

AD-769 673

NATURAL LANGUAGE INFERENCE

Yorick Wilks

Stanford University

Prepared for:

Advanced Research Projects Agency
Department of Defense

August 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

DISCLAIMER NOTICE

THIS DOCUMENT IS THE BEST
QUALITY AVAILABLE.

COPY FURNISHED CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

AD 769673

STANFORD ARTIFICIAL INTELLIGENCE LABORATORY
MEMO NO. AI-211

AUGUST 1973

COMPUTER SCIENCE DEPARTMENT
REPORT NO. CS-383

NATURAL LANGUAGE INFERENCE

by

Yorick Wilks

ABSTRACT: The paper describes the way in which a Preference Semantics system for natural language analysis and generation tackles a difficult class of anaphoric inference problems (finding the correct referent for an English pronoun in context): those requiring rather analytic (conceptual) knowledge of a complex sort, or requiring weak inductive knowledge of the course of events in the real world. The method emp (g) converts all available knowledge to a canonical template form and endeavors to create chains of non-deductive inferences from the unknowns to the possible referents. Its method of selecting among possible chains of inferences is consistent with the overall principle of "semantic preference" used to set up the original meaning representation, of which these anaphoric inference procedures are a manipulation.

This paper owes much to discussions with Annette Herskovits, Bill Strapson, Ken Colby, Bruce Anderson and Horace Enea. The mistakes are all mine of course.

The views and conclusions contained in this document are those of the author and should not be interpreted as representing necessarily the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U. S. Government.

This research was supported by the Advanced Research Projects Agency, Department of Defense (SD 183), USA.

Reproduced in the USA. Available from the National Technical Information Service, Springfield, Virginia, 22151.

20

Reproduced by
**NATIONAL TECHNICAL
INFORMATION SERVICE**
U S Department of Commerce
Springfield VA 22151

INTRODUCTION

The paper describes inferential manipulations of a representation of the meaning structure of natural language. It differs from previous descriptions (1, 2, 3, 4) of this semantics-based system, which have concentrated on the representation itself, and above all on the procedures by which the representation is produced from input sentences and paragraphs in English. In this paper I assume that structure of representation, except for a brief recapitulation, and concentrate on operations upon it for the solution of a class of difficult anaphora problems.

The system described is part of a running system for understanding and translating natural language on the PDP6/10 at the Stanford A. I. Laboratory, programmed in MLISP and LISPL. 6. I shall not in anyway stress the translation-into-French aspect of the work, but its presence provides a continual empirical check of the adequacy of the inference and "understanding" described here.

The earlier emphasis on the construction of the linguistic base is, I believe, fully justified. The present system is, to my knowledge, the most comprehensive producer of meaning structures for unadorned natural language available at present in terms of implementation, vocabulary, disambiguation of many-sensed words and references, etc. Moreover, as I have argued elsewhere (2), it is not the implementation of a conventional theory from linguistics, but is one with somewhat different principles of content.

In what I call its basic mode, the system already resolves anaphoras depending on superficial conceptual content of the text words. This is done in the course of setting up the initial representation. I shall call these type A anaphoras. For example, in "Give the bananas to the monkey, although they are not ripe. They are very hungry", the system in its basic mode would decide that the first "they" refers to the bananas and the second to the monkeys. It can do that simply from what it knows about monkeys getting hungry because they are animals, and bananas having phases like ripeness because they are plants. All this information is, one might say colloquially, part of the superficial meaning of "banana" and "monkey".

This paper describes an "extended inference mode" of the system that tackles the other kinds of anaphora example that I shall call types B and C. Consider the correct attachment of "it" in "John drank the mineral from the glass, and it felt warm in his stomach". It is clear that the pronoun should be tied to "whisky" rather than "glass", but how it is to be done is not immediately obvious. Analysis of the example (see below) suggests that the solution requires, among other things, an inference equivalent to the sentence "whatever is in a part of X is in the X".

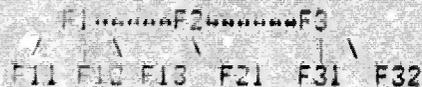
Anaphoras like the last I shall call type B, in that the inferences required to resolve them are analytic but not superficial. By analytic I simply mean that the sentence above, about parts and wholes, is logically true and not in any clear sense a fact about the real world (but rather about the meanings of words). What is meant by "superficial" in the distinction between types A and B will become clear after a discussion of the meaning formalism employed here.

I shall also discuss type C anaphoras, which require inferences that are not analytic at all, but weak generalisations (often falsified in experience) about the course of events in the world. Yet their employment here is not in any sense a probabilistic one. In "The dogs chased the cat, and I heard one of them squeal with pain", we shall, in order to resolve the referent of "one" (which I take to be "cat" and not "dog"), need a weak generalisation equivalent to "animate beings other than other animate beings may be unpleasantly affected". Such expressions are indeed suspiciously vague, and a reader who is puzzled at this point should ask himself how he would explain (say, to someone who did not know English well) the way he knew the referent of "one" in that sentence. It can hardly be in virtue of a probabilistic fact about cats and dogs because the same generalisation would apply to whatever was chasing and being chased. I doubt we could do it if he does not come up with something very like the inference suggested, and it may be the nature of natural language itself that is troubling him.

The inferences for type C, then, are general expressions of partial information (McCarthy's phrase) and are considered to apply only if they are adequately confirmed by the context. What I mean by that will become clear in the course of what follows, but in no case do these expressions yield deductive consequences about the future course of the world. Indeed, they would be foolish if they did because the world's course cannot be captured in that way. In the unhappy example above, it might have been his earlier dinner that made him feel good.

BRIEF REGARD OF THE SYSTEM'S BASIC MODE OF ANALYSIS

In its basic mode, the system fragments texts (into phrase/clause like items) and attaches a template to each. A template is a canonical form of connectivity of semantic formulas as follows (where a formula is a complex item, to be described, corresponding to a sense of an input word):



F1, F2, F3 are the principal formulas of the template and are always agent, action and object (in that order), though any of them may be a dummy in any particular example. (F11, F12, F13) is a list of formulas dependent on main formula F1 etc. It should be said, in view of other current uses of "template", that it is not a surface form at all, but a format underlying meaning representation. Moreover, it does not function within a crude pattern matching technique, such that if some text fragment has no templates matching it, it is thrown away, as it were. Special routines are called in such situations to construct an appropriate template item. All this basic material is set out in earlier papers.

The structure of formulas is explained below in some detail. In brief a formula is a nested list (a binary tree in fact) of semantic primitives called elements (expressed here as LISP atoms). Each such formula expresses a sense of at least one English word.

Let me give an example of a template structure at this point by using the following simplifying notation: any English words in square brackets F1 stand for the meaning representation of those words in the Preference Semantics system. This device is important in the exposition of the material in that the content of the coded forms can be seen immediately, whereas the complex coded forms themselves would be as hard to read as a text, a sentence read a word at a time. But it is important to state that the rules and formalisms expressed within () are really structured primitives, and that their tasks could not be carried out, as some still seem to believe, by massaging the language words themselves to stand for their own meaning representation.

So then, the template for "The black horse passed the winning post easily" could be written (ignoring any ambiguity problems for the moment):

```
[horse]====[passed]====[post]
  ↑           ↑           ↑
 [the black] [easily] [the winning]
```

If any or all of the agent, action or object formulas are missing, the complete code(s) is filled if with a dummy element DTHIS. Thus a template for "The old house collapsed" could be written:

```
[house]====[collapsed]====DTHIS
  ↑
 [the old]
```

In the case of structures like prepositional phrases, we consider the preposition represented as a pseudo action, and the whole template as having a dummy Agent. Thus for "at the Derby", we have:

OTHIS[...](a)[...](Derby)
↑
(the)

The representation of a text (composed of fragments) is then a network of these template networks. The templates are interconnected by case ties. The notion of case is discussed in more detail below, but for the moment a case can be thought of as tying one template to some particular node in another template by a link of a certain type, namely the case type, which specifies the sort of dependence the former template has on the latter. In the sentence "He lost his wallet in the subway" (fragmented at the stroke) we might say that the second fragment of the sentence depends on "lost" in the first, and that the dependence is the locative case. Thus in the representation, the template for the second fragment would be tied to the central, action, node of the first, by a link labelled LOCA. The node on the first template to which the case tie ties is called the marking of the second template.

Type A anaphoras dealt with adequately within this framework, and therefore the purpose of this paper is not to describe the basic mode of operation of the system) because we get a denser network of links by considering the formula for the appropriate referent substituted for the pronoun prouder than with its rivals. A link is reinforced, or strengthened, in the network, when a preference is satisfied. So, if we think of the formula for "ripe" as expressing a preference for application to plants, we see why a denser network arises in the above example for correct solution, rather than for one equivalent to "ripe monkeys". The way in which formulae express preferences of various sorts is described below.

First, solved, these type A anaphoras also constitute links between templates, from the pronoun variable to its correct referent. Thus the completed first form of the whole representation obtained from the parser looks as:

(CA) MARK ANAPHORA F1 F2 F3 (F1 dependents) (F2 dependents) (F3 dependents)

The structural form of representation of a text paragraph is called its IREP (Initial IREPresentation), or "semartic block". No emphasis has been placed here on syntax analysis of the input, and a reader who consults (1, 2, 3 or 4) will see that all of conventional grammar analysis has been done in the course of setting up this form of representation. An example of such an IREP, for the monkey-banana example, is given below as computer output.

Having sketched in the basic mode and its representation, we can now get to our muttons and sketch in the extended inference mode that is the heart of this paper.

21. (whisky (DINE FROM) a+glass)

22. (IN IN) (in) his+stomach)

So, in this infernal representation we have acquired new template-like objects that express, in canonical form, new analytic information extracted from the existing templates, and from which new inferences can be made. It is postulated that the generation of this recipient information from the deeper levels of the formulas is essential to the process of understanding. These new forms differ from standard templates only in that their second node, or pseudo-action, has had a case name CONSd onto whatever the node was formerly. Note here that the form (IN IN) is not redundant since the IN locates the action precisely as containment, while the English preposition can indicate many cases other than containment, as in "in five minutes".

We now describe how these particular extractions were made, even in the absence of any detailed knowledge of the structure of formulas:

11. T_{11} was derived from the template for "John drank the whisky" because from the structure of the formula for "drink" it follows that the liquid drunk is subsequently inside the drinker. This is because, when making up the formula for the action "drink", we express in it that the action consists in causing a liquid to be inside the agent of the action.

Analogously, T_{12} is inferred because the same formula specifies that the liquid enters the drinker through a specific part of the drinker (his mouth, if you please).

T_{13} is inferred from the direction case of the second template, whence we know that it was the whisky that was moved from the glass.

Lastly, T_{14} is inferred from the direction case of the fourth template, because whatever the referent of it is, it is also in John's stomach.

Let us see where we are: we have obtained new template items that grade relative information, but did not appear in the original text. (As we shall see in the detailed treatment below, some of the above are obtained from extraction, more strictly defined, and some from what I shall call "repacking the semantic block".)

In the pool for inference procedures we now have the original templates that mention either the variable pronoun or the possible "answer" referents, plus the extractions. We also have access to an inventory of Common Sense Inference Rules (CSIRs) which are of the form (T1 - T2), where T1 and T2 are T-forms, that is, templates or extractions.

We now try two strategies in turn: first we try a zero-point strategy, and it is to try to identify an answer template (or extraction) and a variable template (or extraction) without the use of any rules.

The general assumption here, and it is a strong psychological assumption, is that in order to resolve these painful ambiguities the understanding system is going to use the shortest possible chain of inferences, if any. And a zero-point strategy will, as it were, have no inferences (in the terms of a chain of CSI inferences) and so if it works, it will provide the shortest chain.

This strategy is adequate for the example under discussion, because in the first step (the initiation of matching) identify extractions of "and" and "it" and identify "it" and the whisky, and we are home. This is the resolution of a B type anaphora, requiring only analytic, non-inductive, true conceptual information. It should be noticed that some B type anaphoras (defined earlier by the need of weak inductive information for their solution) can also be solved by the zero point strategy, because some extractions, and in particular those from the goal side (see below), are inductive and non-analytic.

If the zero-point strategy fails, we bring down all the CSI rules that contain in action subformula occurring in an answer or problem formula in the proof, and attempt to find the shortest chain that leads from some answer to some variable.

Thus, in the sentence "The soldiers/ fired+at the women/and I saw several fall", we extract a form equivalent to [soldiers strike women], since we can tell from the formula for "fired+at" that the action is intended to strike the object of the action. We are seeking for a partial confirmation of the assertion [?several fall DTHIS], and such a chain is completed by the rule [X strike Y] → [Y fall DTHIS], but we need a more rule equivalent to, say, [X strike Y] → [Y die DTHIS], since there is nothing in the sentence as given to partially confirm that rule in a chain, and cause it to fit here. Since we are not really dealing with subformulas in the statement of the rules, rather than natural language words, "fitting" means an "adequate match of subformulas".

It is conceivable that there would be an, implausible, chain of rules and extractions giving the other result, namely that the soldiers fire: [soldiers fire DTHIS] ∧ [X fire DTHIS] → [Y fire+at X] → [X strike Y] → [X fall DTHIS] etc., based on the assumption that though "that fire guns get fired at (...the who lives by the sword shall perish and...)" . But such a chain would be longer than the one already constructed and would not be preferred.

MORE ON THE BASIC MODE

Formulas

Formulas are structures corresponding to senses of words, expressing their meanings. Much of the body of this paper is concerned with the construction of such structures, and the extraction of information from them. It is important to have some general idea of their construction and interpretation.

Formulas are binary trees, expressed as lists, of semantic elements, denoted by right and left brackets. The elements are either case-envelopes or functions such as CAUSE, STRIK, CHANGE, or items such as MAN, WRAP, WRAP, WRAP, as examples here, element names that are usually Anglo-Saxon monosyllables, but there are about 1000 of them. There are also elements like KIND indicating qualifications, and elements (indicated by an initial *) that stand for primitive elements, such as *AN (animate) to cover MAN, BEAST and other animals. In addition, most elements have a negated initial, which is the element name. I assume here that the use of negated primitives of this general sort, that are not logical primitives, need no special defense at this point.

The head of an element in a formula is its rightmost, called its head. This is the most general sort of item, or action, or type of action, which is expressed corresponds to: for example, any word representing a human being will have MAN as its head.

Formulas are binary trees of unlimited depth, they can be decomposed into pairs of elements and subformulas, down to the primitive elements. This process is called building up the formula, or decomposing it into primitive elements. At each stage there is a dependence of the left element on the corresponding right: this dependence is

Formulas are binary trees of unlimited depth, they can be

decomposed into pairs of elements and subformulas, down

to the primitive elements. This process is called building up

the formula, or decomposing it into primitive elements. At each stage there is a dependence of the left element on the corresponding right: this dependence is

Formulas are binary trees of unlimited depth, they can be decomposed into pairs of elements and subformulas, down to the primitive elements. This process is called building up the formula, or decomposing it into primitive elements. At each stage there is a dependence of the left element on the corresponding right: this dependence is

Formulas are binary trees of unlimited depth, they can be decomposed into pairs of elements and subformulas, down to the primitive elements. This process is called building up the formula, or decomposing it into primitive elements. At each stage there is a dependence of the left element on the corresponding right: this dependence is

container, since WRAP is always a qualifier when in the left hand, dependent, position in a pair.

An important notion is that of the semantic preferences that formulas can express. Consider the formula for "grasp" of objects:

```
"grasp" (action) = ((*AN) SUBJ)((*PHYSOB OBJE)((THIS (MAN  
WRAP)) INST (LOCAL SENSE))))
```

In the case of agents, SUBJ and OBJE occur at the top levels on the left of the formula, and at that level in an action formula they express the preferred agent and object of the action concerned.

Thus, grasping, in this sense is an action preferably done by animate beings (*AN) to physical objects (*PHYSOB), and consists in an act of grasping by touch, and done with an instrument (INST is the case element) which is a part of the body. When I say "prefer" here, I mean that, if the preferred agent or object cannot be found, a template is constructed with whatever is available. Thus, "The robot grasped the block" would never be rejected; it would only be rejected if there were a better interpretation than any possible competing interpretation that had an animate agent. I have argued in (2) that this approach to rules of formation has unexpected consequences for linguistic theory.

We would also expect another formula to be available for "grasp", one such as:

```
"grasp" (action) = ((*HUM) SUBJ)((STGN OBJE)(TRUE THINK))
```

In this case, we have an action, preferably done by human beings to grasp (understand) ideas, principles etc), namely of thinking them to be true, or adequate, signs.

The introduction of the basic mode always fit this last formula into a complete structure for "He grasped the principles", and the other formulas for "grasp" into the template for "the boy grasped the toy", by means of the preference and semantic density techniques described in earlier papers. These preferences for agent and object are part of the "supra-lexical conceptual information" referred to earlier.

A few other rules will help to clarify the notion of "knowing our way toward a formula" when interpreting it:

Agents (and implicit ones not be specified by SUBJ case) never attach at the top level in an action formula as described above, or they attach to the head of a formula, as in:

```
"not please" (action) = (NOTPLEASE FEEL) (SUBJ MAN)
```

Here, the normal priority of agents being to the left of (= dependent of) the corresponding action, is violated, since MAN is the agent for FEEL, while at the same time being the head of the whole formula.

This violation of order in search is indicated by also violating the order restriction that normally makes the SUBJ case element the dominant (initial hand member) of the pair in which it occurs. The corresponding rule of analysis is: "On encountering SUBJ as dependent, expect action for the agent to follow to the left".

There is a secondary, but never implicit. Moreover, an object is considered a subject of all actions to its right. This enables us to express the important notion of real and apparent agents of actions. An example is:

"fire+at" (verb+particle)

(AIM) (ON) (X) (AN) (OBJECT) (STRIK) (GOAL) ((THING MOVE) CAUSE) (IT)

This relation (fire preferably by human beings to animate beings) is one of aiming a thing to move (the bullet) with the aim (GOAL case) of striking something. Since XAN is the object of all actions to its right, it is the subject not only of CAUSE, but also of STRIK. Hence, the striking is done by the same animate being. Moreover, THING (the bullet) is internally the agent of MOVE, not the object of CAUSE, which is correct as far as the meaning of "fire+at" is concerned.

At present, we operate with a distinction system of ten cases, which are listed below, together with (in capital letters) the semantic meaning of each case, the questions that define them, and examples of subformulas expressing them. Defining a case is a tricky matter, but the distinction method is reasonably adequate. Note that the subformulas, examples are of those parts of a formula that would express that notion AS PART OF THE MEANING OF A WORD. The subformulas are not, of course, how the system would express the quoted words if encountered in a text, when they would be represented by a template.

recipients: FOR "for a woman" → ((FEM MAN)FOR)
what/why to? what/who for?

instruments: INST "with a stick" → ((LINE THING)INST)
what with? by what means?

directions: x[IB] (see below), TO, FROM, UP
"from the top" → ((UP POINT)FROM)
where to where from? at what? out of where? by what?

possessives: POSS "owned by a man" → ((MAL MAN)POSS)
who owns the thing mentioned?

location: LOCA "at that time" → ((THIS(WHEN POINT))LOCA)
when? where? where at? by what? in what time? near what? at what
time? during what? before when?

in what? "in a glass" → (((FLOW STUFF)OBJE)WRAP)THING)IN)

source: SOURCE "out of wood" → ((PLANT STUFF)SOUR)

goal: GOAL "to strike a woman" →
(((FEM MAN)OBJE)STRIK)GOAL)
to what end? for what purpose?

with: WITH "without a glass" → (((FLOW STUFF)OBJE)WRAP)THING)NOTW)TH)
accompanied by what/who? with what/whom? without what/whom?

subject: SUBJ who did this?

object: OBJ who/what was this done to?

Certain of the above have negative forms leading to additional
elements: NOTOBJ, NOTPOSS, NOTIN, NOTWITH.

These elements have no functions, and occur in two sorts of
constructions: formulas and IREPs. In formulas they express part
of the structure of a word sense. Thus in

"drink" → ((FLOW STUFF)OBJE)WRAP)THING)IN) ((FLOW STUFF)SOUR)THING)

the word "drink" has a liquid source (FLOW STUFF), and is in a
container (WRAP)THING) the other function of these elements is, as
already explained, the name of the tie between the template for some
fragment and a part of another template

Finally, the class of direction case elements (TO and
FROM) and POSS occur only as the indicator of the case of a fragment,
never in formulas. Conversely POSS occurs only in formulas, never as
the indicator of a fragment case.

Finally, "and" is only included in a formula when it is specific:
the aspect of the case is involved. In the formula
for "point", for example, we include a direction specification for
movement (POINT)TO). However, in the formula for "move" we do
not include the element TO or FROM, even though movement must in fact
be of some direction, since we have no reasonable expectation about it
in the case of "point". Sentences containing "move" may very well go
on to specify the direction involved, but its association with "move"
is completely arbitrary, and we cannot expect any confirmation of
expectations that would, say, resolve ambiguities. In this respect

THE IMPLEMENTATION OF THE EXTENDED MODEL

There are three parts to the extended inference mode: the REPACK routine that takes the IREP block and repacks it; the EXTRACT routine which produces extractions, new knowledge not explicit in the text and the INFER which tries to link an answer T-form to one of the free problem variables, that is, a text pronoun giving the problem.

The REPACK routine.

It attempts to replace dummy nodes in the IREP wherever possible before handing the whole representation to the extraction procedure. The replacement is itself a complex form of inference, sometimes as complex as the inference routines on which we are concentrating here. However, there is no pretence that these procedures are mutually independent and done in an ad hoc fashion, so we may legitimately regard it as substituted for one a single process in this way. The difference of activity resulting done by REPACK, in the construction of a new block IREP' from IREP, varies with particular action cases.

If we look back at the informal extractions done from the "John drank the whisky, ..." example, we will see that the new T-forms 21 and 22 are actually obtained by filling in a dummy agent in some template form to make it another template. Thus from [2 DIRE ; DTHIS from 21] we obtained the new T-form, numbered 21, [whisky (DIRE from) 21] which was done by filling the dummy agent node of the template for "with a glass" with the formula for "whisky", and shifting the direction case marker into the pseudo-action. This is repacking, not an extraction proper, since the T-form obtained simply replaces a template "assertion" already in the representation. As we shall see, a true extraction is a new T-form altogether.

Let us now distinguish replaceable and unreplaceable cases roughly as follows. The dummy agent in the second (instrumental) template for "He hit his father / with a club" cannot be replaced to yield any form equivalent to [father (INST with) a club]. So we may say that the instrumental case is unreplaceable. But the dummy agent in the second (recipient) template of "He bought the flowers / for his mother" can be replaced to yield a form for [flowers (RECIP for) his mother], and the recipient case is replaceable, and is replaced by the operation of REPACK.

At the top level REPACK can be written in LISP as:

```
REPACK(IREP) - IREPR
```

The general routine

IREPR takes each T-form, or template-like item, in IREPR in turn and produces, if necessary, in a new block IREPE, better called T-forms extracted from it. At present, extractions are only made from templates that contain either one of the possible actions, or one of the variables of the problem. The former are templates containing a formula for a word on the list ANS, the latter are templates, one of whose nodes is (QUERYMARK THIS). Any templates not containing either an action or a variable are simply transferred unchanged from IREPR to IREPE.

Thus, the general form of the extraction routine at the top level is

EXTRACT (IREPR ANS) → IREPE

Taking each template in turn, we first consider those processes that modify it, and then those which produce new T-forms from it. In the first instance, some manipulations to do with negation, and with the substitution of cases.

If the agent or object formula is negated, the negative item in its formula is removed and the head of the corresponding action formula is negated, to give all the subsequently applied inference rules automatically negate the action. Thus, in \bar{A} notation, we would achieve by this procedure the coded equivalent of

[John drank no+gin] → [John not+drink gin]

Each agent and object formula is then scrutinized by the question "does it satisfy the preference expressed by the corresponding action. If it does not, does any "of-phrase" qualifier of it do so to read". If so, replace the agent or object by that "of-phrase" qualifier as the true agent or object.

Thus:

[John drank of+glass of+wine] → [John drank wine]

[Women drank wine] → [Women drank wine]

The main piece of EXTRACT takes the action formula of a template and moves leftwards through it seeking case heads (other than SUBJ and OBJ). If it finds one, it asks is it replaceable, and, if it is, EXTRACT looks at subsequent fragments to see if REPACK has already replaced it. If it has been replaced it is forgotten, thus avoiding the same case information being extracted twice. It detects that REPACK has made such a replacement by finding the case name itself in the pseudo-action of a succeeding template, and a replaced dummy as the corresponding agent.

With the new case, for example, the dependent of the case element becomes the action of the new T-form. In this case, as with every other, an attempt is made, on finding a potential agent or object for the new T-form at the top level of the action formula of the template under consideration, to identify it with the main agent or object formula of the original template. If this can be done, the agent or object formula of the original template is used, as being more specific. For example, in extracting from the action formula in the template for "John fired at the deer", we find the goal case in [fired+at], with dependent STRIK, which is the attempted action. The object of that action, found to its left, is *AN, which can be instantiated by the formula for "deer" in the main template, namely (THIS BEAST). So the latter is used as the object of the new, extracted, T form [John strikes deer], since "deer" is more specific than "animate being".

For most other cases (agent, direction, location, containment and source): the case element provides the new pseudo-action, and the new subject of it is specified as the dependent of the case element. The new object is found as follows: it is the highest level object actually available in the action that dominates the case (to its immediate right in the formula).

So in the formula for "pour" in "I pour the wine"

```
((AGENT (I)) ((FLOW STUFF) ((TO) ((WRAP THING) TO) MOVE) CAUSE)))
```

we consider (moving leftwards through the formula) the direction case in the sub-formula ((WRAP THING) TO), implying that the (FLOW STUFF) formula, which is the highest level object in the formula, is moved in the direction of a container, or (WRAP THING). The case element TO is dominated by MOVE, whose highest level object is (FLOW STUFF), which would become the pseudo agent of the new extracted T-form, but since it can be identified with the object of the old template, namely "wine", it is, and that becomes the pseudo-agent of the new T-form, since it is more specific than "liquid". It is, that is, (WINE). The container, becomes the object of the new T-form and the dependent case element becomes the action so we get an extracted T-form:

```
(CONTAIN (WINE) ((FLOW STUFF)) (TO) (PDO) (WRAP THING) )
```

where (I) is (I) and (I) is (I) for some+container)

The final routine

This routine has access to the representation IREPE produced by procedure AN and CSIR, the inventory of common sense inference rules. Its form at the top level is

```
INFER(IREPE ANS CSIR) - ANS'
```


The inventory of rules is searched for those containing any action conditions occurring in a T-form in the pool that also contains either an "answer" or a "problem variable". In this case we pull in a rule internally expressed:

[I STRIKE (*ANI 2)] ← [(*ANI 2) fall DTHIS]

Here variables are indicated by numbers ; *ANI expresses a restriction on the variable that any value of it must be animate, and the double ← indicates that this rule can be considered as running in either direction.

This CSIR form is of course a more perspicuous form of:

[I (THIS STRIKE) (*ANI 2)] ← [(*ANI 2) (NOTOP BE) DTHIS]

which, of course covers a wider class of activities than simply the English verb "strike". It would cover at least "hit", "batter" etc.

That a chain of length one is established by the rule from T-forms 12 to 14, since the "animate condition" is satisfied and the variable "x" is identified by the rule with the formula for "women". It should be noted here that the inference rules are very weak in that the application of a rule like the present one is perfectly consistent with the description of a situation where an animate being has struck in some way but does not fall. And this weakness is wholly intentional.

One important inflection in CSIRs is whether or not negation is significant or inert. The negation of the action in a T-form is normally significant. Consider "John drank no gin / in his martini / but it left warm / in his stomach nonetheless". In the template for the first fragment, shifting the negation to the action, and extracting for the containment case from the formula for "drink", we shall obtain a T-form

[gin MATCH BE] container]

and another

[gin MATCH BE] John]

Conversely we shall obtain, by the same method, from the second template [martini (IN BE) container] and [martini (IN BE) John]. In trying to tie only one of these drinks by matching to the extraction [gin (IN BE) John], we shall, without the use of CSIRs be able to get the ANSWER down to a single member, namely [martini], since MATCH will show us that [gin] cannot stay on ANS.

However, if we applied the same analyses to a sentence like "John
saw a car / in the window / and he knew / that he would get it",
the logical content of "it" is "car" and not "window" and he
would not be getting "window" (linking the first and fourth T-forms with some
rule like [(XIII 1) want 2] - [(*ANI 1) have 2]

and, and this is the point, if the same sentence had concluded ". . .
and he knew he would get it" we should have required the same
rule and this rule number, this rule has its "consequent" action
which is the point that question of it is irrelevant to its application.

Also, and this is the point, can be seen to be non-deductive very clearly
and this is the point that it could be said to be of the form $A \rightarrow (B \vee \neg B)$,
and this is the point that any content whatever in a deductive system,
and this is the point that it is not and necessary.

GENERALIZATION AND SOME PARTICULAR COMPARISONS

The system above cannot be considered in any way adequately
adequate, because no one has any very clear idea of what
and this is the point that in this area. But even to qualify, the basic mode
and this is the point that under a considerable vocabulary and range
and this is the point that the extended mode must be shown to be
and this is the point that a sized inventory of CSIRs.

The system above will also be those
and this is the point that a procedure to back up the
and this is the point that to resolve ambiguity within a small
and this is the point that fragments. Ambiguity over a context larger than
and this is the point that just as is ambiguity of the sort discussed
and this is the point that but he should be prepared for it in an adequate
and this is the point that

The system above is a strong, and possibly naive, psychological
and this is the point that chain length is a reasonable metric to
and this is the point that informational interpretations. I think it is
and this is the point that the tension introduced into understanding by
and this is the point that ambiguity has been overlooked. Notice here that
and this is the point that number of CSIRs employed, not counting
and this is the point that different ways of writing down formulas will not
and this is the point that affect chain length.

However, I would justify the principle as being essentially an
and this is the point that called semantic preference (Wilks 2) used in
and this is the point that a representation. That preference was justified as
and this is the point that "semantically deepest" interpretation which was, I
and this is the point that "the least meaning" (in the sense in which a
and this is the point that carries the maximum possible information).
and this is the point that the shortest chain of inferences also minimises the
and this is the point that information flow, and introduces the least extraneous inductive

information into the system. It is clear that such a notion of information based choice is ultimately inadequate. We only have to consider a sentence like "I was named after my father" where it seems clear that we exclude one interpretation simply because it contains virtually no information. This alone shows there must be some qualification to a "minimising information" theory. However, the fact that all available theories are wrong, by no means puts them all in the same position. I think such hypotheses about the overall manner in which an understanding system endeavors to maintain its coherence are well worth making and testing, and that they represent an aspect of human language "competence" almost wholly ignored by current linguistic and artificial intelligence. One could make the point more precisely as follows: virtually all the systems in those areas define "success", that is to say the success of a particular parsed representation with respect to a text. What they do not tell us is what number of successes are registered, as if almost always the only realistic practice. But human understanders do not just accept the first, or opt for the first they find, or pick one at random; they prefer one in particular on some principled basis.

It is for this reason that the subject investigated in this paper cannot be treated in isolation from an adequate linguistic base (which, as some seem to think, The inferring of a correct interpretation is intimately related to the systematic exclusion of competing interpretations, and any system that does not allow itself a multiplicity of sense and structure in at the start can hardly hope to do the job because the difficulty never arises there, but then neither does one essential aspect of natural language either. I have developed elsewhere (6) an abstract view of meaning along these lines: that to have meaning is essentially to have one meaning RATHER THAN ANOTHER. Or, put another way, having meaning essentially involves procedures for the exclusion of alternative interpretations. This, I believe, is the residual truth lurking beneath the "proceduralizing of meaning", a thesis which when taken a face value is patently false.

Let me mention a closely related shortcoming of the micro-world approach to natural language analysis: it concerns what I believe to be the fundamental mistake in AI about the notion of "inference". Let me say this without falling the obvious, not from dry motives of clarity, but because I believe the muddle has important practical consequences in the area of natural language understanding.

There are two references, in the bare sense of that word, of "inference" that people might make from one assertion to another.

(1) All Englishmen are untrustworthy and Cecil is an Englishman, SO he is untrustworthy.

(2) Cecil is an Englishman, SO he is untrustworthy.

(3) This is triangular, SO it is three-sided.

I take it that for a deduction, true in all possible worlds, and words independent of the meanings of the words "Englishman", "Cecil" and "not to be true".

and is an inference, simply and solely, and certainly not valid, whether or not it happens to be true for some English Cecil.

There is a valid inference, true in all possible worlds, as they say, because of the central meanings of "triangular" and "three sided", a fact that is sometimes expressed by saying that the premise missing, for there to be a deduction, namely "all triangular things are three sided" is in fact to be true.

What is the point for our purposes of all this dogmatic and traditional identification? Simply this: the extra-conceptual content, the so-called dictionary, that is CSIR inferences of the sort we have discussed in this paper problems in text, are of type (1). These inferences could function as part of a deductive system by the addition of sufficient inductively unreliable premisses to convert them to form (2). They could then function within established deductive machinery, such as first order PC, PLANNER in one of its modes of operation, etc.

But there may be no need for doing that, at least in the case of natural language analysis, because the conclusions reached can be no more reliable than the dubious generalisations functioning as premisses, whatever the power of the deductive machinery intervening.

In this paper I have described how such weak information can fulfil a problem solving role in natural language analysis, in terms of a notion of "adequately confirmed" inference in context. But that does not require the deductive machinery at all.

The point will be clarified here by noting two research situations where, by contrast, the deductive machinery may pay its way: (1) in logic, and (2) in simulated micro worlds.

In the case of a robot, locally moving about in the world with perceptually mediated information and plans, the world itself can provide a clear sense of contradiction. If the robot's deductions fall at the door is upon, but it bangs into the firmly closed door in fact, then the conclusions are contradicted and the preceding premisses can be discarded, as would be the case with a scientific theory tested by a laboratory experiment. That is to say, the premisses may be discarded, not because there can be contradiction of conclusions by deductive machinery can transfer the "not" back to some premise, as of refuting a scientific theory, though the question of which premise it should be transferred back to is very difficult of course.

This situation is quite different from the analysis of natural language where there is little or no expectation of contradiction. In understanding the text, the understander does not expect to encounter contradictions in the text in the near future. A robot could in principle learn from experiments, and in the case of a dialog in particular, could step back and ask questions of the author, or even understand texts without experimenting on them. The probability of a robot's being able at all to discover such contradiction is a doubtful and hard to trackback having done so, to the next least probable individual. And at the moment no system is in striking evidence of such an ability.

The nature of simulated micro-worlds is different. Here there is no payoff to be obtained, but there is no need for it since all promises made by the system are fulfilled, and no real information can ever enter the system. The robot, after executing the command "Clear off the top of the stack", it is clear, by definition, apart from the possibility of a formal contradiction. No lingering and sticky doubt is allowed to imperil the stability of the house of cards. It will be clear that such situations have no value for the unreliable inductive information required for the analysis of natural language.

It is clear that there has been that if there is no payoff to be obtained from a heuristic inductive approach to natural language understanding, there is no reason for pursuing it. This position is different from that of the search for a solution in formal fields, and urges the pursuit of the search for a solution in a field rather than a search for an answer to a question about whose content we are ignorant. An important difference between the deductive position is that their method does also depend on the content, or human competence in this area, and not just on the formalization of principles that could be expressed in a formal language.

After these qualifications, some very brief comparisons follow between the work described here, and three other AI approaches to language understanding: those of Charniak(7), Schank(8) and Hinton(9). In this comparison and criticism of systems is not meant to be a critique, and I give only brief general remarks, in order to contrast the present system along a number of dimensions; which are (a) the nature of the linguistic base constructed or proposed, in terms of application to everyday texts in English; (b) the degree of implementation and the definiteness of the task proposed as an application of the elusive notion of "understanding"; (c) the location of the system within the inference-deduction opposition, and (d) the implementation of a preference system that both prefers certain interpretations to others on a reasoned principle.

...do not consider the linguistic base essential, and is not particularly interested in the ambiguities of sense of words. ...the most adequate linguistic base of the three, and the present system in general presuppositions. ...restricted to unambiguous simple words, and a ...it is doubtful if the meaning of complex concepts, ...could be expressed in that way.

...is not intensively implemented but has a very ...the resolution of the sorts of anaphora ...I think the strategies we advocate ...whether or not the rules form an ...expressed at the word level, as he does, and ...the facility to chain OSIRs, or what he calls ...implementation through ...at present it ...in the inference field, ...as distinct from the drawing ...of some problem or performance of some task. ...The "inferences to be drawn from x" is ...outside the pages of detective fiction, ...completely implemented within its original ...and assessable task, its merit, as ...theory of grammar, is often said to ...the semantics (ilinsky and ...in an ad context, in that it ...distinction (semantics and ...work in semantic analysis by computer ...found unnecessary in ...the same authors' ...distinction as ...and undeniable.

...I do not feel on sure ground because ...the authors' work, since, naturally, their ...to answer this question of mine. ...the fact that some of the ...that Winograd's system is ...like the present one, are ...inference rules whose ...the possibility of fitting them ...the language context in hand. Any clarification ...the relation of their work to this distinction ...

...of preference and choice between ...nothing has been done by the authors, and ...there will always be one and only one ...in terms of their rules, or that the first ...is, I think, the only worker in the field

also has given attention to this question.

REFERENCES

1. An Artificial Intelligence approach to machine translation. Stanford AI Laboratory Memo #161, and in Schank, R. (Ed.) Computer Models of Thought and Language. San Francisco: W. H. Freeman, 1976 (in press).

2. Reference semantics. Stanford AI Laboratory Memo #162, and in Schank, R. (Ed.) Formal Semantics of Natural Language. San Francisco: W. H. Freeman, 1976.

3. Understanding without proofs, in Proc. Third. Internat. Joint Conf. on AI, 1973.

4. and Herskovits, A. An intelligent analyzer and generator for natural language. in Proc. Internat. Conf. on Computational Linguistics, Pisa, 1973.

5. The generation of French from a semantic representation. Stanford AI Laboratory Memo #212.

6. Denotability and natural language. Mind, 1971.

7. Jack and Janet in search of a theory of knowledge. Stanford AI Laboratory Memo #198.

8. and Rivest, C. Inference and the computer. Stanford AI Laboratory Memo #197.

9. Understanding Natural Language. Edinburgh, 1972.

10. and Papert, S. Artificial Intelligence Progress Report. Stanford AI Laboratory.

11. Semantic memory. in Minsky, M. (Ed.) Semantic Memory. MIT, 1968.