UNIVERSITY OF QUEENSLAND

DEPARTMENT

OF

COMPUTER SCIENCE

St Lucia,

Queensland, 4067,

Australia

TECHNICAL REPORT

No. 58

Context and Linguistic Knowledge in Natural Language Processing

W.H. Wilson

December, 1984

Abstract

This paper investigates the prospect of dividing linguistic knowledge into modules, which could be used, a few at a time, in understanding a sequence of natural language utterances. This approach leads to the need to keep track of the modules which are currently relevant at successive parts of a sequence of natural language utterances. Factors in determining the current context set include recent context history, themes announced explicitly at the start of a natural language passage, knowledge associated with words in the current clause or phrase, or in nearby clauses or phrases, and a bias towards semantic continuity. Benefits include a reduction in the problem of word-sense ambiguity.

1. Introduction

By linguistic knowledge, we mean knowledge which is important for understanding natural language. Clearly this is a matter of degree. Syntax is normally of great importance in understanding a natural language passage. Verbs often have strong preferences about the types of the objects which play roles in the action which the verb describe: these preferences would rate as (semantic) linguistic knowledge. Knowledge about how the human ear converts audio signals into neural impulses would be important only for understanding very specialized natural language passages on that topic, and so would not be linguistic knowledge in the sense in which we are using the term (except in that special case). Section two of this paper describes experimental evidence which led to the hypothesis that much of linguistic knowledge is best organized into coherent modules and that this is because of the need to maintain a context in which to understand natural language passages. Section three considers issues of representation and organization of linguistic knowledge, with this modularization and the need to monitor context as an objective. Section four of the paper discusses and describes components in a context tracking algorithm.

2. Modularity in Human Semantic Knowledge.

We begin this section by describing an experiment performed by the author and discussing the conclusions which might be drawn from it. Subsequent sections describe other relevant experimental work, and relevant work within Artificial Intelligence.

2.1. The 'nock' experiment.

This experiment was performed by the author on a group of 48 final-year undergraduate Australian computer science students. The subjects were presented with a list of character strings (in a column):

bit	nock	sult
freeze	sill	table
mouse	sla	xrqj

ranging from those like 'xrqj' which were obviously nonsense, through ones like 'sult', which have the sound pattern typical of English words, to ones like 'mouse', which really are words in the English language. The were then asked to mark those that they believed were genuine English words, (and give definitions or examples of their use), and those that were not. 43 out of the 48 subjects did not recognize 'nock' as a word. They were then given the cue words 'archery' and 'arrow', and asked if this changed their views on any of the strings. 12 out of the 43 now recognized 'nock', and none 'recognized' any non-word. Some of the 12 could not give a formal definition for 'nock', but said it fitted into phrases such as 'nock an arrow'. Only one subject mentioned the use of 'nock' as a noun.

It is interesting that, although they had been strongly encouraged to give definitions rather than examples if they could do so, many of the subjects gave examples as responses to quite simple words like 'mouse'.

Inferences

It seems that marginally understood words are not always accessible directly to human memory, without a suitable surrounding context, or at least not when there is a question as to whether the string of letters in question is a proper word. When that context is supplied, the word becomes recognized and so understood, at least as part of a phrase used in that context. Of course, many other words are recognized out of context, (although it is not clear that all of these are fully understood out of context). The evidence from the example of this marginal word is that a frame of reference may be important for understanding.

Furthermore, most readers will agree on introspection that non-sequiturs take longer to understand than utterances continuous in theme with past utterances. This all suggests that humans group words and knowledge together into relatively small *modules*, and that access to meanings, or to part of the meanings, may be by way of an appropriate module. If at any one time only certain modules are *active*, that is, in the forefront of consciousness, then it is reasonable to further assume that meanings relating to those modules will be assumed to apply to any utterance heard. This would help to explain why humans have so little difficulty (most of the time) in disambiguating words in utterances.

2.2. Other Evidence for Modularity

This interpretation also ties in neatly with the well-known experiment of Bransford and Johnson (1972), (reported in Bourne, Dominowski, and Loftus, 1979). In this experiment, two groups of subjects were tested for comprehension of a turgid prose passage which makes only implicit reference to the concepts it is manipulating. A sample of the text used is shown below:

> First you arrange things into different groups. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, ... it is better to do too few things at once than too many. ...

One group were told that the passage was about washing clothes, and they understood and remembered it well; the other group were told nothing and could understand or remember little. In this experiment, the text gives relationships between unspecified objects, which can only be assigned to slots in a frame when a particular context is given.

Further evidence from psychology includes the experiments and conclusions of Freedman and Loftus (1971) (reported in Bourne, Dominowski and Loftus, 1979) in which it was observed that mean response time to questions like *What is the name* *of a fruit that is yellow?* is less than that for questions like *What is the name of a yellow fruit?* The explanation, in terms of knowledge modules, would be that *yellow* does not determine a module, whereas *fruit* does. This means that the search for the *fruit that is yellow* begins when the question is complete, while the search for the *yellow fruit* begins after the question is complete *and* the *fruit* module has been accessed. With *fruit that is yellow,* this access can be overlapped with processing the rest of the sentence.

Collins and Quillian (1970), Loftus (1973), and Loftus and Loftus (1974), (again reported in Bourne, Dominowski, and Loftus, 1979) showed that, for humans, a piece of information is easier to produce if they have recently produced a related piece. Subjects first asked to name one member of a category, and soon thereafter a different member of the same category, respond more rapidly the second time. A refinement of this work, due to Swinney (1979), is mentioned later in this section, in connection with a discussion of some work of Charniak (1983).

2.3. Related Work Within Artificial Intelligence

Related work within mainstream Artificial Intelligence research is to be found in Walker (1978, Sections 3 and 4) and related papers, which describe (among other things) ideas of Grosz and Hendrix on focus spaces and partitioning of semantic nets. Clearly there is some commonality of theme between the notion of partitioning semantic networks and modularizing linguistic knowledge. The distinction between that work and this is as follows. Firstly, Grosz's concept of focus usually relates to an object or small group of objects in a semantic network, whereas a module is intended to refer to a natural grouping of all available linguistic knowledge about some topic. Secondly, Hendrix seems to impose his partitioning on an overall semantic net in order, for example, to limit or group information relevant to the current problem or sub-problem, or to provide a representation of the meaning of the text or utterance currently being processed, whereas each module is intended to be sufficient by itself - a largely independent and modular body of knowledge. Overlaps and interfaces are inevitable, but that module names give rapid access to particular parts of the overall body of knowledge. We shall (later) define *domain focus*, a term intended to convey both the relationship and the distinction between this work and that of Grosz and Hendrix.

Wilks concept of a *thesaurus* (Wilks, 1978) is similar to some aspects of our chosen representation of semantic linguistic knowledge. His *preferences* effectively perform, among other things, a role similar to a local form of domain focus, in that they would allow word sense disambiguation in favor of the preferences of nearby words.

The work of Charniak (1978 and 1981) is relevant to the knowledge representation strategy used in this paper, in that he uses stored knowledge templates for words or concepts to reduce semantic processing and reasoning.

Charniak (1983) has also considered contextual influence in language comprehension, with an (un-modularized) semantic net as his model for real-world knowledge. He argues against systems in which the possible word meanings considered are constrained by context, relating his objections to experimental work of Swinney (1979). The theory presented in this paper of Charniak proceeds to constrain possible word meanings in a way which is weak enough to escape the problems posed by Swinney's work. Some details will be presented in a later section, when some definitions have been made. It may be that Charniak's method is formally equivalent in algorithmic terms to the ideas presented in this paper; the emphasis in this paper is primarily on the management and representation of linguistic knowledge, though of course we give an account of an algorithm for exploiting this representation.

Word expert parsing, as espoused by Small (1981), for example, is one way of organizing the large amount of knowledge needed for natural language understanding. For each word, there is a small expert system which contributes information about how that word should interact with neighboring words, and from the combined efforts of the different experts, an overall meaning for a natural language fragment is arrived at. This concept seems useful as a technique to be used in parallel with domain focus: in itself it seems not to use the constraints provided by the discreteness of context in a natural way; domain focus provides some of what is needed or useful here. One of the tasks of a word-expert is word-sense disambiguation, a problem which could be reduced by having sub-experts for the same word in particular modules. There is no natural reason to store the information about, say, *port* = a type a wine, with that about *port* = a transport terminal. Nor is it necessary to have the word expert for *port* able to distinguish, by itself, between the possible word-senses for *port*, if the theme of the passage (wine-tasting, perhaps) has already been deduced in another manner.

However, we agree with Small (1981) that humans extract meaning, at least partly, by knowing "how the particular words interact with each other." Wilks (1978) is saying something similar. We will use such knowledge to enrich our context tracking algorithm so that it will use word interactions to decide between possible candidates for modules.

3. Linguistic Knowledge and Modularization

Two related problems in natural language understanding are the classification of linguistic knowledge, and the task of representing the different types of knowledge emerging from such a classification. The term *linguistic knowledge* is intended to cover both knowledge which is obviously linguistic, including syntactic categories like *noun group* and *verb*, and perhaps definitions of words, and also knowledge which is, arguably, independent of language, including rules like *X eats* → *X is less hungry*, which do not define *eat* or *hungry*, but specify a relationship between them. We shall list some of these varieties of linguistic knowledge, and to comment on the representation of the different sorts of knowledge. The criteria to be used in the classification will include the following. Is the knowledge syntactic or semantic? How will the knowledge will be used? What are natural ways of representing this knowledge? Does the knowledge relate to single words or to various larger groupings of words?

3.1. Classification of Linguistic Knowledge

Syntactic knowledge is well enough understood that, for the purposes of this paper, the following brief comments are adequate. The kind of syntactic knowledge which should be stored about a word corresponds closely to the word's syntactic class: noun, verb, adjective or whatever it is. Syntactic fragmentation of sentences into noun groups, verb groups and prepositional groups froms a basis for refinements to the concept of context tracking. These refinements will be discussed after the context tracking algorithm is described, in section four of this paper. A facet of syntax which is relevant to modularization of linguistic knowledge relates to parts of speech. Thus it is convenient to note here that a given word may suit several parts of speech, either because of ambiguity (*fly* as in insect or as in flap wings), or because of multiple uses for a word with a single meaning (as with *a meeting room*, *went to a council meeting*, and *we are meeting tonight.*)

Semantic linguistic knowledge is oriented towards properties of words, or rather, of the concepts which the words stand for. Categories of semantic linguistic knowledge include:

* *Definitions:* of words not regarded as primitive terms. As pointed out by Bobrow and Winograd (1977), there may be more than one view of the concept underlying a word. The example they use is the word *kiss*, which may be viewed as a physiological event, or as a social event. On top of this, there will also be many words which have more than one meaning, and hence more than one definition. Definitions are regarded here as one way of relating words to other words. It is not intended that all imprimitive words should be reduced to primitive words by a network of definitions as a matter of principle: compare the remarks of Bobrow and Winograd, (1977). Rather, a set of words might

be chosen as *basic* for a particular module of interest, and semantic representations of utterances would be in terms of these basic words. The basic words might well be decomposable from some other point of view, much as, in physical science, atoms are indecomposable from the point of view of molecular physics, but decomposable from the point of view of sub-atomic physics. Thus a word might be decomposed in one module, and basic in another.

* *Operations:* associated with a particular word, such as the activities a noun can participate in. Thus *child* wold participate in *play* and *learn,* and would inherit the attributes of *person*, such as *eat* (though not those of *mammal*, such as *gives-milk*).

* *Root Forms of Words:* and the relationship to the root form. For example, the root form of *buyer* is *buy*, and the relationship is that a *buyer* is (agent of *buys*).

* *Categories to which the word belongs:* as *green* belongs to *color*, and *cat* belongs to *mammal.*

* *Substructure:* the components of a noun (for example the legs of a *table*) or the subprocesses of verb which represents an activity (for example: move food to mouth; masticate food; swallow food - as subprocesses of *eat.*)

* *Related and Opposed Concepts:* as *war* is related to *battle* and opposed to *peace.*

* *Attributes of a Word.* What attributes does the object or activity represented by a word have? What are typical values of those attributes? For example, for *house*, an attribute would be *size*, and the value might be *large*, or *18 squares.*

* *History or Origins:* for concepts represented by some words, notably proper nouns.

* *Expected Future Events or States.* The word *pregnant* carries with it the expectation *will give birth*, for example.

* *Rules:* for manipulating the object, state, or activity underlying the word. For example, a rule for the word *letter* in a module such as 'communication' would be:

  IF *sealed* AND *stamped* → *post.*

* *Relationships:* which the word may participate in (other than ones listed above). For example, *Person owns Pet.*

* *Effects on the meaning of dependent or nearby words:* For each verb, for example, there will be knowledge of surface-to-semantic translations for prepositions in associated prepositional groups.

* *Domains* to which the word is relevant. This information may not be a simple list: it may in some cases be a list of rules helping to specify when certain modules are relevant.

The purpose of introducing relations and activities in which words participate is to make available a few key concepts needed for dealing with this word. For nouns, in particular, it may be compared to defining the fundamental operations on a newly invented data type. It is not intended to clutter up the system with the totality of human knowledge.

Some of these knowledge categories need no further comment in relation to particular parts of speech. Examples and descriptions of the rest are given in later sub-sections.

There is no guarantee that there are not other important categories of semantic linguistic knowledge. The point of the list, then, is that there are many and diverse types of knowledge, and that schemes to find a single underlying representation for all types may be counter-productive. Languages like PROLOG lean towards the idea that most knowledge can be represented in terms of relational assertions and rules, but while the generality of this scheme is unarguable, it is not clear that it is the best way to represent arbitrary units of knowledge. It seems more likely that humans represent knowledge in a way which makes *objects*, as well as relationships and rules, into representational foci. In PROLOG an object is constructed, from the relationships in which in takes part and related rules, in an analogue of short term memory, at the time the object is needed. It would also be convenient to have explicit facilities for handling records, trees, directed acyclic graphs and strongly-typed data, as well as relations, predicates, rules and logical operators.

We shall use a *frame*-like notation, represented as a nested list, as well as rules, to provide a representation of some forms of linguistic knowledge, for purposes of illustration.

3.2. Semantic Knowledge About Nouns.

This section inspects semantic linguistic knowledge from a special angle, namely what can be said about nouns and the concepts or objects they represent, and how this might be related to modularization of linguistic knowledge. In order to illustrate this knowledge we need to develop at least a working representation for the kinds of knowledge we find are useful for nouns. As foreshadowed above, we shall base this representation on frames, lists, and situation-action rules, with relational assertions and logical operators inserted where appropriate. Words preceded by a $ sign are words intended to have special significance to the knowledge representation system.

In this section, nominally about nouns, we will not always confine ourselves to nouns in cases where something more general can be said.

3.2.1. Definitions.

Not all nouns will have definitions: some will have to be regarded by the system as primitive. This does not mean that nothing can be said about such words, only that they are not further analyzed into structural or other components. In

Winograd's Blocks World, for example (Winograd, 1972), it would be pointless to analyze the word *block* into fibers of wood or grains of plastic or whatever. The purposes of the system are served by regarding *block* as a primitive object. One would, of course, still want to specify that a block has attributes such as *color*, *shape*, *size* and so on. A definition may be complex and irrelevant.

Efficiency, utility, and consistency argue against the definitions being themselves expressed in natural language. The system must have a representation for the meanings of the utterances it is processing, and it seems sensible to use the same representation, within reason, for units of meaning built in to the system, as most word definitions must be.

Substructure and Subprocess Definitions

The precise nature of the definition will depend on the word: some words, such as *kill* (as a verb), can be defined fairly simply by some equivalent of *end life of*, while other words, such as *strangle*, can be better defined by giving the substructure or subprocesses of the concept, or perhaps some other attribute such as the purpose. Thus, for example, *table* could be defined as a kind of furniture consisting, typically, of a planar top surface, a support, and being used to support objects at a convenient height. This brings us to the question of defaults and exceptions: it would be convenient to store with the structural description, say, a common value for the support, namely four legs or a pillar. Sinre not all tables are like this, however, it is important to record that this 'four legs or a pillar' is merely a default, or common type of support. Thus the frame-like nature of these definitions begins to emerge.

In a frame-like notation, the definition of *table* might thus appear as

```
table    (($kind-of-concept object)
          ($parts ((top ($must-be planar)) (support ($must-be legs))))
          ($ako furniture)
          ($purpose (to support something)))
```

Multiple Definitions for One Word

A glance at the definition of *table* in a dictionary reminds us that structure and purpose do not tell all about the word *table*. One dictionary lists 25 definitions under *table*, plus several uses in phrases like *table tennis* and *table napkin*, whose meanings could not be readily deduced from those of the constituent words.

What is missing? What sort of problems do the multiple definitions pose? Some of the 25 dictionary definitions referred to above relate to specialized uses of *table* in technical domains such as architecture and printing. Others relate to specialized senses of table in everyday use, such as 'a group of people seated together in a restaurant'. Other meanings may relate to metaphorical uses of *table*. As it is too much to expect a natural language understanding system to regularly unravel figurative language from scratch, the additional meaning must be available to the system if their occurrence in the input text is at all likely.

How do we choose between different definitions of a single word or phrase? One solution is to build an expert system around each word, in order to deduce, from the context, which definition should apply (as in Small (1981).) However, there is no guarantee, with this method, that the expert systems for a group of successive words will produce a coherent sequence of definitions. If the sequence of decisions is coherent, then it is possible that similar decisions have been taken repeatedly for successive words. Another way of handling this ambiguity problem is to divide the universe of discourse into a number of simpler, more limited universes of discourse, in each of which each word would have zero, one, or at worst a small number of possible meanings. This method can be contrasted with the one-expert-system-per-word method by observing that, in effect, the modular domain decomposition idea supports one (larger) expert system per *group of related words.*

Dictionary versus Thesaurus

In fact, the conventional concept of dictionary may have been an unfortunate one for natural language understanding. The normal alphabetical organization which dictionaries possess is suited for humans to learn the meaning of a word, since the human information processor has available an enormous and high-speed associative memory containing myriads of idiomatic expressions and contextual information. This information relates both to the text of the dictionary entry, and to the situation where the word being looked up has been encountered.

Thus, the word store of an natural language understanding system should probably resemble the associative network used by a human as the internal representation of linguistic knowledge, not the list of definitions suggested by the term 'dictionary'. Thus information about the different senses of a word such as *port*, would, as shown below, be found in quite different places. At any given time during the understanding process, a relatively small selection of *modules* would be active, or available, and this would determine which sense of a word such as *port* is assumed.

```
($module fluid ...
         port (($ako wine) ($attributes alcoholic ... )) ... )

($module transport ...
         port (($ako direction) ... )
         port (($ako harbor terminal opening) ... ) ... )

($module luggage ...
         port (($ako suitcase) ... ) ... )

($module computer-hardware ...
         port (($ako connection) ... ) ... )
```

3.2.2. Activities.

Another group of knowledge relates to the activities a noun (or other part of speech) can participate in, and in what role it participates in such activities. Thus this class of knowledge helps specify the words with which the word we are considering frequently and characteristically appears. This activity knowledge can frequently be expressed as a sequence of situation-action rules such as the following examples, for *meat*:

```
meat    ((IF frozen → $may-do thaw)
         (IF raw → $may-do freeze cook)
         (IF cooked → $may-do serve)
         (IF served → $may-do eat))
```

3.2.3. Attributes.

The knowledge about activities, described above, principally relates to the interactions of nouns and verbs. Some interactions of nouns and adjectives are described by the *attributes* a noun can have, and the possible/impossible values of those attributes, where sensible information on this is available. Pursuing our example of the word *meat*:

```
meat     ((color red white brown green)
          (weight $unknown)
          (varieties lamb mutton beef veal pork fish chicken ... ))
```

Other noun/adjective interactions relate to inferences, and can be expressed using rules once again:

```
meat     ((IF green → $must-be rotten)
          (IF brown → $must-be cooked)
          (IF hot → $must-be cooked)
          (IF chicken → $must-be white)
          (IF warm → $must-be $one-of cooked recently-killed badly-stored))
```

3.2.4. Roots, and Relation to Roots.

Another kind of knowledge relates to a root form which exists for some words. For many problems of inference, this information is enough to enable a solution to be found. For example, the information about the words *freeze*, *frozen*, and *freezer*, shown below,

```
freezer  ($root freeze performs)
          ($similar-to refrigerator)
```

would allow one to plausibly infer from the brief dialogue

'I can't find it in the refrigerator.'

'I froze it when we bought it.'

that the object referred to is 'in the freezer', provided a freezer is known to be present, and syntactic information about the relationship between *froze* and *freeze* is available.

3.3. Semantic Knowledge About Verbs.

3.3.1. Definitions of Verbs.

Much of what was said above about definitions applies to verbs: only the slots and slot-values in the frames need to change. Taking the verb *eat* as an example:

```
eat       (($kind-of-concept action)
           ($subprocesses
                   ((move [object] to mouth of [actor])
                   (swallow [object])))
           ($purpose (nourish [actor]))
           ($object  ($default food)
                   ($typical bread cheese meat butter vegetables fruit))
           ($actor $unknown))
```

In practice, subprocesses might well be pointers to other frames.

3.3.2. Other Knowledge for Verbs.

Other types of knowledge for verbs, as for nouns, may describe words which commonly occur with the verb in question. For example, for the verb *eat*, one might list typical instruments, sources, locations, (all nouns), and manners (adverbs).

```
eat     (($instrument knife fork spoon)
         ($source plate bowl packet)
         ($location table restaurant refectory kitchen pantry)
         ($manner greedily hungrily daintily slowly))
```

In a natural language system, as with any definitional system, there have to be primitive terms, which are not further defined. Many systems have words which are treated as commands, (to answer a question, insert information into a data base, etc.) These words may be primitive in the sense of not being defined in terms of other words, while being non-primitive in the sense of that they are specified by executable code within the system. In this paper we are interested in the first sense of primitivity.

Common verbs may not be as amenable to domain modularization as common nouns. Much depends on the choice of modules. An example of the problem is the verb *sustain*: in

man sustains injury

sustain may be paraphrased as *suffer*, whereas, in

food sustains life

sustain may be paraphrased as *support* *support* and *suffer* are, of course, quite different. With an unlucky choice of module, it might well appear that both these sentences referred to **health**, so that module could not be used (by itself) to choose the correct definition. With a lucky choice of module, however, the first phrase might well come down in the **tissue damage** module, and the second in **food** module, and in that case there would be no problem. Clearly much depends on the details of the break-up of knowledge into modules.

3.4. Semantic Knowledge About Adjectives.

In part, the treatment of adjectives as subjects of knowledge is dual to their earlier mentions as modifiers of nouns. Thus, of *green* we would say that it is one possible value of the attribute *$color*, which may be predicated only of physical objects. It is probably unnecessary to record for the adjective *green* the nouns which are likely to be *green*, as this kind of information can more economically be recorded in association with the nouns.

4. Context Tracking

For a considerable time, systems with enough knowledge to understand a range of utterances relating to some limited domain of discourse have been possible (Winograd (1972), for example). Most natural language understanding systems have, indeed, confined themselves to a limited domains, either as a restriction on complexity and the size of the linguistic knowledge base, or because an explosion of processing time or problems of ambiguity forced a limited domain approach.

By limiting the scope of a domain of discourse *at any one time* to a sufficiently small area, (a group of knowledge modules,) it should be possible to achieve an natural language understanding system which is relatively efficient, and yet as general as necessary. We shall refer to the module(s) of interest at a particular time as the *domain focus* at that time. The knowledge modules of domain focus constitute both the *context* in which understanding is attempted, and the *domain* in which the system is currently expert. Hence the words *context* and *domain* or *domain focus* may be used almost interchangeably in what follows. Names of modules will be in bold face, as in **personal relationships** and **car maintenance.**

Understanding a large range of utterances may well depend on understanding a number of smaller domains, and on having an effective system for switching domain focus. For good context switching, it would be necessary firstly to detect when the latest utterance lies outside the scope of the current domain, and secondly to be able to decide on, and access the knowledge of, an appropriate new knowledge module.

4.1. How To Switch Domain Focus

How do we decide what the current domain focus is? That is, on what basis do we calculate the names of the module(s) which will form the domain focus set which will be used to determine the meaning of a word or group of words?

There are several possibilities. We can somehow arrive at an initial domain focus, and thereafter only change domain focus when it proves to be impossible to understand something in the current domain. Alternatively, we can assume that choice of domain focus should be associated, at least in part with the incoming sentence, or perhaps clause, and calculate the module information for each sentence as part of processing it. We can note that sometimes sequences of utterances have a hierarchical structure, going from the general to the particular and back again, and so decide to stack old modules and attempt to understand by 'looking down' into the stack of domain focus sets when an understanding fault occurs.

To hope that a single module will always be an adequate focus for a whole sentence is unrealistic. On the transitions from one module to another, we can expect to find a section of the utterance where both modules are relevant. And indeed, many utterances may relate to interactions between two or more modules, as in

> Lovingly, Steve repaired his girlfriend's car.

Here, **personal relationships** and **car maintenance** are interlaced, and both vie for the role of domain focus. One way to handle this is to allow several knowledge modules to be available at once (and this will surely be necessary in practice, in cases where the compartmentalization of knowledge chosen by the system designer does not agree well with a particular sentence). Another method is to use different focus sets for different components of the sentence: this will be discussed later.

Maintaining a hierarchical domain focus, as suggested above, leads to the question of when to discard modules from the focus set. Again, we can look at the human natural language understanding system for guidance. Humans are not good at thinking about several things (for example, modules) at once, except in cases where they have been trained to do so (and thus presumably have integrated the multiple modules in some fashion). There is certainly some memory of recently mentioned modules, on the other hand.

Forgetfulness

Plausibly, a data structure of limited size is maintained, and new modules are discarded on a least recently used basis when necessary, and perhaps also after a certain length of time has passed without use of a particular module (whether or not there is a need to discard it to make room for new modules.) Human domain focus evidently works this way, and there are informal protocols for returning to earlier topics in a conversation. One might then ask where the old domain focus sets are re-captured from in such cases. However, the relevant information would still be available within the stored semantic representation of the earlier part of the conversation, which we assume is stored semi-permanently, unlike the domain focus structure, which is just a form of scratch-pad for developing the semantic representation.

The difficulty humans have in thinking about several unrelated things at once would arise from a need to switch back an forth between diverse domain focus sets. Whether this structure is accessed on a hierarchical basis or as a single set is not clear from experimental or introspective evidence (see for example Collins and Quillian, 1972). A hierarchy seems likely to be organized as two or three levels (recent, older, and almost forgotten, say), rather than as a strict stack-of-bounded-size. There is, after all, little reason to think that the previous sentence is more (or less) relevant to understanding the current sentence than is the last sentence but one: frequently utterances build meaning from diverse sources, each of which may, in effect, contribute its own sentences to the utterance.

Domain Focus Calculations

To extract domain focus information from the incoming stream of words, a good scheme may be to calculate the set of 'most popular' modules. A method for doing this would be to have the words vote for the modules they are related to, and including in the context set for the sentence all those modules obtaining more votes than some cutoff. This cutoff would need to vary with the length of the sentence, and other factors may help decide whether to include certain modules, such as whether previously mentioned modules are still accessible, and whether words in the sentence will be undefined unless a module referring to them is introduced. Indeed, word sense ambiguity might be introduced by adding new modules to the domain focus, and this might be a good reason for deciding against doing so in a particular case.

4.2. A Simple Context Algorithm.

In order to demonstrate some of the ideas on context just explained, we shall apply the basic voting algorithm to an example text. This will also demonstrate some flaws in this over-simple method, which we shall then try to correct. The context set for the example consists of names like **indoor, food** and **people**.

A major feature of the design of the context algorithm is the data structure on which it operates. This consists of labels on each word in the global vocabulary, which list those modules to which the word belongs. In the case of words like *the*, *one*, *thing*, and *of*, which would normally give little information about context, the label is an empty list. (Other words like *well* are awkward; *well* could be the adverbial form of *good* (and so contribute little), or the noun which one hopes is full of water (or oil), and so quite useful for determining context. Syntax is important for domain focus calculations.) The role of the algorithm, then, is simply to inspect these labels, and then, at the end of each sentence, pool the votes for the module names, and announce the modules with the most votes.

So each vocabulary item in the text shown below must be assigned a module set. Thus *nursery* suggest **indoor** and **infancy** module, and *buttered, scones,* and *milk* all suggest **food** module. Common words like *in, the,* and *and* are assigned a null module set, as mentioned above.

Text for Analysis

> After learning their lessons in the nursery, Steven and Jennifer went into the kitchen. They asked the cook for some buttered scones and milk. 'You will have to ask your mother about that,' the cook replied. 'I cannot be spoiling your appetites with snacks so close to meal times.'

The analysis corresponding to the text above would have been as follows:

Sentence: after learning their lessons in the nursery steven and jennifer went into the kitchen
Counts: **food** 1; **moving** 1; **time** 1; **indoor** 3; **infancy** 1; **people** 3; **education** 2
Likely modules: **people**; **indoor**

Sentence: they asked the cook for some buttered scones and milk
Counts: **fluid** 1; **food** 4; **talk** 1; **people** 2
Likely modules: **food**

Sentence: you will have to ask your mother about that the cook replied
Counts: **food** 1; **people** 4; **talk** 4
Likely modules: **talk**; **people**

Sentence: i cannot be spoiling your appetites with snacks so close to meal times
Counts: **time** 1; **food** 3; **talk** 2; **people** 2
Likely modules: **food**

4.3. Enhancing the Domain-Focus-Tracking Algorithm.

A word in several modules contributes to the probability of each of these modules. This seems undemocratic - an ambiguous word gets more voting power. This raises the question: do some words have more semantic content than others? It seems likely that for context purposes, a word like *black* is less important than a word like *dog* which in turn is less important than the name of a particular (known) entity (such as *Lassie*). Module votes should be weighted to take this into account.

Neither **people** nor **talk** is a particularly good module to understand the third sentence in, although neither is terribly bad. The sentence is really about authority in adult-child relationships, and the module classification used was not sophisticated enough to notice this. This leads to another point: **talk** occurs frequently in the module counts, and in fact the point of the sentences in question is not usually the talk, but what is said and who says it. The role of the **talk** module expert, then, will be an ancillary one: it is there to connect the speakers and some attributes of their speech (such as *reply, shout, plead*) to the things they say.

That is one view of **talk**. Another view, and perhaps one of equally wide application to modules with similar properties, is that the sentences are not really about **talk** in any important way, and that the context-finder should recognize this. The problem seems to arise because in sentences which report speech have a lot of words which reflect this fact. All first and second person pronouns do this, for example. So perhaps words in 'noisy modules' to module should have lower weighting on their vote for that particular module, in order to prevent those modules from dominating the context finder's results.

So far, we have treated domain focus as a constraint on meaning at the sentence level. Another way to enhance the algorithm is to make it more local in its effect. In *The armor-clad hero struck the evil goblin with his gleaming sword*, each of the three noun groups falls naturally into a separate module (say **warrior, monster**, and **weapon** respectively), and the context might as well be fine-grained enough to take this into account. One could have distinct domain focus sets for agents, objects, instruments and other cases. As there is no need to abandon sentence-wide context completely, in favor of local noun group focus, the various noun group context focus sets should be seen as *preferred context sets*; one would fall back to the wider-scoped focus sets when no meaning could be obtained from the local focus sets.

Word-Sense Ambiguity

Charniak (1983) cites the sentence

The astronomer married a star

as an example of a sentence likely to give problems to humans. The problems with the sentence are obvious, but we shall list them for ease of reference:

* *astronomer* activates **astronomy** module.* On the least effort principle, since it is possible to interpret *star* in this domain, *star* is likely to be interpreted as a celestial object, not an entertainer.

* Very strong expectations associated with *marry* request that its object be human. Entertainers are frequently human, celestial objects are not. Compare Wilks' theory of Preference Semantics (see his 1978 paper, for example).

Later during semantic analysis, the problem with marrying a celestial object becomes evident, and we go back and find a definition of *star* compatible with *married*.

How can we draw in some of the Preference Semantics ideas to attack this problem without the need to backtrack? The way around this problem is to stipulate that the verb *married* contributes a vote for **human** in the local context set of its object, the noun group *a star*. This is pooled with the votes from *star* itself,

* or perhaps a module stack with **astronomy** on top and **human** underneath

resulting in resolution of the *star* as the entertainer variety.

Domain Focus and the Arguments of Charniak and Swinney

As promised earlier, we shall now square the concept of domain focus with the arguments of Charniak (1983) on the basis of the experimental data of Swinney (1979). The gist of the Swinney experiments is as follows. The word *bugs* is ambiguous - insects or spying devices. Swinney presented sentences involving *bugs* to subjects, preceded by disambiguating information such as *cockroaches and.* The response time to a word such as *ant* presented shortly after the *bugs* is thereby reduced, compared with the response time with *ant* by itself. However, and more unexpectedly, the response time to a stimulus related to the other sense of *bugs,* such as *spy* is also reduced. One conclusion that could be drawn from this is that, in humans at least, the access pathway to information about both **espionage** and **insects** is in some way activated by the mention of (unambiguously insectoid) *bugs.* However, one has to remember that in humans both pathways could be activated in parallel. This might be an automatic consequence of the fact that a human context tracking algorithm had referred to **espionage** and **insects** in its calculations. In a serial machine, this would not be efficient. **Espionage** and **insects** would be referred to in the domain focus calculations. This would not speed up access to information in the **espionage** module in a computer implementation, as the information in the module is presumably not retrieved from secondary memory if the vote comes down in favor of another module (**insects** in this case). So it boils down to differences between the architectures of human brains and computers.

Effects of Domain on Semantic Actions.

Our commentary so far has not given any deep consideration to the effects of the need to take semantic actions in particular modules on the need to have a particular module expert 'resident' at any given time. As an example of what this might mean, consider the following phrase:

Kit and Diana went into the garden.

In this case, the meaning would be different, at least for a human natural language understander, depending on the ages and health statuses of Kit and Diana. If, for example, Diana was known to be a paraplegic, or Kit a baby whose locomotion consisted of crawling, then humans would tend to visualize the scene in a way which would differ from the case where both were healthy adults. Hence it would probably be necessary to bring in a Kit expert and a Diana expert at the very first reference to these entities. And if there was no Kit or Diana expert in existence at the start of a natural language text, then it might be necessary to create them on the fly.

Continuity Requirement

If a word is interpreted in module x in some sentence, then it should normally be interpreted in module x if it occurs in the next sentence. An exception to this would be utterances which discuss meanings of words.

4.4. Context Tracking in Humans.

If we assume that humans use a method similar to context tracking in their natural language understanding, then it seems likely that different human beings interacting with a computer system for natural language understanding would have different compartmentalizations of the subject matter being discussed, and different degrees of compartmentalization. This would very likely cause problems for man-machine interaction in case of a serious mismatch, as either the human, or more likely, the machine, may 'thrash' in attempts to maintain the correct context. The problem would be likely to occur whether the understander uses a context finder or not, and indeed it might be easier to compensate for such problems in a context-based system.

It seems plausible to suggest that this problem also occurs in human-human interactions already, though to a lesser degree because of the sophistication of the human natural language understanding system.

Humans talking have an advantage over machines reading typed utterances in deciding context. This is because oral utterances may encode information about changes in context in the form of inflection, pause, stress, volume, pitch, or, in suitable situations, facial expression and gesture. No human is ever likely to be any doubt that a change of context is appropriate when someone first screams 'Fire!'

Some of this information is encoded in large bodies of text in the form of paragraphing and sentence break-up, and, in some cases, in exclamations such as *'Fire!!' he screamed hoarsely* and would thus be available for machine understanding.

4.5. Summary of Context Tracking Algorithm

We summarize the components of a context tracking algorithm in Tables 1, 2, and 3. Some items of knowledge are expressed in more than one way; for example as special-case caveats in the use of various sources or structures, and also as general rules. Table 1 describes the underlying structures used for context tracking, and the basic voting and weighted voting methods. Table 2 lists heuristics to try when the basic method proves inadequate (or as refinements to the basic method). Table 3 contains an initialization rule, and rules for manipulating the heuristics of Table 2.

TABLES 1, 2 AND 3 ABOUT HERE

5. Conclusions

Context tracking, by maintaining an idea of which modules of knowledge are currently relevant, is an approach to reducing the problems of word-sense ambiguity, to managing and organizing the vast amounts of semantic linguistic knowledge needed for natural language understanding systems, and for regulating access to that knowledge by the system.

It also offers the prospect that compatibly designed single domain systems might be incorporated with relative ease in systems whose areas of expertise are potentially much larger than a single domain of discourse.

References

Bourne, L.E., Dominowski, R.L., and Loftus, E.F., *Cognitive Processes*, Englewood Cliffs, New Jersey: Prentice-Hall, 1979.

Bobrow, D.G. and Winograd, T., An overview of KRL: a knowledge representation language, *Cognitive Science* **1** (1977) 3-46.

Bransford, J.D., and Johnson, M.K., Contextual prerequisites for understanding: some investigations of comprehension and recall, *Journal of Verbal Learning and Verbal Behavior* **11** (1972) 717-726.

Charniak, E., On the use of framed knowledge in language comprehension, *Artificial Intelligence* **11** (1978) 225-265.

Charniak, E., A common representation for problem-solving and language-comprehension information, *Artificial Intelligence* **16** (1981) 225-255.

Charniak, E., Passing markers: a theory of contextual influence in language comprehension, *Cognitive Science* **7** (1983) 171-190.

Collins, A.M., and Quillian, M.R., Does category size affect categorization time? *Journal of Verbal Learning and Verbal Behavior* **9** (1970) 432-436.

Freedman, J.L., and Loftus, E.F., Retrieval of words from long-term memory, *Journal of Verbal Learning and Verbal Behavior* **10** (1971) 107-115.

Loftus, E.F., Activation of semantic memory, *American Journal of Psychology*, **86** (1973) 331-337.

Loftus, G.R., and Loftus, E.F., The influence of one memory retrieval on a subsequent memory retrieval, *Memory and Cognition* **5** (1974) 467-471.

Small, Steven, Viewing word expert parsing as linguistic theory, *Proceedings of the Seventh International Joint Conference on Artificial Intelligence* (1981)

70-76.

Swinney, D.A., Lexical access during sentence comprehension: (re) consideration of context effects, *Journal of Verbal Learning and Verbal Behavior* **18** (1979) 645-659.

Walker, D.E., (Ed.), *Understanding Spoken Language*, New York: North-Holland, 1978.

Wilks, Y., Making preferences more active, *Artificial Intelligence* **11** (1978) 197-223.

Winograd, T., *Understanding Natural Language*, New York: Academic Press, 1972.

Table 1: Structures Used in Context Tracking Algorithm

Structure	*Method of Use and Comments*
List of relevant modules for each vocabulary item	Words in a sentence or phrase vote for the most popular modules. Voting needs to be weighted against "noisy" modules, for significant words
Context focus set	Only the knowledge in modules in this set is available for use in understanding the sentence. The structure is actually a stack (with limited depth) of sets of modules; if no understanding is possible with modules in top set, then try lower sets.
History of context	Focus set should retain modules from previous sentences or phrases for a limited period of time, unless there are cues suggesting a major topic change.

Table 2: Knowledge Used in Context Tracking Algorithm

Knowledge	*Method of Use and Comments*
Future contexts	In case of understanding failure, context from following sentences or phrases may be incorporated in the context set for the current phrase, unless other knowledge shows that the two phrases are not related.
Forgetfulness	Unused modules tend to drift out of the context sets. Modules are not "unused" if they are tacitly referred to by using knowledge in closely related modules.
Preference semantics	In phrase by phrase context tracking, the verb's preferences for restricted types of item to fill some of its semantic cases should influence the context set voting for those cases.
Syntactic clues	Determining the syntactic role of a word may help to decide its module vote.
Continuity requirement	Words or noun groups already identified should keep that identity even if the current context set would suggest otherwise, unless understanding failure occurs. This does not apply in utterances discussing the meanings of words
Cues for sudden change of context	For example, some phrases like "Let's talk about ..." may signal a change in context. Speech understanding systems can have extra cues from tone, inflection etc.

Table 8: Rules Used in Context Tracking Algorithm

Rules	*Method of Use and Comments*
Initial context rule	Obtain initial context by analysis of title of passage or first sentence or two. Sometimes passages start with a minor theme before flowing on to major theme(s).
Don't-try -too-hard rule	If understanding is possible with the current domain focus, then do not change that domain focus.
Try-another rule	If understanding does not follow from the use of one of the above knowledge sources, then try another
Backtrack rule	If understanding does not follow after all heuristics above have been tried, then backtrack, ignoring the don't-try-too-hard rule