American Journal of Computational Linguistics *Microfiche 50 : 2*

AN ORGANIZATION FOR A DICTIONARY OF WORD SENSES

DICK H. FREDERICKSEN
IBM THOMAS J. WATSON RESEARCH CENTER
YORKTWON HEIGHTS, NEW YORK 10598

ABSTRACT: This paper describes a lexical organization in which "senses" are represented in their own right, along with "words" and "phrases", by distinct data items. The objective of the scheme is to facilitate recognition and employment of synonyms and stock phrases by programs which process natural language. Besides presenting the proposed organization, the paper characterizes the lexical "senses" which result.

1. Introduction.

This paper describes an internal lexical organization which is particularly designed to capture the facts about synonymy. Besides recording the inclusion of each word in one or more synonym sets (identified with its various "senses"), the scheme attempts to distribute attributes perspicuously between "senses", "wordings", and the intersections of the two. In addition, there is provision to record multi-word idioms, stock phrases, and the like, and to include these as elements in synonym sets when appropriate.

Briefly, "senses" are represented in their own right, along with "words" and "phrases", by distinct data items. Each word or phrase is associated with a list of the "senses" which it can express; conversely, each "sense" is associated with a list of "alternative wo lings" Additionally, each word is associated with a list of phrases in which it occurs.

Grammatical category, features, selection restrictions, and the like are applicable at three different levels: to words or phrases as such, to "senscs" as such, or to particular usages of words or phrases (equivalently, to particular wordings of "senses").

This lexical organization has been implemented at IBM Research, Yorktown Heights, N.Y., by a program -- not to be described here -- which builds such dictionaries in a very compact form, giving interactive assistance to the person making the entries. (For example, the program points out the possibility of merging "senses" whenever their wordings overlap and their attributes are compatible, and merges them if so directed.) There are suitable facilities for saving the results, retrieving them in various ways, and for altering such things as schemes of classification without scrapping previously prepared work.

The ultimate intent is that the "dictionary of senses" should serve as the lexical component in a natural language fact-retrieval system. Pending its incorporation in that role, it will be used to amass and organize information on the semantic relations among words and phrases.

The balance of this paper comes in two sections:

Section 2 presents the proposed lexical data structures, and suggests how they are to be used. Included is a sketch of how various types of grammatical and semantic "attributes" fit into the scheme.

Section 3 discusses the character of the "senses" encoded in the resulting dictionary. Reasons are advanced for regarding lexical "senses" as something far short of semantic primitives. At the same time, synonym sets are defended against the view that "true paraphrases are rare or nonexistent".

2. The Internal Representation.

It will be our purpose in this section to say just enough about internal representation to lay bare the organizing principles of the lexicon. The focus is on architecture and motivations; details of field layouts, internal codes, etc. are not at issue here.

To make the discussion concrete, suppose we are interested in the senses of the word "change" Assuming that none of the words are unfamiliar, the following should put us in mind of two senses:

change: 1. *v* alter;
2. *n* small coin.

This, of course, is just a dictionary entry in the traditional format (though with synonyms offered in lieu of definitions).

On the other hand, we might approach the same information from a different direction: starting with the two concepts, we might seek words to express them. It is difficult to picture this latter situation without assigning artificial labels to the concepts. Call them concepts 1 and 2, and suppose for a moment that there were a practical way to look the concepts up *(without* having thought of either word for either concept). Then the information to be retrieved might be envisioned this way:

1. *v* change, alter
2. *n* change, small coin

It is this duality of viewpoint -- that words have senses, while senses have wordings -- that our lexical representation must reflect.

The starting point, then, is that words, phrases, and "senses" are separately represented. There are three principal types of data item, plus a standard connector:

1. A *"Key Data Item"* (KDI) represents a single word.

2. A *"Phrase Data Item"* (PDI) represents a string of two or more words which are to serve as a unit in some context.

3. A *"Sense Data Item"* (SDI) represents one distinct sense common to a set of words and/or phrases. In general, a word or phrase may be usable in more than one sense, while a given sense may have alternative (synonymous) wordings. Both these types of variability are recorded making use of the next data item:

4. A *"Sense Link Element"* (SLE) is a connective item, to be explained shortly.

Three principal fields will engage our attention in each type of data item. Fig. 1 summarizes the fields for each type.

"Alternative Senses" Link	Global Attributes	"Phrase Involvements" Link

KDI (Key Data Item)

"Alternative Senses" Link	Global Attributes	"Component Word" Links

PDI (Phrase Data Item)

"Alternative Wordings" Link	Global Attributes	"Sense Chain" Link

SDI (Sense Data Item)

"Alternative Senses" Link	Local Attributes	"Alternative Wordings" Link

SLE (Sense Link Element)

Fig. 1
Schematic of Data Items, with Principal Contents

Each KDI (Key Data Item) or PDI (Phrase Data Item) contains an "alternative senses" link -- a pointer to the first SLE (Sense Link Element) in a chain of SLE's which represent the various senses of the word or phrase. The SLE's are chained via their own "alternative senses" links, and the final member points back to the KDI or PDI. Thus, we shall speak of such a chain as a ring

specifically, an "alternative senses ring". If no senses are on record for a particular word or phrase, the "alternative senses" link in the KDI or PDI is self-referent.

Reciprocally, each SDI (Sense Data Item) contains an "alternative wordings" link. This leads to a chain of SLE's which represent more-or-less synonymous wordings that express the sense. These SLE's are chained through their own "alternative wordings" links, and again the chain is closed into a ring -- this time beginning and ending with the SDI.

The structure that is shaping up may now be seen in Fig. 2. *The crucial point is that each SLE represents the intersection between an "alternative senses" ring and an "alternative wordings" ring.* From the standpoint of the word or phrase, it represents a particular sense; from the standpoint of the sense, it represents a particular wording.

Starting from a KDI or PDI, one gets to the SDI for a particular sense by advancing along the "alternative senses" ring to the relevant SLE, then detouring along the ring which connects the latter to the SDI (as one of the SDI's "alternative wordings"). Starting from an SDI, one gets to a particular wording by the reverse process. Since each "alternative senses" ring contains exactly one KDI or PDI, while each "alternative wordings" ring contains exactly one SDI, each SLE is tied to exactly one sense of one word or phrase. (Equivalently, it is tied to one wording of one sense.)

The next point of interest is that "attribute" fields are present in all four types of data item -- even in the connectors (SLE's). The attributes which may be recorded in each, however, come from different bags.

To begin with, the attributes found in an SDI characterize all the wordings of a given sense whenever the wordings are used in that sense. In Fig. 2, for example, sense "1" should be marked as a "verb" sense, while sense "2" is a "noun". One would not wish to record the attribute "verb" in the KDI for the word "change", for the KDI represents facts about the word itself, irrespective of sense, and "verb" does not hold for all uses of the word "change". On the other hand, "verb" does characterize all wordings of sense "1", whenever they're being employed to express that sense. It would furthermore apply to any additional wordings which we might think of, such as "modify", provided they are really used in a synonymous way.

As a matter of fact, it turns out that the traditional parts of speech -- noun, verb, adjective, preposition, etc. -- fit best in this scheme as global attributes of senses, recorded in the SDI's.

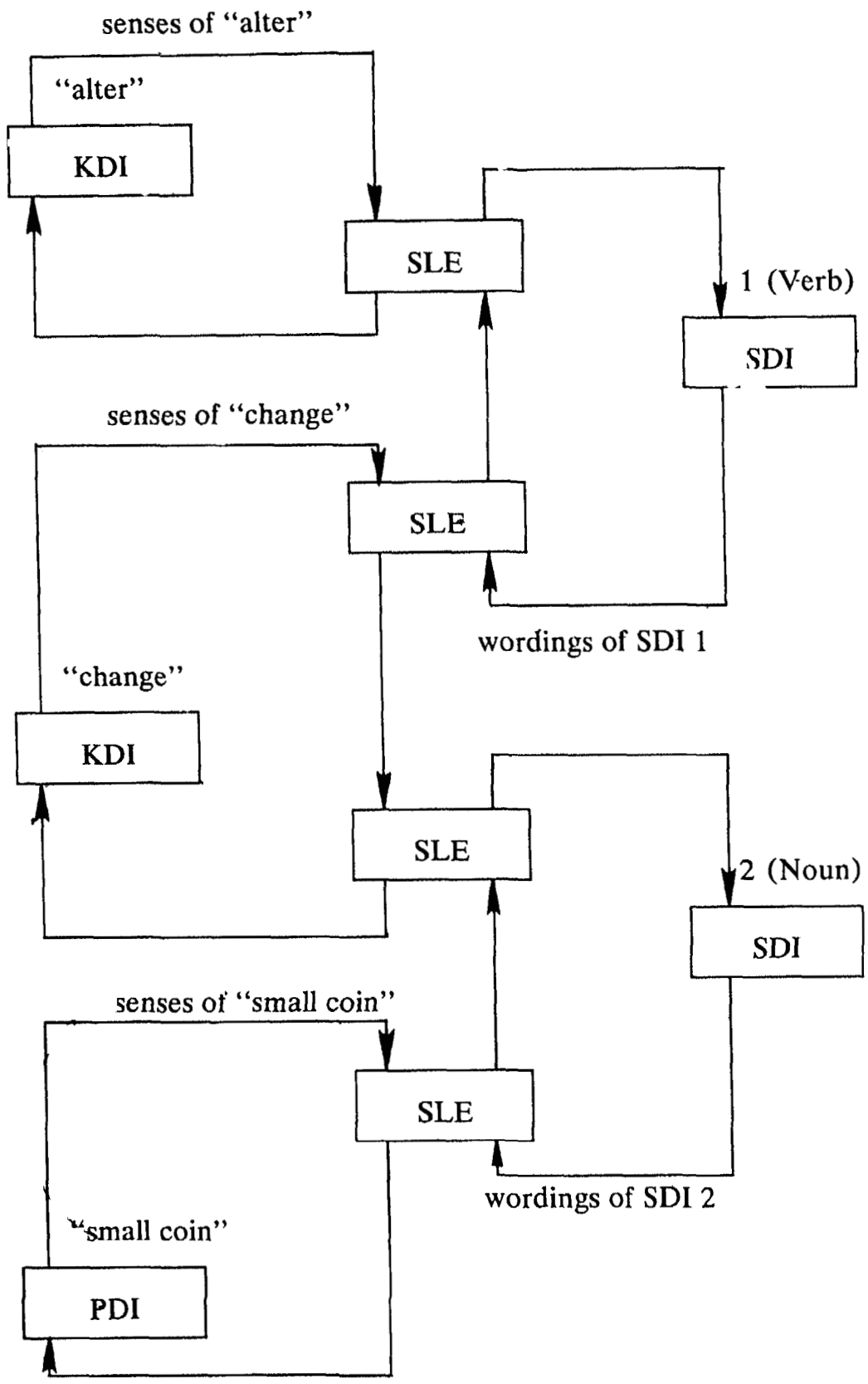

Fig. 2
"Alternative Senses" and "Alternative Wordings" Rings

(The first sense has two wordings: "alter" and "change". The second sense has wordings "change" and "small coin". Two senses are recorded for "change", and one sense each for "alter" and "small coin".)

A different sort of attribute may be recorded in a KDI, as a global feature of the word itself. For example, we may note of the word "change" that it is "regularly conjugated". That is, when used

as a verb, it forms the third person singular by adding "s", and both past and past participle by adding "ed" To be sure, this "global" attribute applies only to the "verb" senses of "change"; but a moment's reflection will confirm that "change" has more than one "verb" sense, and the regularity of its conjugation is common to all of them. Thus, it is useful to note this regularity as an attribute of the word itself. (Contrast this with the behavior of the word "can", which is regular when it means "to pack in cans", but irregular when it means "is able to".)

Various other attributes suggest themselves as global characterizers of the words themselves, to be recorded in the KDI's. For example, one might wish to note of "change" that it drops its final "e" when adding "ing" (this is the normal rule) but of "singe" that it doesn't.

Still other attributes are appropriate when characterizing multi-word units (in PDI's). A string of words whose meaning is not evident from the mere juxtaposition of its constituents (such as "give up") may be classified as an "idiom". A string of words whose meaning could be figured out from the meanings of its constituents, but which occurs with enough frequency to warrant inclusion in the dictionary, might be classed as a "stock phrase". (Example: "drop dead".) A string like "perform in a subordinate role", which one would not normally expect to encounter in its own right, might be classed as a "definition" (for a certain sense of the word "accompany", difficult to reword except with a definition).

Perhaps the most unexpected site for recording attributes is in the connective elements (SLE's). These are the logical place, though, to note features that apply to a specific sense of a word, without being global to either the sense or the word. Consider the following four sentences:

On the way to the office, he stopped daydreaming.
On the way to the office, he ceased daydreaming.
On the way to the office, he ceased to daydream.

versus:

On the way to the office, he stopped to daydream.

Suppose we choose to view this as a restriction upon the (surface) object of the verb: "stop", when applied to an action, must take a gerund as its object; "cease" can take either a gerund or an infinitive. (It wouldn't affect the point being made if we said that "stop" inhibits a certain grammatical transformation en route to surface structure, while "cease" permits it.)

Now, we wouldn't want to mark "gerund object only" as a global attribute of the sense, for we have just shown that "cease" and "stop", two wordings of the sense, differ with respect to this restriction. On the other hand, it doesn't belong among the global attributes of the word "stop" as such, for "stop" has other verb senses, even transitive ones, to which the restriction is completely inapplicable. (Consider "stop a hole in the dike", "stop a catastrophe", etc.) That leaves the alternative we are suggesting: treat the restriction as an attribute of one particular usage of the word (equivalently, one particular wording of the sense).

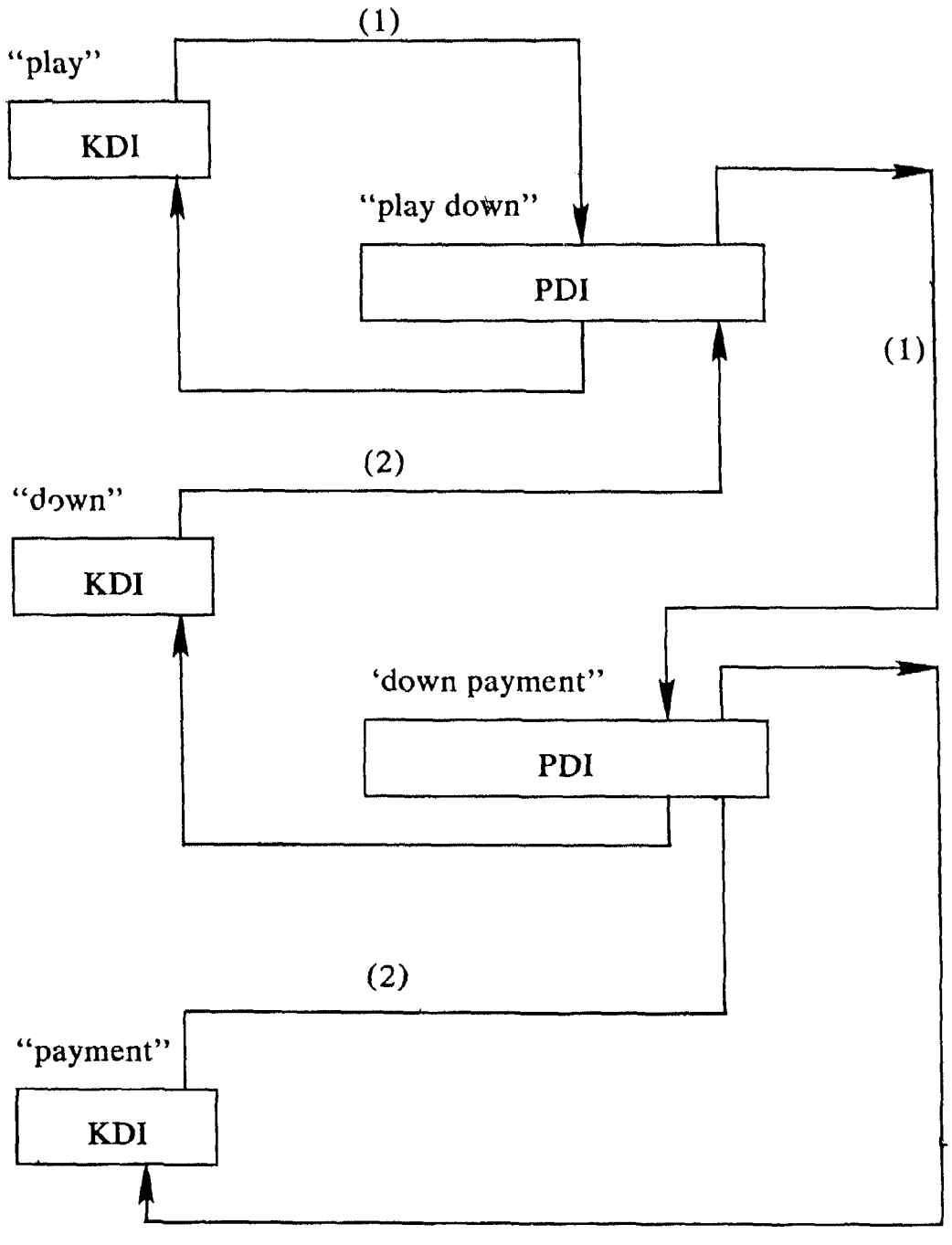

Fig. 3
"Phrase Involvement" Rings

(Where numbers are shown on connecting links, they indicate the position of the word in the phrase which is linked to.)

Besides having senses, individual words are involved in phrases, and this fact is also represented in our data structure. Fig. 3 shows the plan of attack. In the KDI for each word, there is a link connecting it to the PDI for the first phrase in which the word is known to occur, together with a number designating the position of the word (1st, 2nd, 3rd, etc.) in that phrase. In the PDI itself, there is a continuation link for each word of the phrase, together with its number in the next phrase. In the final PDI involving a given word, the link for that word points back to the KDI. Thus, independent of its "alternative senses" ring, each KDI may have a "phrase involvements" ring.

This structure makes it possible to retrieve all the idioms, stock phrases, definitions, etc., in which a given word has made its appearance, anywhere in the dictionary. As the same structure is used to encode every multi-word unit, no occurrence of a word is ever lost sight of, and a phrase can be looked up via any of its constituent words.

Of the fields to which Fig. 1 calls attention, we have discussed all but one. In the SDI for each "sense", there is a "sense chain" link field. This links the SDI to its successor in a global chain of "senses". Using this chain, it is possible to make an exhaustive, non-duplicative list of all the "senses" recorded in the dictionary. The listing program has only to proceed down the chain, retrieve from each SDI its attributes, decode them, then chase around the "alternative wordings' ring of the SDI and list the wordings alongside the attributes.

One more feature of the internal representation deserves mention: the data items for words occur as "leaves" in a lexical tree (Fig. 4). That is, the KDI for a word can be looked up letter-by-letter, following a chain of pointers that correspond to successive letters. The chain ends at a KDI after following a substring sufficient to distinguish the word from the nearest thing like it in the dictionary. The lexical tree has the advantage that words can be looked up either at random or in sequence.

Recapitulating, these are the essential features of the representation:

*1) "Senses" are represented separately from "wordings", and the mutual connections between them are made explicit in both directions.

*2) "Wordings" may be either single words or multi-word phrases. These are represented by distinct types of data item, and may be subject to distinct schemes of classification, but they are on the same footing with regard to "sense" connections. With each word is associated an exhaustive list of the phrases in which it occurs.

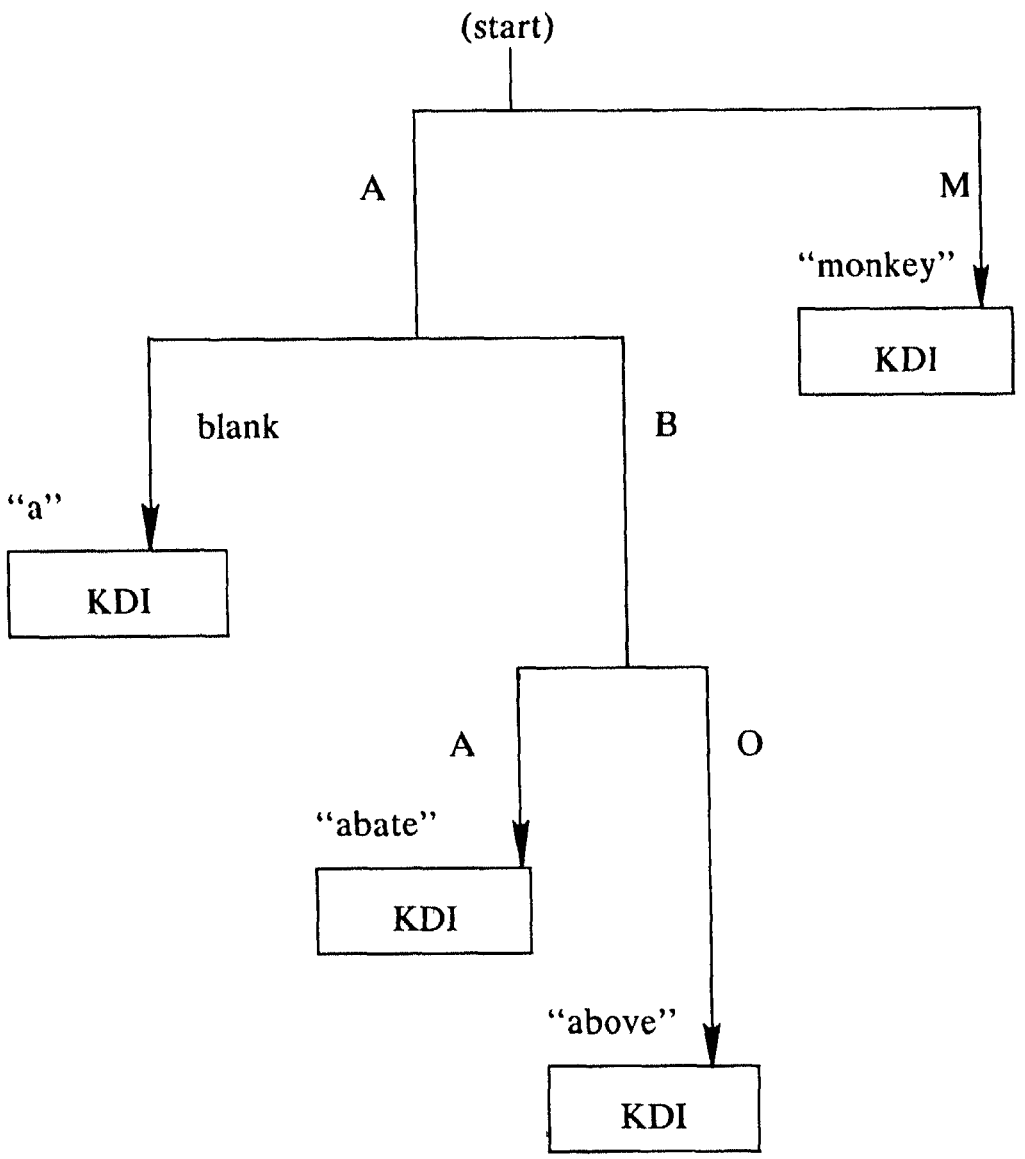

Fig. 4
Lexical Tree

(For a dictionary containing only the words "a", "above", "abate", and "monkey", this would be the full tree. The path to each word is only as long as needed to distinguish it from the neighbor with which it shares the longest leading substring.)

*3) Classifiers and features, drawn from appropriate sets, may be attributed separately to words, to phrases, to senses, or to particular senses of words or phrases (i.e., to particular wordings of senses).

*4) The data items which represent senses are globally chained, and may be exhaustively listed.

*5) The data items which represent words are accessible as "leaves" of a lexical tree; hence they may either be retrieved by lookup (in response to presentation of the words) or volunteered in alphabetical order.

Given a commitment to represent a lexicon as suggested by points *1 through *5 above, various implementations would be possible. Alternative implementations of individual points (though not of the scheme as a whole) have in fact been described by other writers. The lexical tree (*5), for example, is no great novelty: Sydney M Lamb and William H. Jacobsen describe implementation details of one such tree [5]. [10] also concerns a dictionary which uses this general style of organization for lookup. For that matter, the lexical tree is reminiscent of Feigenbaum's "discrimination tree." [1]

More interestingly, the separate representation of senses and wordings has been incorporated in other systems by R. F. Simmons ([11], [12]) and by Larry R. Harris [3]. This way of looking at matters led Harris to remark some of the same points that we have been stressing: that senses have alternative wordings just as words have alternative senses; that multi-word phrases might occur on the same footing as individual words in the expression of a sense; and (interestingly enough) that part-of-speech information really adheres to the "sense", not to the "word" Similarly, Simmons associates his "deep case" information with lexical nodes representing "wordsenses", while words themselves are treated as "print image" attributes of the wordsenses.

Harris's dictionary was only a minor component in a small-scale model of concept acquisition. No great number of either words or concepts was required to illustrate the principles at stake, so Harris programmed the dictionary as an array, with words represented by rows and "concepts" by columns. Elements of the array were merely frequencies, indicating the strength of association between each word and each concept.

Needless to say, for a full-scale vocabulary of words and concepts, such an array is mostly empty; nobody would dream of expanding it in that form. From a programming standpoint, the only thinkable choice is some form of list structure. Having decided in principle to use "some form of list structure", though, one might well ask: Why chains? Why rings? Why not just include in each Key Data Item a full list of pointers to the corresponding Sense Data Items, and vice-versa?

The answer is simply one of convenience. It's easier to handle insertions and deletions when they don't require the movement of expanded items to new quarters, or the provision of "overflow" pointers. It's easier to reclaim freed storage when deleted items come in a handful of standard

sizes. As for "rings", they eliminate the need for two-way pointers, since one can break into a ring at any point and follow it to its source.

It should be noted that to make rings an attractive representation, the details of the material being represented must cooperate. In particular, the rings must not become too long, or the processing required to follow them becomes excessive. It happens that "alternative senses" rings and "alternative wordings" rings are typically short rarely more than a dozen links per ring. "Phrase involvement" rings, on the other hand, can become spectacularly long, especially for words like "a" and "to". In practice, it's necessary to provide these rings with short-cut links.

Any of these programming details could be altered, however, without abandoning the essence of the scheme, which is given in points *1 through *5 above.

3. The Character of Lexical Senses.

Perhaps the first thing to get straight about the "senses" represented in this dictionary is what they are *not*. They are not "concepts"; they are not a set of "primitives" into which human experience an be decomposed. No conjecture is put forward here that any such collection of discrete, atomic concepts even exists, let alone that it might be finite.

Rather, the "senses" of the dictionary are in the nature of fuzzy equivalence sets among words. (This is only a metaphor; we shall do more and more violence to the technical notion of an "equivalence set" as we proceed.) Each "sense" groups a set of words which, in a set of appropriate contexts, might be used more or less interchangeably. That the equivalence sets are fuzzy, one can convince oneself with but the briefest immersion in the materials of the language -- trying to decide whether particular words belong in particular groups or justify the creation of new groups.

Consider, for example, the following set of words and phrases:

(abandon, give up, surrender, relinquish, let go, desert, leave, forsake, abdicate)

Clearly, there is a common theme that can run through all of these, given the right circumstances. It might be expressed as "reluctant parting from somebody or something". This can be seen by coupling the verbs with various possible objects:

(abandon, give up, surrender) a town to the enemy

(abandon, give up) all hope

(give up, relinquish) one's claim to an estate

(give up, let go) our entire stock at a loss

(abandon, desert, leave) one's wife and children

(desert, forsake) a friend in need

(give up, abdicate) the throne

(abandon, desert) an exhausted mine

(forsake, give up) all other, keeping thee only to her/him

(abandon, desert, leave) the area threatened by the storm

Should we, then, declare this group of words to be a "sense"? There are difficulties. The various words carry nuances, which it may or may not be easy to ignore in a particular context. "Forsake", for example, can suggest that there is something reprehensible about the action. It can also connote formal renunciation, and the above example from a marriage vow shows that the formality can be present without the reprehensibility. Nuances get in the way of interchangeability; it would sound strange to substitute "desert" into the marriage vow.

Besides nuances, the individual words have conventional areas of application. One does not normally say that the doctors "deserted" all hope, or that an errant husband "surrendered" his wife and children. The minister officiating at a wedding would be considered daft if he adjured the bride and groom to "abdicate" all others, and a merchant would not advertize that he was "relinquishing" his entire stock at a loss. (Somehow, the latter situation calls for more pedestrian language.)

At the opposite extreme, overawed by this lack of interchangeability, we might decide to respect the unique personality of each word, abolishing equivalence classes altogether. The inconvenience of such a cop-out is obvious: we then have to introduce some *other* mechanism for recognizing the equivalence of utterances that are intended synonymously, though they employ different words. But beyond being inconvenient, the exclusion of equivalence sets is a denial of linguistic facts -- just as bad, in its own way, as the naive attribution of unconditional synonymy.

For it is a commonplace of everyone's experience that the speaker and the listener agree to ignore the nuances of words, whenever nuances get in the way of communication. A writer who has used the word "give up" eight times in five lines will surely cast about for some alternative ways of saying the same thing. If "relinquish" and "abandon" would normally be too flowery, or if "surrender" would in other circumstances call to mind an armistice ceremony in a railway wagon, that will not deter the writer from tossing in a few occurrences of those words -- once a context has been established that discourages the overtones. Nor will the reader understand matters any differently. It is as if writer and reader conspired: "We're fed up with *that* word, let's hear another." Or, perhaps, the writer simply connives at jolting the reader awake with frequent changes of idiom, maybe even an occasional incongruity. In any case, synonymy is imposed upon

the words, and this literary behavior merely exaggerates what people do habitually in common speech.

Not only can words be stripped of nuances normally present; they can *take on* colorations suggested by the context. The suggestion of "reluctance" conveyed by all the verbs of our example can be inferred, in at least one case, from the setting alone; and in this case, a variety of more neutral verbs could be used synonymously:

(part with, take leave of) our entire stock at a loss

One could even substitute the word "sell", and it wouldn't change the meaning that was already read into the utterance. But to admit context-dependent synonymy of this degree is to stretch the equivalence sets" to the point of uselessness.

It comes to this: neither the grouping nor the separation of words can be fully justified. Grouping is nearly always conditional, and separation is often so. If one could anticipate all possible contexts in which a group of words could occur, one could perhaps enumerate all possible equivalence sets -- one for each combination of word group with a set of contexts making the words interchangeable. Anyone, however, can see the futility of that aspiration.

In the end, one settles for messy compromises. Words are grouped if a largish set of contexts in which they are interchangeable springs readily to mind. They are separated (into perhaps overlapping groups) if the imagination readily suggests contexts in which their meanings differ "significantly" -- whatever "significantly" may mean. In doubtful cases, when words are grouped somewhat questionably, one promises oneself to add markings some day that will prevent misuse of the equivalence. When words are separated somewhat questionably, one promises oneself to add a mechanism some day that will recognize their relatedness.

In the end, too, one assigns internal structure to the equivalence sets. That's the effect of assigning local attributes to the alternative wordings ("animate subject", "object a vehicle", etc.): constraints are imposed upon the interchangeability of the wordings. More radical structuring can be accomplished if, for example, one notes "government" as an alternative wording of the sense "govern, rule, control", with the attribute "nominalization".

A trenchant discussion of such difficulties may be found in Kelly and Stone [4]. There the emphasis is upon disambiguation: given a word in a passage of text, they seek to identify (by selection from a fixed list of possibilities) the sense in which it is used. Building a computerized

dictionary for the purpose, they soon became concerned with the arbitariness and the proliferation of target "senses", as taken from standard desk dictionaries. They argue, with persuasive examples, that what lexicographers conventionally distinguish as separate senses of a word are often just applications of the word's underlying concept to different contexts. To cover the various contexts, the underlying concept has to be stretched a little, by a process of metaphoric extension. This metaphoric process is beyond our present power to computerize, but for the long run looks indispensable for successful language processing. Meanwhile, the authors advocate a dictionary which records for each word as few discrete senses as practicable, combining into one sense all the usages which can reasonably be united by a common underlying thought.

It is interesting to re-examine Kelly and Stone's argument with a different task in mind: not the disambiguation of one word, but the recognition of synonymy between two words. A metaphorical capability would be as useful for the one task as for the other, but in the case of synonym recognition, some of the considerations which have guided traditional lexicography remain pertinent. In particular, it is necessary to ask not merely whether the concepts overlap, but whether the one word may in fact be used in place of the other. As noted before, usage is restricted by conventional domains of application; for example, an "alteration" is conceptually both a "change" and a "modification", but one wouldn't *call* it a change or a modification when painting a sign for a tailor's shop.

The arbitrariness of the equivalence sets is not all that disqualifies them as "conceptual primitives". There is a much deeper difficulty in the fact that practically all "senses" can be paraphrased in terms of other "senses". Take, for example, the intransitive sense of "change" (as in "My, but you've changed!"). Surely, one would suppose, the concept of "change" must be primitive? Change of state is what well-nigh a third of all verbs are about.

But if "change" is a "primitive", it's a peculiar sort of "primitive", for it can be paraphrased in a variety of ways:

> (change, become different, cease to be the same, assume new characteristics, make a transition into a new state)

Note that the multi-word paraphrasals are not idioms; the individual words contribute their usual meanings to concatenated meanings which express the concept "change".

But perhaps we were merely unlucky? Perhaps we chanced upon a concept which looked elemental but actually turned out to be complex. Maybe the real primitives are "become", "be", "cease", "different", "same", etc. Let's dig into that possibility.

What does it mean to "become X", where X is an adjective? The meaning can be variously expressed:

(become X, come to be X, get to be X, get X, turn X, grow X, assume the characteristic X)

That's a discouraging number of ways for a "primitive" to be re-expressible -- though if we choose to regard "come to be" and "get to be" as idiomatic concatenations of words, only one of the alternatives makes use of other concepts to explain the one at hand.

As for "different", it implies a whole underlying anecdote about somebody making a comparison, after first making a judgment about relevant things to compare. In the combination of the two concepts -- "become different" --, we furthermore drop mention of the objects being compared. It's simply understood that they are certain attributes of the subject at two points in time.

It is tempting to invent ad-hoc "transformational" explanations for these phenomena. One might conjecture, for example, that "The man changed." is a surface realization of four underlying sentences:

(Man be X at time m. Man be Y at time n. X not equal Y. Time n greater-than time m.)

The trouble with explanations of this sort -- apart from the fact that they introduce growing complexity into the understanding of straightforward utterances -- is that they assign arbitrary primacy to some concepts at the expense of others. Why should

"time n greater-than time m"

be an assumed primitive? May we not equally well conjecture that "time n greater-than time m" is a surface realization of these?:

(Time be m. Time change. Then time be n.)

For that matter, why not view

"Time elapsed."

as a surface form of this?:

"At least one thing in the universe changed."

After all, what is "time" but a nominalized way of talking about the presence and partitioning of change?

The difficulty, it would seem, lies in the very notion of context-independent "conceptual primitives". The metaphor itself is at fault: it calls to mind a fixed set of elements, like those of which matter is composed, out of which all ideas must be compounded. But where concepts are concerned, primitivity is a matter of focus. Shift the perspective a little, and new elements swim into view as fundamentals, while former simples become complex.

A more promising metaphor is the analogy to a vector space. A set of basis vectors is, in a way, a set of "primitives" out of which all the entities in the space can be composed. These primitives have the appealing property that they are only primitive relative to one frame of reference. Rotate your point of view, and what used to come natural as basis vectors are now at an angle; they become easier to express as sums of vectors that lie along new axes. That bears a resemblance to what we have seen in the case of lexical "primitives".

Thus far and no further may the analogy be pushed, however. The elements which span "conceptual space" can be no such uniform set of objects as those in a vector space, while the rules of composition are coextensive with grammar -- at a minimum. Composition of concepts itself contributes to the meaning. (For that matter, it is arguable whether concepts are sufficiently separable to model them as discrete objects at all -- whether simple or composite.) Moreover as "conceptual space" must encompass all things thinkable, the rules of composition must themselves be part of the space. That is, the operators as much as the things operated upon lie within the space to be spanned.

A seeming counterexample to these remarks may be found in the "primitive ACT's" of conceptual dependency theory, as propounded by Schank, Goldman, Rieger, and Riesbeck ([2], [7], [8], [9]). On a close reading, however, the "primitive ACT's" turn out to be *verb paradigms* -- powerful, semantically motivated *generalizations* about large classes of verbs. The names of these paradigms replace specific verbs as building blocks in the "conceptual" representation of an utterance. The

effect is to provide strong guidelines for the inference of unstated information, for the comparison of related utterances, for paraphrasal, etc.

To represent a particular verb in terms of these ACT's, however, it is necessary to augment each ACT with various substructures which detail the manner, the means, the type of actor or object, etc. No reduced set of representatives is as yet offered for the adverbs, nouns, adjectives, etc. in terms of which the "primitive ACT's" are qualified. If such additional condensation were attempted, the elaboration of a given utterance in terms of the *full* set of "primitives" might well ramify without practical end. In other words, reduction of the set of names for nodes (and labels for arcs) must be purchased at the expense of extending the number of them required to represent each utterance.

In conceptual dependency representation, just as in the semantic networks" of Quillian [6], Simmons ([11], [12]), Slocum, and others, reality ultimately appears as a shimmering web, every part of which trembles when any part of it is touched upon. Taken in its totality, the system -- as yet -- is entirely compatible with skepticism about a *comprehensive* set of "conceptual primitives"

In any case, the verbal "senses" proposed here lie at a far lower level of generality than the "primitive ACT's" used in conceptual dependency theory. In terms of that theory, they come closest to the so-called "CONCEXICON entries" used by Goldman in realizing surface expressions of a concept from its conceptual representation [2]. Given a primitive ACT, Goldman narrows it down to a particular "CONCEXICON" entry by applying the tests in a discrimination tree to the rest of the structure in which the ACT appears.

Our lexical "senses", therefore, are left with a humbled role. If they span anything, it might best be thought of as "communication space", not "conceptual space". Even in this light, they are a hugely redundant basis, and a not at all unique one. They form no inventory of the experiences being communicated about; "meaning" is still a step removed, still evoked rather than embodied by the elements of this basis.

If we persist in calling these things "senses", it is because that is the traditional term for what is *brought to mind* as the synonym sets of a given word are enumerated. The tie-in with meaning is tenuous, but the human user is able to supply it. There is at least this much justification for the term: synonym sets, more forcefully than words, direct attention to the points at which a tie-in must be made between the tokens of communication and the underlying representation of "world knowledge"

In a full-fledged system for processing natural language, then, we must envision the "dictionary of senses" as a component stretching vertically across the "upper" layers. Its "sense data items" must link, in some way, to the deeper-lying data structures which encode "knowledge of the world" (the "pragmatic component"). The "key data items" and "phrase data items" register tokens to be expected or employed in "surface" utterances. Global and local attributes recorded in the various data items guide parsing and interpretation. Where one takes it from there depends upon the linguistic approach to be used.

References:

[1] Feigenbaum, Edward A. (1963), "Simulation of Verbal Learning Behavior", in *Computers and Thought*, eds. E. A. Feigenbaum and J. Feldman, McGraw Hill.

[2] Goldman, Neil (1975), "Sentence Paraphrasing from a Conceptual Base", *Communications of the ACM*, February, 1975, Vol. 18 No. 2.

[3] Harris, Larry R. (1972), "A Model for Adaptive Problem Solving Applied to Natural Language Acquisition", Cornell University, Ithaca, N.Y. PB-211 378.

[4] Kelly, Edward, and Stone, Philip (1975), "Computer Recognition of English Word Senses", Chapter IV, North-Holland Publishing Co, Amsterdam.

[5] Lamb, Sydney M., and Jacobsen, William H., Jr. (1966), "A High-Speed Large-Capacity Dictionary System", in *Readings in Automatic Language Processing*, ed. David G. Hays, American Elsevier Publishing Company, New York.

[6] Quillian, M. Ross (1968), "Semantic Memory", in *Semantic Information Processing*, ed. Marvin Minsky, The MIT Press, Cambridge, Massachusetts.

[7] Schank, R., Goldman, N., Rieger, C., and Riesbeck, C. (1973), "Margie: Memory, Analysis, Response Generation, and Inference on English", *Proceedings, Third International Joint Conference on Artificial Intelligence*, Stanford Research Institute, Stanford, California.

[8] Schank, Roger C. (1973), "Identification of Conceptualizations Underlying Natural Language", in *Computer Models of Thought and Language*, eds. R. Schank and K Colby, W. H. Freeman & Co., San Francisco.

[9] Schank, Roger C. (1973), "The Conceptual Analysis of Natural Language", in *Natural Language Processing*, ed. Randall Rustin, Algorithmics Press, Inc., New York.

[10] Schmidt, Charles T. (1970), "A Dictionary Structure for Use with an English Language Preprocessor to a Computerized Information Retrieval System", Naval Postgraduate School, Monterey, California. AD 710 363.

[11] Simmons, R. F., and Slocum, J. (1972), "Generating English Discourse from Semantic Networks", *Communications of the ACM,* October 1972, Vol. 15 No. 10.

[12] Simmons, R.F. (1973), "Semantic Networks: Their Computation and Use for Understanding English Sentences", in *Computer Models of Thought and Language,* eds. R. Schank and K. Colby, W. H. Freeman & Co., San Francisco.