

MODELS OF THE SEMANTIC STRUCTURE OF DICTIONARIES

Kenneth C. Litkowski
16729 Shea Lane
Gaithersburg, Maryland 20760

CONTENTS

Summary	3
1. Introduction	6
2. Attitudes Toward Dictionaries	8
3. Previous Research on Dictionaries	10
4. Description of Dictionary Contents	13
5. Basic Model	15
6. Expansion of the Model: Points as Definitions	17
7. Semantic, Structural, and Syntactic Parsing of Definitions	21
8. The Ultimate Model: Points as Concepts	30
9. Procedures for Finding the Primitives	33
10. Relationship to Efforts to Represent Knowledge in Frames	44
11. Final Remarks	47
References	49

Figures

1. A typical subgraph of the dictionary digraph using the basic model	17
2. Subgraph of model with points representing both single and multiple definitions	19
3. Subgraph of expanded model with grouping of definitions	20
4. Subgraph of a model incorporating a parsing system	22
5. Basic model, verb subgraph example subject to Rule 1	35
6. Basic model, verb subgraph example subject to Rule 2	36
7. Basic model, verb subgraph example subject to Rule 3	37

Table

1. Recognition rules for semantic components	43
--	----

SUMMARY

Ordinary dictionaries have not been given their due, either as sources of material for natural language understanding systems or as corpora that can be used to unravel the complexities of meaning and how it is represented. If either of these goals are ever to be achieved, I believe that investigators must develop methods for extracting the semantic content of dictionaries (or at least for transforming it into a more useful form).

It is argued that definitions contain a great deal of information about the semantic characteristics which should be attached to a lexeme. To extract or surface such information, it will be necessary to systematize definitions and what they represent, probably using semantic primitives. In this paper, I describe procedures which I have developed in an attempt to accomplish these objectives for the set of verbs in Webster's Third New International Dictionary (W3). I describe (1) how I have used the structure of the dictionary itself in an attempt to find semantic primitives and (2) how it appears that the systematization must incorporate a capability for word sense discrimination and must capture the knowledge contained in a definition.

The body of the paper is concerned with demonstrating that semantic information can be surfaced through a rigorous analysis of dictionary definitions. The first step in this process requires a comprehensive framework within which definitions can

be analyzed. In developing this framework, we must remember that each word used in a definition is also defined in the dictionary, so that we must be able to uncover and deal with vicious circles. The framework must also be capable of representing traditional notions of generative grammar to deal with the syntactic structure of definitions. A suitable framework appears to be provided by the theory of labeled directed graphs (digraphs).

Using points to represent dictionary entries and lines to represent the relation "is used to define", two models of the dictionary are described. From these models and from digraph theory, we can conclude that there may exist primitive units of meaning from which all concepts in the dictionary can be derived.

To determine primitive concepts, it is necessary to subject definitions to syntactic and semantic parsing in order to identify characteristics that should be attached to each definition. Syntactic parsing such as that implemented for systemic grammar by Winograd is the first step. A semantic parser must next be developed. It appears that definitions themselves, and particularly definitions of prepositions (which are used to express sense relations), will be of significant help in developing such a parser. Further work is necessary to develop procedures for surfacing from definitions information about the context which must be associated with each sense. It appears as if this parser will have more general use for ordinary discourse.

These notions lead to the ultimate model of a dictionary, where points represent concepts (which may be verbalized and symbolized in more than one way) and lines represent relations (syntactic or semantic) between concepts.

Based on these models, procedures for finding primitive concepts are described, using the set of verbs and their definitions from W3. Specific rules are described, based on some elementary graph-theoretic principles, structural characteristics of dictionary definitions, and the parsing of the definitions. These rules have thus far reduced the initial set of 20,000 verbs to fewer than 4,000, with further reduction to come as all rules are applied.

It is argued that this approach bears a strong relationship to efforts to represent knowledge in frames. Although much work is needed on the parser and on a computerized version of this approach, there is some hope that the parser, if expectations are borne out, will be capable of transforming ordinary discourse into canonical frame representations.

1. INTRODUCTION

During the past 15 years, scientists in many fields have been building a reservoir of knowledge about the semantic characteristics of natural language. Perhaps somewhat inexplicably these developments have for the most part ignored the semantic content of dictionaries, despite the fact that even a small one contains a vast amount of material. Some attempts have been made to dent these repositories, but the steps taken have been tentative and have not yet borne significant fruit, perhaps because the sheer volume and scope of a dictionary is so overwhelming. As a result, most studies have dealt with only a few definitions without a comprehensive assault on the whole. While such studies have led to many insights, it seems that the full usefulness of a dictionary's contents will be realized only when a comprehensive model of its semantic structure is developed.

Any system intended to provide natural language understanding must necessarily include a dictionary. If any such system is to achieve broad applicability, its dictionary must cover a substantial part of the natural language lexicon. For this to occur, the developers of a system must either create a dictionary from scratch or be able to incorporate an existing dictionary. Given the amount of effort that usually goes into development of an ordinary dictionary, the former alternative is rather impractical. However, little has been done toward meeting the latter alternative; with what follows, I will

describe the approach which I believe must be followed in transforming the contents of an ordinary dictionary for use in a true natural language system.

In order to be used in a language understanding system, a dictionary's semantic contents must be systematized in a way that the sense in which a word is being used can be identified. Before this can be done, it is necessary to characterize what is already contained in each definition. To do this, it seems necessary to write the meaning of each definition in terms of semantic and syntactic primitives. My purpose in this paper is (1) to describe how to use the dictionary itself to move toward identification of the primitives, at the same time (2) showing how this process can be used (a) to provide the capability for discriminating among word senses (i.e. characterizing the frames into which a given word sense will fit) and (b) to characterize knowledge contained or presupposed in a definition.

Before embarking on the description, it is necessary to point out some limitations which should be kept in mind as the reader proceeds. First, in trying to present an overview of my approach, I have had to forgo describing the detailed steps which I have followed to date. Second, even had I presented a full description, I would still have been short of providing sufficient details to enable computer implementation of any procedures. Third, since the approach presumes that concepts represented by the lexicon are the realizations of many as yet unknown recursive functions to be discovered by stripping away

one-layer at a time, results other than procedures to be used in stripping will not emerge until all layers have been removed. (However, I do argue that the "stripping" procedures are inherently useful, in that they will constitute a parser even in the intermediate stages.) Fourth, since I have not had access to a computer, which has become essential for significant further progress, I have been unable to determine how far the procedures I have developed would take me, so there is an inherent uncertainty as to how much further development is needed. Notwithstanding these limitations, I am hopeful that what is presented will provide a satisfactory framework for further investigations into the contents of dictionaries. I will comment further on these limitations and how they might be overcome at the end of the paper.

2. ATTITUDES TOWARD DICTIONARIES

Many of the significant contributors to the present understanding of meaning (such as Katz and Fodor 1963, Fillmore 1968 and 1971, Chafe 1970, Jackendoff 1974, Winograd 1972, and Schank 1972) have generally ignored dictionaries. Yet, each has presented a formulaic structure for lexical entries to serve as a basis for the creation of a new dictionary. Although their perceptions about the nature of language are well-established, their formalisms for lexical entries have not taken advantage of the equally well-established practices of lexicography.

The rationale underlying the development of new formalisms, expressed in some cases and implicit in others, is that lexical

entries in dictionaries are unsatisfactory because they do not contain sufficient information. These formalisms thus require that semantic features such as "animate" or "state" be appended to particular entries. While it is true that ordinary dictionary entries do not overtly identify all appropriate features, this may be less a difficulty inherent in definitions than the fact that no one has developed the necessary mechanisms for surfacing features from definitions. Thus, for example, "nurse" may not have the feature "animate" in its definition, but "nurse" is defined as a "woman" which is defined as a "person" which is defined as a "being" which is defined as a "living thing"; this string seems sufficient to establish "nurse" as "animate". In general, it seems that, if a semantic feature is essential to the meaning of a particular entry, it is similarly necessary that the feature be discoverable within the semantic structure of a dictionary. Otherwise, there is a defect in one or more definitions, or the dictionary contains some internal inconsistency. (Clearly, it is beyond expectation that any present dictionary will be free of these problems.)

The possibility of defective definitions has also generated criticisms, more direct than above, on the potential usefulness of a dictionary. On one hand definitions are viewed as "deficient in the presentation of relevant data" since they provide meanings by using "substitutable words (i.e. by synonyms), rather than by listing distinctive features" (Nida 1975:172). On another hand, the proliferation of meanings

attached to an entry is viewed as only a case of "apparent polysemy" which obscures the more general meaning of a lexeme by the addition of "redundant features already determined by the environment" (Bennett 1975:4-11). Both objections may have much validity and to that extent would necessitate revisions to individual or sets of definitions. However, neither viewpoint is sufficient to preclude an analysis of what actually appears in any dictionary. It is possible that a comprehensive analysis might more readily surface such difficulties and make their amelioration (and the consequent improvement of definitions) that much easier.

Even though dictionaries are viewed somewhat askance by many who study meaning, it seems that this viewpoint is influenced more by the difficulty of systematically tapping their contents than by any substantive objections which conclusively establish them as useless repositories of semantic content. However, it is necessary to demonstrate that a systematic approach exists and can yield useful results.

3. PREVIOUS RESEARCH ON DICTIONARIES

Notwithstanding the foregoing direct and indirect criticisms, some attempts have been made to probe the nature and structure of dictionary definitions. A review of relevant aspects of two such studies will help the material presented here stand out in sharper relief.

Olney 1968 describes the conceptual basis of many projected routines for processing a machine-readable transcript of

Webster's Seventh New Collegiate Dictionary (W7). The primary objectives of these routines were the development of

- "(a) rules for obtaining certain of the senses described for W7 entries from other senses described for the same entries or from senses described for other W7 entries from which the first (at least in typical cases) were derived morphologically; and
 - (b) semantic components and rules for combining them to yield specifications of senses that cannot conveniently be obtained by rules referred to in (a) above."
- (ibid.:6)

Although these objectives are reasonable, they do not take advantage of the possibility that the semantic structure of a dictionary might be a unified whole. As a result, any routines that are developed seem to require the serendipitous perception of patterns. Further, if a dictionary does have a unified semantic structure, it is not clear that a rule relating meaning to form will be relevant to a model of the semantic structure even though interesting results might emerge. It seems necessary to have some comprehensive view that will permit us to know whether a particular rule is well-formed. This lack of objective criteria also imperils any analysis that selects a subset of definitions for detailed analysis. The selection of a subset of the dictionary should arise from well-defined a priori considerations rather than an intuition that a particular

subset seems to be related. An example of this intuitive approach appears in Simmons 1975 and 1976.

In Quillian 1968, the analysis of dictionary definitions was part of a study of semantic memory, and for that reason was not concerned with the full development of a dictionary model. In that study, a person determined the meaning of a concept when he "looked up the 'patriarch' word in a dictionary, then looked up every word in each of its definitions, then looked up every word found in each of those, and so on, continually branching outward until every word he could reach by this process had been looked up once." This process was never actually carried out because (1) not all words in a dictionary were used in the computer files, (2) the process was terminated when a common word was found in comparing the meanings of two words, and (3) there was a belief that there are no primitive word concepts. The termination of a search as designed was necessary in any event since, without any restrictions, it is likely that a large part of the dictionary would have been reached on every occasion. More importantly, Quillian did not fully consider what was happening when branching led to a word already encountered, namely, that a definitional circularity was thereby uncovered. Such circularities which might be vicious circles, must be treated specially (as will be shown below), and hence, Quillian's unrestricted branching should have been modified. Quillian also overlooked the possibility that a concept common to two patriarchs is more primitive than either. The continued

comparison of more and more primitive concepts, along with restrictions on the outward branching, implies that primitive concepts actually do exist.

Based on these observations, I take, as a working hypothesis, the assumption that a dictionary may be a unified whole with underlying primitive concepts.¹ With this beginning, it is necessary to articulate a model of the dictionary which will permit an identification of the primitive concepts through the application of well-defined rules or procedures. It is proposed that what follows constitutes the first steps toward meeting this objective.

4. DESCRIPTION OF DICTIONARY CONTENTS

Since a dictionary contains much material, it is first necessary to delineate exactly what is to be modeled.² For this purpose, it is assumed that the semantic content of a dictionary essentially resides within its definitions, thereby excluding from formal analysis such things as the pronunciation, the etymology, and illustrative examples. As presently conceived, the analysis will focus on the word being defined (hereafter called the main entry), the definitions (including sense numbers and letters used as delimiters), part-of-speech

¹ No dictionary is likely to satisfy this assumption, which is only a theoretically desirable characteristic. The assumption enables us to exclude the definienda from the models.

² In the interests of space, I have glossed over a large number of intricacies that would have to be dealt with in arriving at a machine-readable transcript suitable for analysis. Several pages would be required to describe them fully.

labels, status or usage labels, and usage notes. The manner in which these features will be employed will be made clear as the analysis proceeds.

The hypothesized unified nature of a dictionary arises from the fact that definitions are expressed by words which are also defined³ (i.e., there is no semantic metalanguage). If we wish to understand the meaning of a given definition, then we must first understand the meanings of its constituent words. Since each constituent corresponds to a main entry, then, in order to understand the meaning of the given definition, we must understand the meaning of the constituent words' definitions. Continued repetition of the process is nothing more than the outward branching process described by Quillian; however, as mentioned before, we must make this branching more disciplined in order to deal with vicious circles and avoid unwanted circularities.

If we are to have a fully consistent dictionary, its model must show how each definition is related to all others. Thus, for each definition, X, the model should enable us to identify (1) those definitions of the constituent words of X that apply and those that do not apply, and (2) the production rules that generated X from these definitions. For example, in the definition of the noun broadcast, "the act of spreading abroad",⁴ it

³ There are some exceptions to this assertion, such as proper names, biological taxa, and other special symbols, as pointed out by the Journal's referee.

is necessary that the model indicate (1) which of the definitions of the, act, of, spread, and abroad apply, and (2) the production rules by which the and act and all other collocations) occur together. If this can be done for each definition in the dictionary, and if any inconsistencies are reconciled, then, as will be shown, it should be possible to find the primitive concepts in the dictionary and to transform each definition into a canonical form.

5. BASIC MODEL

The theory of (labeled) directed graphs (digraphs)⁵ is used as the formalism for the models. Digraph theory deals with the abstract notions of "points" and "directed lines"; its applicability to the problem before us therefore depends on how these notions are interpreted. In this respect, it is important to distinguish the manner in which this theory is used here from the manner in which it previously has been used in semantics and linguistics. The two most common uses are (1) where trees display phrase and syntactic structures (cf. Katz and Fodor 1963), or (2) where directed graphs portray the sequential generation of words in a sentence or phrase (cf. Simmons 1972). In these cases and others (cf. Quillian 1968 and Bennett 1975) graphs are used primarily as a vehicle for display

⁴ All definitions used in this paper are taken from Webster's Third New International Dictionary, Encyclopaedia Britannica, Chicago, 1965.

⁵ Terminology for digraphs follows Harary 1965.

and no results from graph theory are explicitly employed to draw further inferences. However, as used here, graphs constitute an essential basis for the analysis and hence will play an integral role in a number of assertions that are made.

In the simplest model, a point can be interpreted as representing all the definitions appearing under a single main entry; the main entry word can be construed as the label for that point. The part-of-speech labels, status or usage labels, and usage notes are considered integral to the definitions and may be viewed as part of a set of characteristics of the individual definitions. A directed line from x to y will be used to represent the asymmetric relation " x is used to define y "; thus, if the main entry x appears exactly or in an inflected form in a definition of y , then xRy . (This does not preclude a distinct line for yRx or xRx .) Therefore, we can establish a point for every main entry in a dictionary and draw the appropriate directed lines to form a digraph consisting of the entire dictionary. (This digraph may be disconnected, but probably is not.) An example, which is a subgraph of the dictionary digraph, is shown in Figure 1 on the next page. Except for broadcast, only the labels of each point are shown, but each represents all the definitions appearing at its respective main entry. The directed line from act to broadcast corresponds to the fact that "act is used to define broadcast", since its token appears in "the act of spreading abroad". In this model, the token "spreading" is not represented by a point, since it is not a main entry.

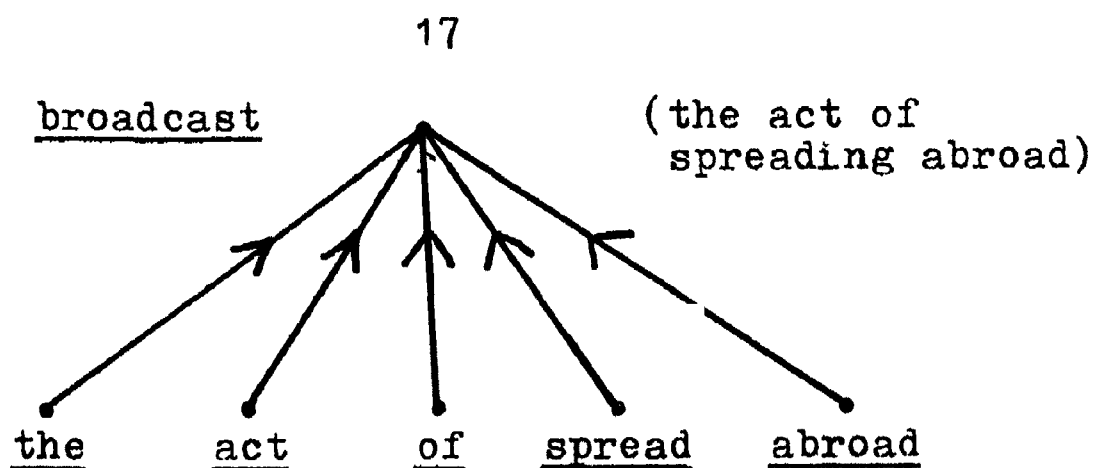


Figure 1. A typical subgraph of the dictionary digraph using the basic model.

Since the definition shown is not the only one for broadcast, this point has additional incoming lines which are not shown.

The resultant digraph for even a small dictionary is extremely large, perhaps consisting of well over 100,000 points and 1,000,000 lines. Clearly, such a digraph provides little fine structure, but even so, it does have some utility. The manner in which it can be used is described in Section 9.

6. EXPANSION OF THE MODEL: POINTS AS DEFINITIONS

Letting each point in the basic model represent all the definitions of a main entry provides very little delineation of subtle gradations of semantic content. As a first step toward understanding this content, it seems worthwhile to let each point represent only one definition. However, the basic model will not trivially accommodate such a specification (primarily because of the interpretation given to the directed line), and thus it must first be modified.

In the basic model, the existence of a line between two points, x and y , asserts that xRy , i.e., " x is used to define y ". Since the points represent all the definitions under the

main entries, the existence of a line arises from the simple fact that x appears in at least one of y 's definitions. If the point y represents only one definition, say y_j , there is no difficulty in saying that xRy_j . However, if we wish every point to represent only one definition, then we must find the definition of x , say x_i , for which x_iRy_j is true. Referring to the subgraph in Figure 1, this amounts to determining, for example, which definition of abroad is used to define the token "abroad" in "the act of spreading abroad", that is, finding the i such that "abroad _{i} Rthe act of spreading abroad" or "abroad _{i} Rbroadcast _{j} ".

It should be intuitively clear that interpretation of points as single definitions is desirable. However, there are no a priori criteria by which the appropriate value of i can be determined, and hence there is no immediate transformation of the basic model into a model where each point represents one definition. Since this objective is worth pursuing, it is therefore necessary to develop criteria or rules according to which the desired transformation can be made.

In the application of rules that may be developed, it will be convenient to make use of a model intermediate between the basic one and the one with points as definitions. For this purpose, we can combine the two models by employing a trivial relation, x_iRx , which says that the i th definition of x is used to define x ; this holds for all definitions of x . The line reflecting xRy would remain in the model, so that the digraph

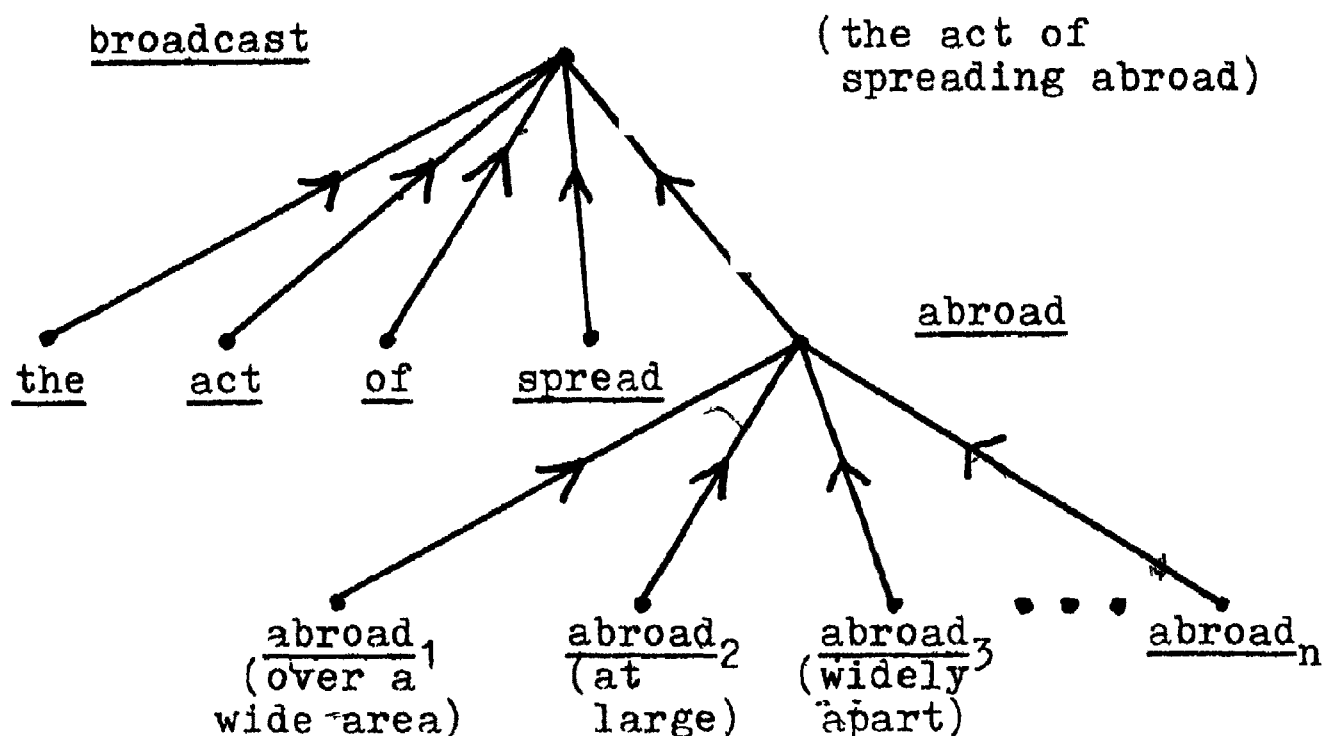


Figure 2. Subgraph of model with points representing both single and multiple definitions.

would show both $x_1 R x$ and $x R y_j$, and x would be a carrier, as illustrated in Figure 2. In this case, the unsubscripted abroad represents all the definitions of abroad (only some of which are shown). If and when suitable criteria establish, for example, that abroad₁, but not abroad₂, abroad₃, ..., fits the context of the token "abroad" in the definition of broadcast, it would then be possible to draw a line directly from abroad₁ to broadcast without the intermediation of the unsubscripted point abroad, thus eliminating paths from abroad₂, abroad₃, ... to broadcast.

This model thus includes the points of the basic model and adds points to represent each individual definition in the dictionary. The lines between these points ensure that no relation in the basic model is lost. As described in the example, it is necessary to develop rules according to which the points repre-

senting more than one definition can be eliminated or bypassed, so that the only relations, xRy , that remain are such that x and y are points which represent one definition.

It may happen during the application of rules that some lines to a carrier will be eliminated with more than one still remaining. In such a case, it will still be useful to modify the digraph as much as possible. For example, if xRy in the basic model, where x has m definitions and y has n , and xRy_j in the expanded model, then $x_1, \dots, x_m R y_j$. It may be that some criterion indicates that, say $x_1, x_2 R y_j$ but that $x_3, \dots, x_m \not R y_j$. When this occurs, we can create two points x_a and x_b such that $x_1, x_2 R x_a$ and $x_a R y_j$, and $x_3, \dots, x_m R x_b$, but with no line from x_b to y_j , as illustrated in Figure 3. The utility of this type of

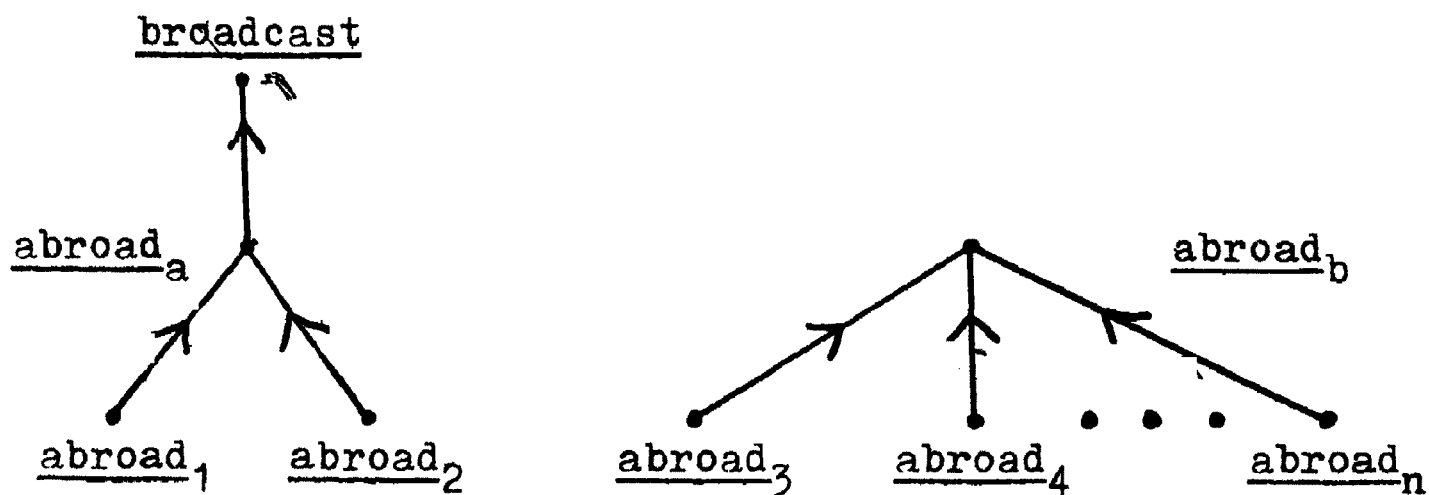


Figure 3. Subgraph of expanded model with grouping of definitions.

grouping will be demonstrated in Section 9. In any event, since many criteria will eventually be required in the elimination of points representing two or more definitions, this ability to group definitions is a necessary mechanism for modeling intermediate descriptions of the dictionary. (It should be noted

here that all such points will not be eliminated; those that remain will indicate an essential ambiguity in the dictionary; this is further discussed in Section 8.)

7. SEMANTIC, STRUCTURAL, AND SYNTACTIC PARSING OF DEFINITIONS

The basic and expanded models, exemplified in Figures 1, 2, and 3, do not portray any of the meaning of the dictionary, but rather indicate where particular relationships exist. In fact, these two models portray only the relation "is used to define" as if there is no other relation between definitions. This approach does not capture some very important elements that go to make up a definition.

Instead of being analyzed directly into its ultimate constituents, as in Figures 1 and 2, the definition, "the act of spreading abroad", should first be broken down into subphrases and then into its ultimate constituents, as in Figure 4, shown on the next page. A desirable property of the new points is that they have the syntactical structure of definitions. Thus, "the act" and "spreading abroad" have the form of noun definitions; "spread abroad" has the form of a verb definition; and "of spreading abroad" (not shown, but feasible under a different parsing) has the form of an adjective definition. This would eliminate such combinations as "act of" or "of the". The points representing phrase constituents of a definition thus have the form of definitions, but lack a label.

The absence or presence of a label seems to make no difference in understanding the definition represented. In fact,

digraph. (This is not to say that the words disperse, scatter, and distribute have the same meaning, but rather that in some instances these words can express the same concept.)

The definition, X, "the act of spreading abroad" is essentially an entity unto itself. The definitions of its component words have similar independence. However, like atoms in molecules, we need to identify those forces which hold the components together and which endow the whole with whatever characteristics it has. The definitions of the component words may require several words for their expression, but they are symbolized by one word in the definition X; even so the symbol and the definition both represent the same entity, which has certain characteristics enabling it to be acted upon by certain forces. These characteristics are the semantic, structural, and syntactic properties of definitions, and the forces are the production rules by which the entities (i.e. the component definitions or their symbols) are brought together. A definition may be viewed as the realization of such rules operating on the characteristics of other definitions. The Herculean task before us is to build a parsing system or recognition grammar which will articulate the characteristics to be attached to each definition and which will capture the production rules necessary to portray the relationships between definitions. The remainder of this section will present my ideas on how to approach this task.

The process which I have used for finding primitives entails showing that one definition is derived from another thereby excluding the former as a candidate for being primitive. Such a demonstration of a derivational relationship requires a parser. Each pattern which I observe between definitions helps to exclude further definitions and simultaneously becomes part of the parser. As a result, identification of the characteristics to be attached to each definition does not have to be accomplished all at once; as will become clear below, our purposes can be served as the components of the parser are delineated. Thus, success does not require full articulation of the parser before any parsing is initiated. The following represents general observations about the form of the parser as it has emerged thus far.

The first set of characteristics would result from the syntactic parsing of each definition. The purpose of this step would be simply to establish the syntactic pattern of each definition. The output of this step would be similar to that generated by Winograd (1972) in his parser. The 'dictionary' for the parser would be the very dictionary we are analyzing, although only the main entry, its inflectional forms, and its part-of-speech label would be used in this step. Ambiguous parsings and failures would be kicked out; the failures in particular, would provide an excellent source for refining the parser used by Winograd. Clearly, this step is not trivial, and it might even be argued that it is beyond the state-of-the-art.

However, by using a corpus as large as a dictionary and by kicking out failures and ambiguities, I believe that this step will significantly advance the state-of-the-art

The second set of characteristics would be determined from a semantic parsing of the definitions, that is, an attempt to identify the cases and semantic components present within each definition. For this study I have found the following distinction to be useful: A case is a semantic entity which is not intrinsic to the meaning of a word, e.g. that someone is an agent of an action, whereas a component is an intrinsic part of the meaning, e.g. a human being is animate. It is necessary to articulate recognition rules for determining that a particular case or semantic component is present. The little that has been done to develop such rules has been based primarily on syntactic structures or a priori assertions that a given case or component is present. Despite the recognized deficiencies of dictionaries, I believe that it is possible to bring much greater rigor to such rules with evidence gleaned directly from the definitions. For example, cut has a definition, "penetrate with an instrument"; this definition would be parsed as having the instrument case. (Note also that this definition makes the instrument case intrinsic to cut.) However, in most cases, it will be necessary to examine the definitions of the constituent words. For example, the verb knife has the definition, "cut with a knife"; although it is quite obvious in this instance that a knife is an instrument, rigor demands that we go to its

definitions where we find, "a simple instrument ...". A great deal of analysis may ultimately be required to discern the intrinsic characteristics to be attached to a definition, but I believe that many of these can come from the dictionary itself rather than from intuition.

Although the number of cases and components discussed in the literature is not very large, the number of ways in which they may be expressed, at least in English, is significantly larger. In addition, there is still a large amount of ambiguity, i.e., not every form specifically indicates the presence of a particular case. For example, a definition, "act with haste" does not indicate that "haste" is an instrument: rather, "with haste" expresses a manner of acting. Unraveling all these nuances requires a great deal of effort. However, it appears that a particularly good source of help in this endeavor might be found in the definitions of prepositions (which are used primarily to indicate sense relations).

Bennett 1975 found it possible to express the meaning of spatial and temporal prepositions (a high percentage of all prepositions) with only 23 components. However, in Webster's, 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 prepositional definitions contain information about the context surrounding

the preposition, particularly what sort of entities are related by the prepositions. My examination of verb definitions containing prepositions has led to the observation of many noticeable word patterns, i.e. collocations, which appear to be useful in the recognition of cases. For example, one definition of of states that its object indicates "something from which a person or thing is delivered". In examining verb definitions, there appears to be a distinct set of verbs with which this sense is used in the following frame "(transitive verb)(object) of (something)". The verbs that fit the slot are exemplified by free, clear, relieve, and rid. Thus, if this pattern appears, the object of the preposition can be assigned the meaning "something from which a person or thing is delivered". Through the use of prepositional definitions in this way, I have therefore been able to articulate some semantic recognition rules by which the sense or case of a noun phrase (the object of a preposition) can be identified. My use of this technique has barely begun, so that it is presently unclear whether this approach will suffice to disclose all the case information that we wish to identify with a semantic parser, but if not it will certainly make significant strides toward this objective.

Parsing of a definition according to the preceding notions is still not sufficient to identify the semantic components which should be attached to a main entry, since much of the semantic content is only present by virtue of the definition's constituent words. Thus, a complete rendering of a definition's

semantic content must be derived from the semantic characteristics of its constituents, in a recursive fashion, all the way down to the primitives. Although identification of these primitives is the primary goal of the approach being presented here, and hence, intrinsically incomplete until the analysis is completed, the set of semantic characteristics for a particular definition can be developed as we proceed toward our goal. To do this, it will be necessary to articulate rules which indicate how semantic characteristics may be transmitted from one definition to another. An example of such a rule is: If the noun X possesses the semantic component "animate", and if X is the core noun (i.e. genus) in definition y_f of the noun Y , then Y will also have the component "animate". Another example is: If a verb X has a definition x which has been parsed as having an instrument case, and X is the core verb of a definition y_j of Y , and y_j also has been parsed as having the instrument case, then the instrument in y_j is "a type of" the instrument in x_i . It will also be necessary to articulate other derivational (such as the application of a causative derivation to a state verb) and transformational (such as the application of a gerundial transformation to any verb) rules. This process of delineating how semantic characteristics are transmitted will at the same time give more meaning to the lines of the dictionary digraph than simply "is used to define".

The third, and final, set of characteristics that must be attached to a definition is a specification of the context that

must be present if that definition intended. The context restrictions may require that the definiendum must be used in a particular syntactical way, for example, as a transitive or intransitive verb. Usage restrictions may specify the presence of particular words such as particles or objects. For example, there is a distinct set of definitions for the idiom take out, which thus requires the presence of the particle "out" in addition to the verb. One definition of the transitive verb chuck requires the object "baseball". Other definitions may require a specific subject. Finally, there are semantic restrictions that may be discernible only from the definition itself. For example two definitions of the verb cheer are: "to give new hope to" and "lift from discouragement, dejection, or sadness to a more happy state"; if the second definition is intended, it seems necessary that the context indicate the prior state of discouragement, dejection, or sadness, since we cannot presume such a state, for someone might have been in a happy or non-sad state and simply received some new hope. In the absence of the necessary context, we would default to the first definition.

Thus far in my research, I have not devoted any effort toward developing procedures for prescribing the context based on the definition. I expect that initiation of this step will benefit from further results of the first two steps.

Although the parsing system outlined in this section may appear to be exceedingly complex, such an eventuality is not unexpected. The characteristics to be attached to each defini-

tion are not significantly different from those proposed by Fillmore 1971. It is also important to note that some of the goals of analyzing the contents of a dictionary are to reduce the amount of redundancy, to remove vicious circles, and to represent the meaning of a word in a more efficient way. Hopefully, this type of analysis would eventually lead to a substantial reduction in the size of a dictionary; the prospects for this are considered further in the next section.

8. THE ULTIMATE MODEL: POINTS AS CONCEPTS

At this juncture, it is necessary to ask whether the points of the digraph models sufficiently correspond to meaning as we wish it to be represented. In the two models described thus far, the analysis of a definition was deemed complete when the appropriate definitions of the constituent words had been identified. This situation is not entirely satisfactory, since, if a constituent word has more than one definition that applies, the definition being analyzed is subject to more than one interpretation and hence may be called ambiguous with respect to that constituent. For example, if the two definitions of abroad, "over a wide-area" and "at large", fit the definition of broadcast to yield either "the act of spreading over a wide area" or "the act of spreading at large", it is not legitimate to exclude one. This situation is only a reflection of the fact that natural language is almost always somewhat ambiguous. However, in accepting this fact, it is necessary that we incorporate it into our models.

Parts of the parsing system described in the last section will help to discriminate and select those definitions of a constituent word which fit a given context. As the parser is refined, the candidates for a particular context will be narrowed as described in Section 6, but many instances will remain where more than one definition fits the context. We might say that any point representing more than one definition thus constitutes an ambiguity. Viewed differently, we might also say that the context is not sufficient to distinguish among all the definitions of a word. In other words, we can 'blame' the ambiguity on the context..

We must expect that ambiguity will be present in the dictionary and deal with it on that basis. For purposes of illustration, let us say that abroad shown in Figure 4 is one such point. To remove such points from the d graph, we must make two points for the definition of broadcast, one representing "the act of spreading abroad₁" and one representing "the act of spreading abroad₂". These two points use the same words for expressing a definition and will be distinguishable only by the fact that their underlying definitions are different. Because of this situation, it is no longer valid to say that a point of the model represents a definition: rather, we will say that a point represents a "concept".

It is also possible that the concepts represented by two or more points can be shown to be equivalent. The concept, "the act of spreading abroad", has been shown to be equivalent to

"the act of spreading over a wide area". If the latter phraseology appears under some main entry, say distribution, then both it and the definition of broadcast would eventually be analyzed in the same way. We will say that both expressions may represent the same concept and hence are equivalent at least to this extent. (Since the other definitions of these words would be different they are not totally equivalent.) This concept will thus be represented by one point, labeled by either broadcast or distribution and equivalently verbalized as "the act of spreading abroad" or "the act of spreading over a wide area". This interpretation is a reflection of the fact that in ordinary speech a single concept may be verbalized in more than one way.

The observations in this section lead to the following description of the 'ultimate' model: The semantic content of a dictionary may be represented by means of a digraph in which (1) a point represents a distinct concept, which may be verbalized in more than one way and may have more than one label, and to which is appended a set of syntactic, semantic, and usage features, and (2) a line represents an instance of some one of a set of operators which act on the verbalizations or labels of a point according to the features of that point to yield the parametric values of another point. It should go without saying that the complete portrayal of a dictionary according to this model requires a considerable amount of further work; nonethe-

less, I believe that the model provides the appropriate framework for describing a dictionary.

9. PROCEDURES FOR FINDING THE PRIMITIVES

In Section 3, I stated that the model of a dictionary should permit the transformation of each definition into its primitive components. Based on the preceding descriptions, it is suggested that the full articulation of the ultimate model will satisfy this objective for the following reasons: (1) An elementary theorem in the theory of digraphs asserts that every digraph has a point basis, that is, a set of points from which every point in the digraph may be reached. Since points represent concepts in the ultimate model, it seems reasonable to assert that the point basis of its digraph represents the set of primitive concepts out of which all others in the dictionary may be formed. Based on the characteristics of the points in that model, it is possible (and perhaps even necessary) that each primitive concept would be verbalized in several ways and symbolized in several ways (as will be shown below) (2) Since the digraph has a finite number of points and lines, the sets of primitive concepts and operators are also finite.

It only remains to find the primitive concepts; this will be done by applying rules, based on the models and the parsing system, to identify words and definitions which cannot be primitives. Essentially, the assertion that a word or definition is non-primitive requires a showing that it is derived from a more primitive concept and that a primitive cannot be derived from

it. These non-primitives can be set aside and their full syntactic and semantic characterization can be accomplished after the primitives have been identified. Although no primitives have yet been identified (since the described procedures have not been fully applied), their form and nature will be delineated.

To demonstrate the validity of my approach, I have been applying rules developed thus far to the set of verbs in Webster's Third New International Dictionary (20,000 verbs and their 111,000 definitions). This set was chosen because of their importance (cf. Chafe 1970) and the (bare) feasibility of coping with them manually (although it may be another 3-4 years before I am finished, at my current rate of progress). I have attempted to formulate my procedures with some rigor, keeping in mind the ultimate necessity of computerization. I have developed some detailed specifications for some of my procedures, envisioning the use of computer tapes developed by Olney, but have not completed these since I do not presently have access to a computer.

Despite the focus on verbs, it will become clear that words from other parts of speech are inextricably involved in the analysis. Also, the rules that are presented can, for the most part, be applied to other parts of speech. Notwithstanding the fact that the meaning of many verbs is derived in part from nouns and adjectives, I believe that each verb definition also contains a primitive verb constituent.

Each verb definition consists of a core verb (obligatory) and some differentiae (optional). (The definitions of other parts of speech have a similar structure, i.e. a core unit from the same part of speech and some differentiae.) The subgraph of the total dictionary digraph formed by core verbs accords fully with the models described in Sections 4, 5, and 7. Therefore, any rules developed on the basis of those models will apply equally to the verb subgraph. We need only keep in mind that the differentiae come from other parts of speech and become embodied in the core verb. This is how the verb cut comes to have the instrument case intrinsically. To begin the analysis, we will let E represent the set of those verb definitions which have been identified as non-primitive; initially, this set is empty.

Rule 1. If a verb main entry is not used as the core unit of any verb definition in the dictionary, then all its definitions may be placed in E. (This rule applies to points of the basic model which have outdegree 0, i.e. no outgoing lines.) Since no points can be reached from such a verb, it cannot be

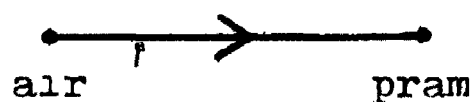


Figure 5. Basic model, verb subgraph example subject to Rule 1.

primitive. In Figure 5, the point labeled by pram represents the definition "to air (as a child) in or as if in a baby carriage"; since pram is the core unit for no definition in the

dictionary, all its definitions may be excluded as non-primitive. In W3, this rule applies to approximately 13,800 verbs out of 20,000; the number of definitions in the verbs excluded is not known.

Rule 2. If a verb main entry is used only as the core unit of definitions already placed in E, then all its definitions may also be placed in E. (This rule applies to points of the basic model with positive outdegree. The uses of such verbs as core units follow definitional paths that dead-end; hence, they cannot be primitive. Figure 6 shows a portion of the dictionary

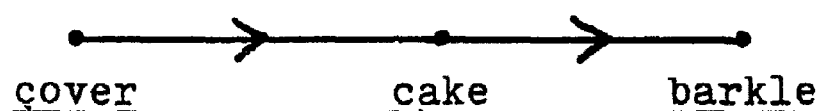


Figure 6. Basic model, verb subgraph example subject to Rule 2.

digraph where the verb cake defines only barkle, which in turn is not used to define any verb. Thus, the definitions of cake may be included in E after the definitions of barkle have been entered. In W3, this rule applies to approximately 1400 of the 6200 verbs that remained after application of Rule 1.

Rule 3. If the verbs forming a strong component are not used as core units in any definitions except those in the strong component or in definitions of verbs already placed in E by Rules 1, 2, or 3, then the definitions of all verbs in the strong component may be placed in E. (This rule applies to points of the basic model which constitute a strong component, i.e. a maximal set of points such that for every two points, u

and v , there are paths from u to v and from v to u . This rule does not apply when the strong component consists of all points not yet placed in E .) A strong component consisting of the verbs aerate, aerify, air, and ventilate is shown in Figure 7.

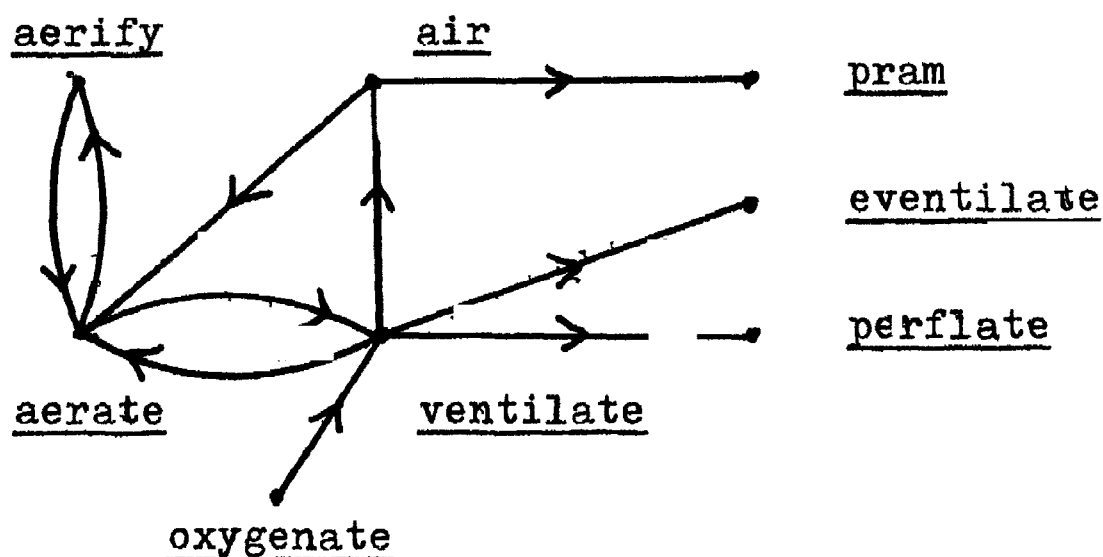


Figure 7. Basic model, verb subgraph example subject to Rule 3.

Except for oxygenate, the other verbs defining the set constituting the strong component are not shown. Since it is possible to start at any of the four and follow a path to any other of the four, there is no real generic hierarchy among them. It is possible to emerge from the strong component and follow paths to pram, eventilate and perflate, to which, however, Rule 1 applies. If we follow a definitional path that leads into this strong component, we can never get out again or if we do we will only dead-end. Hence, the definitions of all the verbs in the strong component are not primitive and may be placed in E . In W_3 , this rule applies to approximately 150 of the 4800 remaining after the application of Rule 2. Actually, Rules 2 and 3 may be applied in tandem; based on those placed in E . Thus,

after Rule 3 places the definitions of aerate, aerify, air, and ventilate in E, it so happens that Rule 2 then applies to the definitions of oxygenate.

After Rules 1, 2, and 3 are applied to the digraph of the basic model, the remaining points constitute a strong component of approximately 4500 points. This differs from those to which Rule 3 applies in that there would be no points left if we placed all its points in E. This final strong component is the basis set of the basic model, that is, any point of the basic model (i.e. any main entry in the dictionary) may be reached from any point in the final strong component (but not conversely).

At this juncture, we can proceed no further with the basic model alone; it is necessary to expand the points of the final strong component into two or more points each representing a subset of the definitions represented by the original point, as previously shown in Figure 3. In part, this can be accomplished by identifying individual definitions which are not used.

Rule 4. If any definition can be shown to be not used as the sense of any core unit (or only those already in E), it may be placed in E. This rule is essentially a restatement of Rule 1 for individual definitions and includes the following two subrules, among others not presented.

Rule 4a. If all the remaining uses of a verb are transitive (intransitive), then its intransitive (transitive) definitions are not used and may be placed in E. The expansion of a

point into transitive and intransitive uses is a good example of how the points of the basic model are transformed into points of the expanded model.

Rule 4b. If a definition is marked by a status label (e.g. archaic or obsolete), a subject label, or a subject guide phrase, it may be placed in E. Lexicographers creating W3 were instructed not to use such marked definitions in defining any other word..

Other rules have been developed in an attempt to identify the specific sense of the core verb, or those senses of a verb which have not been used in defining other verbs, but are not presented here. However, there are too many instances where the differentiae of a definition do not provide sufficient context to exclude all but one sense (for example, many senses of move fit into a definition phrased "move quickly"). In order to continue toward the primitives, we must shift gears slightly and ask whether a definition can be characterized as "complex", that is, derived from more primitive elements. For example, one definition of make is "cause to be", which can be labeled as complex because it consists of a causative component and a state component, each of which is more primitive by itself than "cause to be".

The importance of the notion of a complex definition becomes evident when we try to visualize how a primitive concept will be identified. To understand this, we must consider some further properties of the digraph. After the application of

rule and any subsequent rule), the remaining graph is a final strong component. (Recall that in a strong component, for each two points, u and v , there is a path from u to v and one from v to u .) Assuming that each point represents a concept (as in the ultimate model), the fact that two concepts are in the same strong component means that they are equivalent. In more traditional terms, what we have is a definitional vicious circle, that is, a definitional chain which adds nothing to our understanding of the meanings involved.

Using the digraph of the final strong component, we can identify (and examine one by one) all putative definitional cycles or vicious circles; these will fall into three classes. The first class will consist of improper cycles, which can be removed by determining that one point is more complex (and hence not equivalent to the definition from which it is derived) Further rules for characterizing a definition as complex are given below. The second class of cycles will be real vicious circles, which fortunately can be removed, but only under certain conditions. For example, one definition of jockey is "maneuver for advantage", while one definition of maneuver is "jockey for position"; these two definitions constitute a vicious circle. In order to remove it, there must be some other definition of either verb which constitutes its meaning; in this case, it is found under maneuver, specifically, "shift tactics". Thus, in order to remove a vicious circle, we must find some way out. If we cannot, we have the third class of

cycles; this class will comprise the set of basic concepts. If there had been no way out for the example of jockey and maneuver, we would have said that no meaning was conveyed by either verb, but rather that the meaning was established by use. This third set of cycles is what is sought by the procedures described in this paper.

As mentioned above, the crux of the analysis after the application of Rules 1 to 4 is the identification of complex concepts. Essentially this entails a showing that, for any definition y_j of verb Y, with Y as the core verb of definition x_j of verb X, the differentiae of x_j make y_j generic to x_j . For example, all transitive definitions of cut would be generic to a definition in which "cut" is used with an object, even without narrowing down to one definition. The general rule may now be stated.

Rule 5. If any definition is identified as complex, it may be placed in E. The net effect of this rule is to break one or more putative cycles of equivalent definitions or concepts, enabling them to be transformed into a strict hierarchical order which will eventually be subject to Rule 4. Thus, the complex definition and all definitions that can be shown to be derived therefrom can be placed in E, because they cannot be part of a primitive cycle.

Rule 5 is implemented only by very specific recognition rules, which are essentially part of the parser. The specific rules entail a showing that some component has been added in

the differentiae of a definition that is not present in the meanings of its core verb. For example, the "manner" component is not intrinsic to the meaning of the verb move; therefore, when a definition has the core verb "move" with an adverb of manner, it can be marked as complex. In establishing a component as non-intrinsic, it is necessary to articulate rules for recognizing the presence of the "manner" component (such as a phrase in a _____ manner" or an "-ly" word with a definition "in a _____ manner") and then to determine if that component is present in any definitions of a particular verb. If not, then the verb can be labeled as complex whenever it is used as the core verb in a definition with differentiae that fit the recognition rule. In addition to move, I have determined that, for the manner component, the verbs act, perform, utter, speak, express, behave, and many others follow the rule. Table 1, on the next page, identifies some specific components, a brief description of how they are recognized, some of the verbs to which the particular rule applies, and an example of a definition labeled as complex by the rule and hence placed in E.

If a definition has a core verb whose applicable sense is one which has been marked as complex, it too can be so marked, since it is derived from a complex definition. For example, all definitions of the form "make adjective", i.e. with an adjective complement, are derived from the definition of make, "cause to be or become" and hence can be marked as complex. In addition, if all definitions of a verb have been marked as

Table 1Recognition Rules for Semantic Components

<u>Name of Component</u>	<u>Recognition Rule</u>	<u>Applicable Verbs</u>	<u>Examples of Definitions</u>
1.Aspect	Verb + Infinitive	cease, begin, strive, continue	<u>commence vi 2,</u> "begin to be"
2.Causative	Causative verb + Infinitive	cause, force, compel, induce	<u>confront vt 2a,</u> "compel (a person) to face, take ac- count of. or endure" <u>make vt 10a,</u> "cause to be or become"
3.Instrument	Verb + "with" + noun defined as instrument, device, etc.	apply, fasten, cut, beat	<u>knife vt 2a,</u> "cut with a knife"
4.Means (Process)	Verb + "by" + Gerund	make, prepare, form, shape	<u>draw vt 4e4,</u> "shape (glass) by drawing mol- ten glass from the furnace over a series of automatic rollers"
5.State Entry	Verb + "into" + noun defined as "the state of ..."	bring, put, throw, fall	<u>disorder vi,</u> "fall into confusion"
6.Deliverance	Verb + "of" or "from" + noun	free, relieve, rid, empty	<u>clear vt 2g2,</u> "rid (the throat) of phlegm"

complex then all definitions in which it appears as a core verb can be similarly marked and placed in E.

Through the development and application of further parsing rules under Rule 5, I am hopeful that I will eventually arrive at the set of primitive verb concepts (i.e. cycles or vicious circles with no way out). I have already reduced the number of verbs from 20,000 to less than 4,000. This number would be much lower, but for the fact that I am applying the rules manually and I must exercise time-consuming care to ensure correctness.

After the primitive concepts have been identified, it will be necessary to go back to all the definitions that were set aside in the process of finding the primitives, so that their semantic characteristics can be articulated. I fully expect that the parsing system which will have been developed will be able to accomplish much of this task. I also expect that the parsing system will have equal applicability as a general parser capable of formally characterizing ordinary discourse in a canonical form. Of course, verification of this expectation will have to await a full presentation of the parser.

10. RELATIONSHIP TO EFFORTS TO REPRESENT KNOWLEDGE IN FRAMES

The process which has been outlined in the preceding sections is closely akin to current efforts to represent knowledge in frames. (Cf. Winston 1977 for an elementary presentation of this notion.) Briefly, a frame consists of a fixed set of arguments, some of which may be specifically related to others, and some of which may have specific values. A frame is intended to

represent a stereotyped situation, with the arguments identifying the various attributes which the situation always possesses. In terms of case grammar, for example, a movement frame will contain arguments or slots for an agent, an instrument, and a destination. By tying frames together in specific relationships, we can build larger and larger frames to represent more and more knowledge, perhaps constructing a series of events, an inference structure, or a description of a scene.

Before building these large structures, it is necessary to represent very small pieces of knowledge. Heretofore, this has been done by postulating the components of frames to represent such things as actions and state changes. But this can be accomplished on a more rigorous basis. For example, if we first locate all definitions using "move" as its core verb and then identify all the case structures in which it is used, we will have a generalized frame which characterizes most if not all of the possible uses of "move". (This approach is currently being followed by Simmons 1977.) Each definition in which "move" is used could then be represented by the generalized frame with some of its slots filled. This process can be followed for any word for which we wish to develop a frame.

If, in addition, we analyzed the definitions of move, we will find that they, in turn, represent instantiations of still other frames, which will be even more generalized than those developed for the uses of "move". The difference between the frames representing the definitions of move and those represent-

ing the uses of "move" is that the latter are the same as the former with some slots filled. Within the bounds of the ambiguity present in the dictionary, this slot-filling will identify which definition of move are employed in which uses of "move". It seems to me that this is nothing more than the process which has already been described using a graph-theoretic approach, except that the generalized frame for each verb will not be carried along through each step. Moreover, since the semantic parsing system which has been described will be based largely on the relationships derived from the definitions of prepositions, and these comprise most of the case relationships, the parsing system will effectively circumscribe the permissible elements (i.e. slots) which can be present, given any particular context. Thus, although the phraseology is different, the effect is the same.

If there is an essential equivalence between these two approaches, then, since frames purport to represent knowledge, the process described, if successful, will result in an articulation of whatever knowledge is contained in a dictionary. What this implies is that the lexicon contains a great deal of knowledge about the world and not just information which will enable us to understand such knowledge.

Frames provide a great deal of insight to the approach which has been described here, but the reverse also seems to hold true. If the semantic content of each definition can be captured, then it may be possible to articulate the frame for

any utterance by combining the characteristics of the definitions of its constituent words within what is permitted by the parsing system.

11. FINAL REMARKS

In Section 1, I described some limitations of this paper and my research. This paper suffers from a lack of sufficient detail to enable a reader or researcher to replicate what I have done or to take the next steps of computerizing the procedures which I have developed. I will provide further details on the specific steps I have followed in reducing the set of verbs from 20,000 to 4,000 to anyone requesting. With respect to computer specifications, I have prepared some, but stopped because I have no access to a computer. However, if any researcher is interested in pursuing this (or setting graduate students to work), I am prepared to develop the necessary specifications and to work hand-in-hand for the further advancement and refinement of this methodology.

I also indicated in Section 1 that my research presently shows no final results and that I do not even know how much further effort will be necessary to explicate the parsing system which has been described. Clearly, there are great distances yet to be covered toward a goal of being capable of transforming ordinary discourse into a canonical form. I believe that characterization of the contents of an ordinary dictionary is an essential step in attaining this goal, and I am hopeful that my approach can be used to develop such a characterization.

If it seems worthwhile to pursue this approach, despite the limitations, I believe the best way to do so would be to establish a single computer-based repository for a dictionary, preferably W3, with on-line access to researchers across the country, and to build the parser and definitional characterizations piece by piece. (I have noted how the parsing system which I have described can be built incrementally.) The magnitude of this effort precludes much progress by individual researchers. Olney tried to do something similar with the collegiate dictionary based on W3, but by distributing bulky computer tapes. He was unfortunately premature; it may be that now is the time to try again.

REFERENCES

1. Bennett, David C. Spatial and Temporal Uses of English Prepositions: An Essay in Stratificational Semantics (Longman Linguistics Library, 17), Longman, New York 1975.
2. Chafe, Wallace L. Meaning and the Structure of Language, University of Chicago Press, Chicago, 1970.
3. Encyclopaedia Britannica. Webster's Third New International Dictionary, Chicago, 1966.
4. Fillmore, Charles J. The case for case. Bach, E. and Harms, R.T. (Ed.) Universals in Linguistic Theory, Holt, Rinehart and Winston, New York, 1968, 1-88.
5. Fillmore, Charles J. Types of lexical information. Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology, Cambridge University Press, Cambridge, 1971, 370-392.
6. Harary, F., Norman, R.Z., and Cartwright, D. Structural Models: An Introduction to the Theory of Directed Graphs, John Wiley and Sons, Inc., New York, 1965.
7. Jackendoff, R. Morphological and Semantic Regularities in the Lexicon, Indiana University Linguistics Club, Bloomington, Indiana, 1974.
8. Katz, J.J. and Fodor, J.A. The structure of a semantic theory. Language 39(1963), 170-210.
9. Nida, E.A. Componential Analysis of Meaning, Mouton, The Hague, 1975.
10. Olney, J., Revard, C., and Ziff, P. Toward the Development of Computational Aids for Obtaining a Formal Semantic Description of English, SDC Document SP-2766/001/00, System Development Corp., Santa Monica, 1968.
11. Quillian, M.R. Semantic memory. Minsky, M.(Ed.) Semantic Information Processing, MIT Press, Cambridge, 1968, 216-270.
12. Schank, R.C. Conceptual dependency: A theory of natural language understanding. Cognitive Psychology 3(1972), 552-631
13. Simmons, R.F. and Slocum, J. Generating English discourse from semantic networks. Communications of the ACM 15(1972) 891-905.

14. Simmons, R.F. and Amsler, R.A. Modeling Dictionary Data, University of Texas Department of Computer Sciences, Austin, 1975.
15. Simmons, R.F. and Lehmann, W.P. A Proposal to Develop a Computational Methodology for Deriving Natural Language Semantic Structures via Analysis of Machine-Readable Dictionaries, University of Texas, Austin, 1976. (Research proposal submitted to the National Science Foundation, September 28, 1976)
16. Winograd, T. Understanding natural language. Cognitive Psychology 3(1972), 1-191.
17. Winston, P.H. Artificial Intelligence, Addison-Wesley, Reading, Mass., 1977.