...)))]

These representations could appear at terminal nodes of an SSN and one could be the contribution made as a result of parsing the verb “change,” unless further analysis were to lead to one of the subsenses of these definitions.

The final aspect of representing a verb definition requires the incorporation of selectional restrictions into the representations. As with the predicate and the argument configuration, selectional restrictions on what can fill particular arguments are derived from the definitional matter. As mentioned before, whatever representational formalism is used must have a capability for identifying the selectional restrictions that must be satisfied by the context. For the subsenses of definition 1 of “change,” the selectional restrictions may be so detailed as to fill in some slots of the basic frame for definition 1 or lead to the necessity for additional slots. As a result, using the terminology of Norman et al. (1975), the concept satisfying an argument of the basic frame may be completely determined in representing a subsense.

For the most part, the subsenses of definition 1 follow the basic frame shown in Figure 1 by providing information about the “respect” in which the subject of the verb “becomes different.” To determine that this is the case, it was first necessary to examine the definitions of the main verbs of each subsense. In each instance, the examination showed that the notion of “becoming different” is part of the meaning of the verb. Having arrived at this finding, it was then determined that most of the remaining information in the subsenses pertains to the “respect” in which the change occurs. These “respects” are shown in Table 3 for each subsense and would be used to replace the word “RESPECT” in the basic frame for definition 1. It should be noted that, for subsenses 1b(2), 1c, 1f, part of 1g, part of 1h, and 1i, it was necessary to search for the

“respect” in the definitions of the subsense's constituents. (It should be added that it was this analysis of the subsenses that led to the placement of the “RESPECT” slot under the slot for “ACCIDENTAL ATTRIBUTES” which in turn modifies the subject of the verb. Each “respect” in which the change could occur was required, via the phrase “without becoming something else,” not to change the essential nature of the subject.)

Table 3

Selectional Restrictions on “RESPECT” Slot of the Frame for Definition 1 of “change”

<u>Subsense</u>	<u>Selectional Restrictions</u>
1a	characteristic, property, or tendency
1b(1)	form, appearance, position, state, or stage
1b(2)	facial complexion
1c	size, quantity, number, degree, value, intensity, power, authority, reputation, wealth, amount, strength, etc.
1d	customs, methods, or attitudes <u>specif</u> religious attitudes
1e	phase of the moon
1f	capacity of being sour (e.g. disposition, taste, smell, acidity) capacity of being tainted (e.g. subject to putrefaction, corruption, moral contamination)
1g	means of conveyance vehicle or transportation line being used
1h	register of the voice voice's tone, pitch, or intensity
1i	method, tempo, or approach

The subsenses may also provide further selectional restrictions about the direction of the change. These restrictions, as shown in Table 4, would be added to the “FROM-STATE,” the “TO-STATE,” or as a relation between the two states. Other information may add new arguments (as in definition 1e) or give values to other slots (as in definitions 1e and 1h).

Identification of the predicate, the argument pattern, and the selectional restrictions for all definitions in itself requires a sophisticated semantic parser. Identification of the predicate can be accomplished in part by a taxonomic analysis of the type proposed by Lehmann (1976); for example, all the definitions in Table 2, the intransitive uses of “change,” could be assigned the predicate “CHANGE.” However, this is not valid, since ultimately the definitions in Table 2 should be assigned the predicate “BECOME DIFFERENT” or whatever primitive turns out to be appropriate.

Table 4

Other Selectional Restrictions of the
Frame for Definition 1 of “change”

<u>Subsense</u>	<u>Selectional Restriction</u>
1a	becomes deprived of (“lose”) comes to have (“acquire”)
1b(2)	becomes red (“blush”)

	becomes deprived of color or luster (“pale”)
1c	becomes diminished (“decrease”)
	becomes greater (“increase”)
1e	SUBJ = moon
	(TIME _x) (THROUGH-STATE = new moon)
1h	SUBJ = voice

In part, identification of the argument pattern can also be accomplished by taking advantage of the taxonomic relationships. Thus, argument patterns for the verbs in whose definitions “change” is the main verb can be used as the starting point for identifying the slots necessary to represent the definitions of “change.” Lehmann (1976) proposed to identify case argument patterns based on an analysis of the uses of “high-level verbs;” for “move,” 17 cases were identified in this way. However, this approach does not ensure that the definitions of “move” will be represented nor does it identify the obligatory arguments necessary to discriminate among the senses of “move.”

The development of procedures for identifying arguments and selectional restrictions ultimately must rely on the observation of patterns and the development of procedures for recognizing those patterns. As noted in the previous section, definitions of prepositions will play a significant role in the development of such recognition rules. For example, the phrase “in one or more respects” in definition 1 of “change,” combined with the fact that one definition of “in” is “with reference to,” could lead to the recognition rule that, whenever “change” (or any verb derived from it) is used in conjunction with a prepositional phrase beginning with “in,” the object of the preposition should replace the word “RESPECT” in the basic frame for definition 1 of “change.”

Although the development of such rules is not the goal of my research, it has become clear that they are necessary for developing sense selection networks. They are also necessary in identifying derivational paths within a dictionary in the search for primitive verb concepts.

4.4 Building Nets and Driving the Parser

The development of sense selection networks which incorporate semantic considerations should follow directly from whatever semantic representations of the senses have been created. Essentially, each semantic representation would be a terminal node of the SSN. However, because the set of definitions may have some internal ordering corresponding to subsenses, it may be necessary to permit nonterminal nodes of the network to provide a distinct sense of the word being parsed when it is not possible to reach a terminal node. Since an SSN is essentially an ordered set of questions leading to a terminal node (or one that has an attached semantic representation), its development consists of identifying the questions and putting them into the order in which the search through the net should be conducted. Although it would seem that the SSN for each word will have to be developed on its own, the semantic representations themselves may possibly be used to identify the questions and to determine the order of the search.

Each semantic representation (as it should be developed, not necessarily as shown in Figure 1) is an ordered n-tuple, perhaps with several sublists, which should contain all syntactic, contextual, and semantic information about the sense represented. The first elements of the n-tuple could contain syntactic information (e.g. noun or verb), followed perhaps by contextual information, and finally semantic information. The necessity of questions and branching nodes could conceivably be recognized by comparing and contrasting the n-tuples corresponding to each sense to find the first differences between them.

In the example of “change,” the first difference would be whether the semantic representation corresponds to a noun or verb; a second difference for the set of verb definitions would be whether the sense is transitive or intransitive, and so on down to the representations of the subsenses where the difference would be, for definition 1, the “RESPECT” of the change. Thus, the differences between the sense representations would identify where the branching nodes should be placed and could be used to develop the questions that should correspond to each branch. Although I have not attempted to follow this procedure, I believe that experience in analyzing such differences will make it possible to develop rules for identifying questions to be used in the SSN.

The questions that are developed for searching through an SSN will be essentially the same as those used in an ATN parser, except that it will be necessary to add semantic paths to the syntactic ones. One difference will be that, instead of developing an *a priori* model of the semantic grammar within which to conduct the search, each node will have to contain information which will tell the parser what to look for next. For example, the fact that the subject of “change” may be in the PAT or AGT case would require the parser to make particular searches in the context. Thus, in the use of “change” in defining “coalify” (see Table 2), the agent of the change is specified as “the process of coalification,” hence relegating the subject of “coalify” to the PAT case. It is possible to conceive of a standard set of procedures for making such a search (see, for example, Chafe (1970, pp. 243-244)).

When the question at a branching node requires a semantic resolution, the parsing requirements may be quite complex. Some of the difficulties that may arise can be illustrated by considering how a search would try to identify the appropriate sense of “change” in definitions where it is used as the main verb, i.e. those shown in Table 2. When “change” is used in defining another verb, three things may happen: (1) an argument of the frame for “change” may be given a value, (2) some further selectional restrictions may be added to an argument slot, or (3) an additional slot may be created. These possibilities may or may not help identify the applicable sense.

Since all the definitions in Table 2 use “change” as the main verb, the first thing to examine is whether the contextual matter of the definition provides a value that fills a slot in the basic frames for “change.” The basic difference between definitions 1 and 2 of “change” is that, in the latter instance, the subject of “change” is the “FROM-STATE” while the “TO-STATE” is the object of a preposition, usually “into” or “to.” Therefore, if there is an “into” in the context, there might be a presumption that definition 2 is the applicable sense, such as in the definitions for “become,” “coalify,” “curdle,” “diphthongize,” “flash,” “gel,” “gelatinize,” “resinify,” and

“turn (vi 6b(1)) .” However, even in these instances, it would be necessary to compare the subject and the object of “into” to determine if there has been an essential change in nature. For example, in the definition of “coalify,” the “process of coalification” is one in which (from the definition of “coalification”) “vegetable matter” undergoes a change “into coal;” this supports the choice of definition 2. On the other hand, the definition of “caramelize” seems to lead to the possibility that both definitions 1 and 2 are applicable because of the disjunction in the object of “to.”

In many of the definitions, the use of the verb “change” bears the same relation to definition 1 of “change” as its sub-senses, i.e. these definitions (shown in Table 5) indicate more specifically the “respect” in which the change occurs. For these definitions, it would be necessary to determine which subsenses would be applicable by comparing the “respect” indicated in Table 5 to the “respect” indicated in each subsense. For those definitions which contain both “from” and “to” prepositional phrases, such as those for “come over,” “devitrify,” “differ,” “melt,” “quarter,” “transfer,” “transship,” and “weaken,” the inference can be made that definition 1 is intended, although further analysis would be necessary to determine which “respect” and hence which subsense is intended.

Table 5
Selectional Restrictions on “RESPECT”
Slot Added in Uses of “change”

<u>Definition</u>	<u>Selectional Restriction</u>
break (vi 5c)	purport, mood, or attitude
break (vi 6b)	line or set
come round (vi 2)	direction or opinion
cut (vi 3g)	direction
deform (vi)	shape
fade (vi 6a)	loudness or visibility
push (vi 5b)	quantity or extent

However, for many of the definitions, such as those for “chop and change,” “graduate,” “hold,” and “specialize,” identification of the appropriate sense of “change” is not possible from the given context. In these situations, all definitions of “change” might apply and it would be necessary to await their use with further context before sense selection can take place.

The crux of these comments is that sense selection based on semantic questions inherently involves further computational analysis which delves into the definitions of the words used in context with “change.” The dictionary shows the complexity of such processing, but at the same time it shows the availability of large amounts of information that will aid this process.

To generalize from the observations made about disambiguation of “change” from whatever context in which it is used, e.g. determining the “respect” in which something “becomes different” or determining the fact that the patient of “change” “becomes something materially different from before,” we are first of all involved in what Small calls “conceptual case resolution,” although in a sense that may go beyond what he envisioned. We are no longer

dealing solely with identifying cases like “agent,” “patient,” and “experiencer” or concepts like “setting,” “event,” and “picture.” We are also concerned with whether concepts like “phase of the moon,” “tendency,” “complexion,” “quantity,” “strength,” “attitude,” or “taste” have been invoked by a particular context. Thus, we can say that semantic representation requires a deeper level of “conceptual case resolution” and an even greater level of complexity. We need to recognize that, when this is done, conceptual relationships between and among the definitions are being delineated. Hence, through this process, we are capturing the factual world knowledge which is embodied in definitions and thus providing some part of what Rieger finds necessary to incorporate separately in his AI Systems.

4.5 Ambiguity

It should be emphasized that a correct sense selection net does not automatically ensure that the correct sense will be identified. The context may simply be ambiguous. This is particularly noticeable in dictionary definitions where sufficient context for sense discrimination is seldom available. I suspect that the same is true for a large number of utterances, particularly if the context is limited to a single sentence. Clearly, it is necessary that a minimum context be present in order to permit discrimination.

When faced with ambiguity, a text processing system could try to find the plausible interpretation using world knowledge or dynamic expectancies of an inference model. Although such systems have been used with some success in limited domains, there are two significant difficulties that may arise. The first is that insufficient world knowledge has been provided to the system. (It may even be that the requisite world knowledge does not exist.) The second difficulty is that no inference model can be an accurate representation of how everybody reasons. (I suspect that each person follows a unique inferencing system.)

Based on the description of sense selection networks which has been laid out in the preceding sections, it would seem that another avenue for dealing with these difficulties has opened up. Since each SSN is designed to ask questions, perhaps it would be desirable simply to go as far as possible in interpreting and then send back the questions that currently stymie further disambiguation. This is, after all, what we would hope that a listener would do if we say something that does not make sense.

Such questioning could have great intrinsic value, because it would indicate (1) an inadequacy in the parsing system itself, (2) the lack of specific world knowledge (either in the system or in the world), or (3) the necessity for making an inference. Given the state of our knowledge about representing utterances, I would prefer a system that simply tries to capture what is present in an utterance. Accurate sense selection networks for relevant parts of the lexicon make it possible to build representations of scripts out of the components that we have at hand, rather than attempting to develop *a priori* inference models.

4.6 Movement Toward Primitives

As previously pointed out, it will eventually be necessary that semantic representations at

the terminal nodes of an SSN be given in terms of primitives. It was also mentioned that the procedures described in this paper have been used as part of a research effort designed to move toward identification of primitive verb concepts. The full set of procedures are described more fully in section 9 of Litkowski (1978), but it will be useful to describe how the notions described here are incorporated in that effort.

The basic procedure used in moving toward identification of primitives is through the development and application of rules which establish that particular words and definitions cannot be primitive. This requires a showing that a word or definition is derived from a more primitive concept and that a primitive cannot be derived from it. The nonprimitives are then set aside and further efforts focus on those words and definitions not yet eliminated.

The notions in this paper are applied by trying to show an explicit derivational relationship between two definitions. If the specific sense of the main verb of a definition can be identified and it can be shown that the definition contains differentiae which provide a value to an unbound argument in the semantic representation for the main verb, then the definition in which the binding takes place can be characterized as nonprimitive. For example, the definition of “diphthongize” (shown in Table 2) gives a value to the SUBJ of the basic frame for “change” and is thus inferred to be nonprimitive.

This notion of filling a slot is used more generally by developing recognition rules that identify particular word government patterns. For example, if a verb has a definition with the phrase “with an instrument” (thus creating an “instrument” argument) and is used in defining another verb accompanied by a “with” prepositional phrase whose object is defined as an instrument, the latter definition is characterized as nonprimitive. Some verbs in the first category are “apply,” “fasten,” “cut,” and “beat;” an example in the second category is the verb “knife.”

Other verb definitions are characterized as nonprimitive when they contain an optional component, i.e. one not necessary to discriminate among the senses of its main verb. Recognition rules are needed to identify different realizations of such components, such as the “manner” component which is optional or fills a slot for such verbs as “move,” “act,” “perform,” “utter,” “express,” and “behave.” Other definitions are characterized as nonprimitive when recognition rules establish that the definition consists of at least two distinct verb concepts. This is true of aspect verbs such as “cause,” “cease,” “begin,” “attempt,” “refuse,” and “serve.” Verb definitions of this type are very similar to those characterized as lexical relations by Evens and Smith (1978) or lexical functions by Mel'cuk (1978).

If relations or functions are used in representing definitions, it is important to understand that the function and the argument are parts of the lexicon, rather than the argument alone. In other words, we would have to do more than indicate that one definition is derived from another in building a semantic representation of an utterance. We would also have to represent the operator which gives rise to the derivation. For example, if we have the use of the “cause” operator, we would have to provide slots in our semantic representation for all the kinds of information which should be associated with the use of “cause.” (See Byerly (1979) for such a detailed specification.) Therefore, whenever we say that one entry in a dictionary is derived from

another by the application of some operator, we must be prepared to bring the representational contribution of the operator to the construction of the semantic representation of the derived entry.

5. MULTISENTECE PARSING

It seems that semantic representations of definitions in the form described must ultimately constitute the elements out of which semantic representations of multisentence texts must be created, perhaps with two foci: (1) describing entities (centered around nouns) and (2) describing events (centered around verbs). In parsing a single sentence, it seems clear that open variables, i.e. unfilled slots, will remain. Many such slots can be filled by later processing and parsing. Thus, at least part of multisentence text processing must recognize this fact, strip away the arbitrary bounds (which some would say are only convenient breath stops) of periods, and build semantic representations that deal with the entities and events by collapsing sentences which are used only to fill in some slots not yet filled. If multisentence texts can then be studied empirically, the structure of ordinary discourse will then be based on observations rather than theory.

6. CONCLUSIONS

Although the paradigm presented in this paper is complex, I believe that it is nothing more than what the lexicons of present AI systems are becoming. I believe that more rapid progress can be made with an explicit effort to exploit and not to duplicate the efforts of lexicographers.

REFERENCES

- Bennett, D.C. (1975). *Spatial and Temporal Uses of English Prepositions: An Essay in Stratificational Semantics*, Longman Linguistics Library, Vol. 17, Longman, New York.
- Bobrow, D.G. and T. Winograd (1977). "An overview of KRL, a knowledge representation language," *Cognitive Science*, Vol. 1, No. 1, pp. 3-46.
- Bolinger, D. (1975). *Aspects of Language* 2nd ed., Harcourt Brace Jovanovich, Inc., New York.
- Byerly, H. (1979). "Substantial causes and nomic determination," *Philosophy of Science*, Vol. 46, No. 1, pp. 57-81.
- Chafe, W.L. (1970). *Meaning and the Structure of Language*, University of Chicago Press, Chicago.
- Earl, L.L. (1973). "Use of word government in resolving syntactic and semantic ambiguities," *Information Storage and Retrieval* Vol. 9, pp. 639-664.
- _____ (1966). *Webster's Third New International Dictionary*, Encyclopaedia Britannica,

Chicago.

Evens, M.W. and R.N. Smith (1978). "A lexicon for a computer question-answering system," *American Journal of Computational Linguistics*, Microfiche 83 and Microfiche 81, Frames 16-24.

Lehmann, W.P. and R.F. Simmons (1976). *A Proposal to Develop a Computational Methodology for Deriving Natural Language Semantic Structures via Analysis of Machine-Readable Dictionaries*, University of Texas, Austin, Texas, September 28.

Litkowski, K.C. (1978). "Models of the semantic structures of dictionaries," *American Journal of Computational Linguistics*, Microfiche 81, Frames 25-74.

Mel'cuk, I.A. (1978). "A new kind of dictionary and its role as a core component of automatic text processing systems," *T.A. Informations*, No.2, pp.3-8.

Norman, D.A., D.E. Rumelhart, and the LNR Research Group (1975). *Explorations in Cognition*, W.H.Freeman, San Francisco.

Olney, J., C. Revard, and P. Ziff (1968). *Toward the Development of Computational Aids for Obtaining a Formal Semantic Description of English*, SP-2766/001/00, System Development Corporation, Santa Monica, California, 1 October.

Olney, J. and D. Ramsey (1972). "From machine-readable dictionaries to a lexicon tester: Progress, plans, and an offer," *Computer Studies in the Humanities and Verbal Behavior*, Vol.3, No.4, November, pp. 213-220.

Rieger, C. (1977). *Viewing Parsing as Word Sense Discrimination*, TP-511, Department of Computer Science, University of Maryland, College Park, Maryland, January.

Rieger, C. and S. Small (1979). *Word Expert Parsing*, TR-734, Department of Computer Science, University of Maryland, College Park, Maryland, March.

Simmons, P.F. and R.A. Amsler (1975). *Modeling Dictionary Data*, Computer Science Department, University of Texas, Austin, Texas, April.

Small, S. (1979). "Word expert parsing," *Proceedings of the 17th Annual Meeting of the Association for Computational Linguistics*

Smith, R.N. (1972). "Interactive lexicon updating," *Computers and the Humanities*, Vol.6, No.3, January, pp. 137-145.