FINDING THE CONCEPTUAL CONTENT AND INTENTION IN AN UTTERANCE IN NATURAL LANGUAGE CONVERSATION*

Roger C. Schank

Computer Science Department
and
Committee on Linguistics
Stanford University
Stanford, California, U.S.A.

Abstract

The conceptual dependency analyzer described in the first IJCAI (8) has been modified so as to function more conceptually with less reliance on syntactic rules. In order to have an analyzer be conceptually driven, it is necessary for the system to know what it is looking for. That is, it must make predictions as to what can follow conceptually at any point in the analysis. This paper discusses the extension of conceptual prediction to include predictions based on context and the structure of the memory model that operates with the analyzer. Such predictions make use of relations between conceptual actions and the implications of those actions. This enables the conceptual analyzer to discover not only the conceptual content of an utterance but also the intention of that utterance in context. We are concerned with the extraction of the conceptual content both explicit and implicit in an utterance in order to analyze effectively in an interactive conversational situation.

I. Levels of Expectation

The primary emphasis that has been given to the study of the sentence by linguists and computational linguists has brought about some peculiar ways of studying natural language. Clearly people do not understand nor generate sentences in isolation. It has been in fashion among linguists who like to attack other linguists' ideas of grammaticality, to refute a statement of ungrammaticality by finding a situation in which the supposed ungrammatical sentence makes sense. Lakoff (6) has recently noted the need for using presupposition - sentence pairs before one can discuss grammaticality. It has long been our assertion that, while it seems reasonable that linguists who are studying grammaticality should take context into account, the study of grammaticality itself seems a bit misguided (see Schank (11)). What would seem to be more reasonable Is to realize that people talk in order to communicate something and it is the discovery of what this something is that is the proper domain of study for researchers interested in natural language. This point-of-view necessitates looking at language analytically rather than from the generative view of transformational grammar. It is this kind of viewpoint that eliminates notions that semantics consists of selectional restrictions which tell you what cannot be said. Clearly if something was said it must be dealt with regardless of its grammaticality.

But even if we recognize that the analytic study of language might yield some fruitful results, the possibility of falling into some of the traps left lying around by generative grammarians is extant. Of these traps, by far the most troublesome is the notion that the sentence is the core of the problem. Theories that are sentence-based simply miss the essence of the problem, namely that something is attempting to be communicated by the speaker and it can be ascertained by taking the entire situation in which it was uttered into account. Here we mean not only the linguistic context, but the physical, mental, emotional, and social contexts as well. Now this is not to say that we must disregard all work that has been done on sentence analysis up until now. On the contrary, many of the techniques used there have their analogues on other levels of analysis. But just as it was important to realize that it simply made no sense to analyze a sentence so as to detect all four or fifty possible syntactic arrangements for it (as the Kuno-Oettinger parser did for example (5)), likewise one does not wish to find more than one conceptual analysis of a sentence if the prevailing context clearly eliminates all but one of the choices.

One element which humans rely heavily on during the understanding process is that of expectation. At the sentence-level, we can predict at any point in an analysis what type of syntactic element is most likely to follow. Thus, if we have just seen a noun the likelihood that a verb will appear next is good assuming one has not already appeared. By the same token, an auxiliary or adverb is likely to appear but with a different probability. Some elements are much less likely to appear (an adjective for example) and some likely to appear depending on some of the semantic information contained within the noun. At any rate, guesses can be made based on what one might expect will occur. Guesses of this kind perform three major functions. First they point the way in searching a data base for an item. Second, they allow for disambiguation. On the sentential level, this means being able to choose between alternative senses of a word that are based on syntactic category. Third, they enable one to predict occurrences of information related to items that have already been discovered. This is important in establishing dependency information.

*This research is supported by Grant PHS MH 06645-09 from the National Institute of Mental Health, and (in part) by the Advanced Research Projects Agency of the Office of the Secretary of Defense (SD-183).

At the conceptual level expectations work in roughly the same way. That is, we can guess the type of conceptual information that is needed to make an unfinished sentence sensible. If we formalize these expectations we can enable a machine to know what it is looking for when it searches through a sentence attempting to find the conceptual structure underlying it. We can use this information for searching the data base, disambiguating, and creating dependencies.

There are predictions that can be made, however, that are based on criteria other than that directly derivable from the stratified linguistic system that comprises conceptual dependency theory (9). Consider the following situation and conversation:

> John meets his friend Fred on the street. Fred is holding a knife. John is angry because his wife Mary has yelled at him.
>
> Fred: Hi.
> John: What are you doing with that knife?
> Fred: Thought I'd teach the kids to play mumblypeg.
> John: I could use a knife right now. (agitated tone)
> Fred: What's the matter?
> John: Damn Mary, always on my back. She'll be sorry.
> Fred: I don't think a knife will help you.
> John: You're just on her side. I think I ought to...........

Now what can Fred expect that he will hear next? There are at least six distinct types of information with which we can answer this question. Sententially, Fred expects a verb. Conceptually, there is a conceptual representation of what John has just said which requires a complete actor-action-object construction (which we call a conceptualization) in order to be complete. Thus, conceptually a conceptualization is expected. But we can also make predictions based on context. According to the context, there are only a certain set of concepts which will fit into the needed conceptualization such that the conceptualization makes sense in context. We most certainly would be surprised if the next piece of information would be 'I think I ought to have fish for dinner'. It is knowing what we do and do not expect at any point in any analysis which allows us to be surprised, shocked or whatever other emotional attribute by a piece of information. You are not able to be surprised if you don't anticipate and it is therefore necessary for a system such as this to anticipate.

What we anticipate here are the following four types of statements based on their contextual likelihood: (1) hurt someone, (2) end relationship with somebody, (3) go to someplace, and (4) emote.

These are classes of actions. We don't know which sentential form 'hurt' 'go' or 'emote' will take but we can estimate the likelihood of the class on the basis of the conceptual category and the prevailing semantic categories that have been used in context. All of these above actions are predicted on the strength of their likely consequences. That is, a desired consequence is known (John feel better) and the above actions would each lead to John's feeling better, but each in a different way. This will be explained at length later on.

A fourth type of expectation or prediction is conversational. That is, people talk for a reason, usually to communicate something or to gain some desired effect in the hearer. Here, it is either to arouse sympathy or to inform about something he is about to do. But the use of ought implies he might not do this, so that his probable reason in making this statement has to do with the effect which it will create on the hearer. Thus we can predict what kind of effect is intended to be made by the speaker and then expect certain types of utterances.

A fifth kind of expectation information has to do with a world view of the situation based on his own individual memory model. Thus, if he knows John to be a convicted murderer his expectation of John's completion of this sentence ought to be different from his expectation if John were an avowed pacifist.

A sixth type of expectation is based on a memory-structure that is common to the cultural norm rather than the particular language or particular individual. The results of this kind of expectation have to do with the options that Fred can take as a result of the expected input from John. That is, the conversation is heading towards death (this idea will be explained below) and Fred's expectation of this can avert the situation by appropriate action, either physical informative conversational or emotional conversational. It is his expectation that decides the appropriate action and his expectation is based on the 'life → death' memory structure explained below.

Basically then, we must recognize that any complete processing system for a natural language utterance takes place within a context that is extremely complex because there are humans in the conversation. Each has a complex memory to begin with and is now in a new complex situation. Part of this problem is being able to anticipate. Therefore, getting a machine to be able to make predictions is an important part of the language understanding problem.

We have dealt elsewhere with sentential and conceptual predictions that aid a computer analysis, so these will be only briefly discussed in the next section. In the remaining sections we will be primarily concerned with discovering the intention of an utterance in a given situation, once the conceptual content has been ascertained.

II. Conceptual Dependency Analysis

Before we can begin to make predictions based on the last four expectation types, it is necessary to have a conceptual representation of each input utterance. That is, we must have analyzed what was said before we can know what the intention of a given utterance was.

Conceptual networks have been developed (see (9), (10)) that are intended to represent the conceptual content of a natural language utterance. We require of these conceptual networks that there be only one such network for any number of natural language utterances that have the same meaning. Thus, the first task in natural language analysis is to get the input utterances into some representation of the meaning of that utterance.

The conceptual representation schema presented here makes the following assumptions and notation:

(1) Underlying natural language sentences there are abstract conceptualizations.

(2) A conceptualization is either an actor-action complex (denoted by «») ; or an attribute statement (denoted b y)

(3) Actions have labeled dependents denoting, object (→), recipient («B), direction (D), or instrument (J).

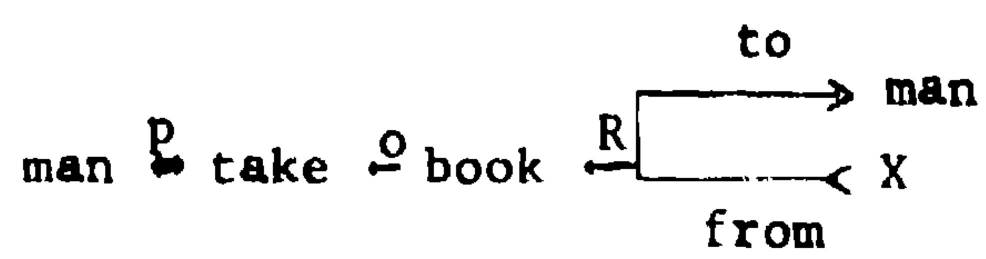
(4) Conceptualizations can relate to other conceptualizations by nesting as instruments or objects, or by causation (denoted by ⤊).

(5) Conceptualizations are modified as to tense by: p(past); f(future); c(conditional); t(transition); t_s/t_f(transition starting or finishing); k(continuant) ; A(timeless);Ø(present); /(negative); ?(interrogative), written over the •

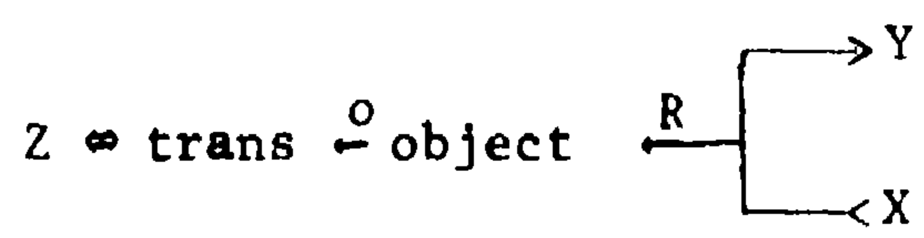
(6) A concept within a conceptualization can be modified by vertical arrows denoting attribution (t) and relation to another concept (⤊)

(7) Conceptualizations can be modified by times (t) or by locations (ft).

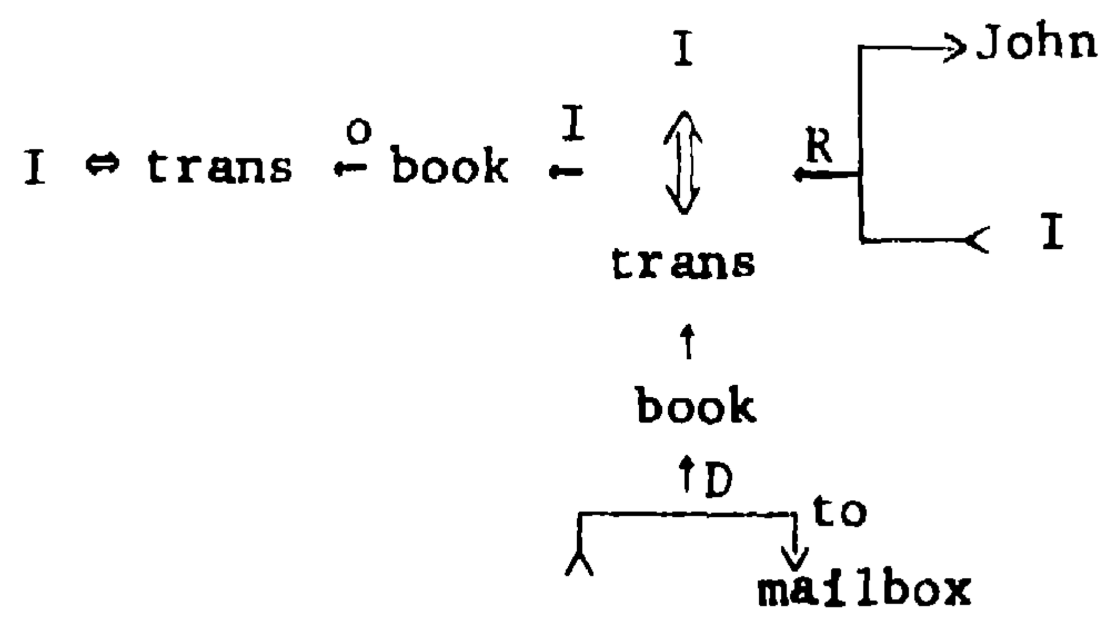
To see what the analysis of an actual sentence is like we can consider the simple sentence 'The man took the book.'. Using the above notation we might analyze this as: man » take . book But, in attempting to uncover the actual conceptualization underlying a sentence, we must recognize that conceptually a sentence is often more than its component parts. In fact, a dialogue is usually based on the information that is left out of the various conceptualizations. For example, in this sentence, we know that there was a time and location of this conceptualization and furthermore that the book was taken from 'someone' or 'someplace' and is, as far as we know, now in the possession of the actor. Thus, conceptually, there exists here a two-pronged recipient case, dependent on the conceptual verb (ACT) through the object. We use this recipient case to denote the transition in possession of the object. Thus we have the following network:



In this instance the recipient and the actor are the same. We can posit an underlying ACT here that is more abstract than 'take' but denotes the transition that is taking place, which we call 'trans.'. This abstraction allows us to recognize paraphrases at the conceptual level without losing syntactic information. We can define the English word 'take' as the instance when Z = Y in the following network:



'Give' Is the realized verb when Z = X. 'Receive represents the same diagram as 'give'. Similarly other 'trans' verbs, for example, 'send' and 'steal[1], have conceptual realizates where other aspects of the network are defined in some manner. Thus the following network, with a conceptualization as instrument, can be sententially realized with 'send'.



Each ACT (such as 'trans') belongs to an ACT category that requires certain conceptual cases (or ACT dependents). 'Trans', for example, must have an object, instrument, and recipient conceptually even if none were explicitly stated in the original sentence. Conceptual analysis by comput utilizes a verb dictionary which locates the appropriate sense of a verb by use of contiguous semantic information and then proceeds to rewrite the verb into a conceptual construction. This conceptual construction is a dependency network consisting of the underlying primitive action (e.g. 'trans', 'go') and the conceptual case dependents that are required by this action.

Thus, we can set up a network that has blank spaces in it with very definite requirements for what will fill them. The system can then know what it doesn't know and needs to find out. This predictive capability is very important in an analysis system that can function in a dialogue situation. Furthermore, semantic predictions enable the system to function without first doing a syntactic analysis. Syntax is called into play only as a searching mechanism for already known semantic information.

Conceptual Dependency uses four conceptual cases, objective, recipient, instrumental and directive. These cases, while not being too disparate from Fillmore's (3) ideas on syntactic cases, have their justification on conceptual grounds. We note that there is a difference between a conceptual instrument and the instrument as it functions syntactically. To better explain this it is necessary to digress for a moment to discuss a certain class of English verbs which we call 'pseudo-state'.

An example of a pseudo-state verb is 'grow'. When we say 'John grew plants', we usually mean that it was the 'plants' that 'grew' and not 'John'. But 'John' was an actor. However, the action that John did, which we call 'growing', was complex and probably consisted of weeding, hoeing, adding fertilizer, watering and so on. What we are really saying is that his action 'doing' (not 'he') caused the plants to 'grow'. The two conceptualizations are related causally (||) with the direction of the arrow denoting the governor and dependent conceptualizations. Thus, the above sentence is realized as:

```
       p
John   ⇔ do      where the 'do' is a dummy ACT.
       ⇑
plants ⇔ grow
       p
```

Now we can see that the sentential instrument of 'fertilizer' in the following sentence 'John grew the trees with fertilizer' is conceptually the instrument of one of the 'do's' associated with the verb 'grow', and not the ACT 'grow'. (In fact 'grow' belongs to the class of intransitive ACTs (IACT) which cannot take any conceptual case.) The most likely analysis of this sentence then, is:

```
              I
John  ⇔ do  ← fertilizer
      ⇑
trees ⇔ grow
```

Of importance here is the fact that the instrument is dependent on 'do' and not 'grow' (nor on 'cause'). However, the verb 'grow' can take an instrument of 'fertilizer'. This is an important distinction which is used by the parser.

The conceptual analysis technique that is referred to here, while reliant on syntactic information, does not first do a syntactic analysis. It is interesting to see how sentences that are similar syntactically are dealt with.

Consider the sentence 'John's love of Mary was beautiful'. This has the same surface structure as 'John's can of beans was edible' but they have different underlying conceptual structures. We consider that 'love' is an ACT no matter how it is realized and thus the NP in the first sentence is graphed:

```
                o
John ⇔ love  ← Mary
```

The graph of the active sentence is then

```
John
 ⇕   ⇔ beautiful
love
 ↑o
Mary
```

Here we have an abstract noun that is realized as an ACT. In the latter sentence we have the abstract adjective 'edible' and this too is an ACT conceptually. Thus the conceptual realization of this sentence is:

```
              o
one ⇔  eat  ← beans
      Poss ⇑    ↖  Loc
       John    can
```

'One' is a dummy actor. In order to find these analyses in the second sentence, 'edible' is discovered in the dictionary as

```
     c       o
one  ⇔  eat  ←  'food'
```

. This means that this conceptualization is conditionally true (denoted by c) and the potential object is of the semantic class 'food'. Syntactic rules indicate where to look for this 'food'. In this sentence, 'can' is examined first, and put aside in lieu of a better candidate for 'food'. Since 'beans' fits, it is used and semantic relations determine the dependencies involved.

In the first example, 'love' is the underlying action (ACT). A conceptualization is set up with 'love' as the ACT and the syntactics are used to determine the correct placement of the conceptual nominals (PP's). A predication about this conceptualization (is beautiful) is found and is placed in the network.

The primary task of the conceptual analyzer, then, is to discover the underlying action present in a sentence, if there is one, then to go through its experience to find out what kinds of syntactic combinations it is likely to find. That is, we want to know what to expect next at any point in the parse. These expectations are discovered on the basis of the syntactic and conceptual categories associated with a given verb.

The analyzer must then be able to choose between the set of senses assigned to a given verb by the dictionary. This it does by the use of what we might choose to call 'semantic'information, or 'selectional restrictions'.

The primary problem in conceptual analysis is finding the verb. To do this we make use of low level syntactic information, such as inflection and agreement rules and certain simple syntactic rules. This process has its parallel on the conceptual level in the heuristics that the parser employs to discover where to go in the conceptual network whenever it has become confused. These heuristics are then low level conceptual rules. That is, they operate on language-specific information that serve as commands to the conceptual apparatus.

Thus the parsing process consists of searching for a central element at each level (the verb sententially, and the actor-action combination conceptually, given that these exist); then using the expectation information provided as a guide to putting together the pieces of the puzzle. To see how predictions work in actual conceptual analysis we can consider parsing 'I saw the Grand Canyon flying to New York.' A parse of this sentence that was conceptually motivated would have to attempt to attach 'fly' to 'Grand Canyon' since we can observe that most English speakers find this sentence to be amusing, implying that they have tried just that.

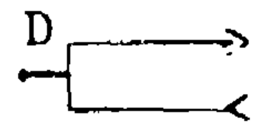
We begin by looking for ACTOR = ACT combinations. We place 'I' in the network and look for an ACT that can have a semantic connection to 'I'. 'See' satisfies this requirement and we create a two-way dependency between them. Upon encountering 'Grand Canyon', we look for the nearest concept that will form an acceptable dependency with it. The nearest concept is the ACT 'see' which is a semantically acceptable connection. Next, 'flying' is input and we have an ACT the nearest concept to which is 'Grand Canyon'. We then try to connect them. However, this choice is disallowed upon consulting the verb-ACT dictionary (which we do each time we find a verb).

Verb	Verb Category	ACT Realizate	Subject	Object
fly	vio	fly_1	birds, plane, insect	None
fly	vio	plane ⇕ go_1 ←I fly_1	animal	None
fly	vt	X ⇔ do ⇑ Y ⇔ fly_1	human	plane

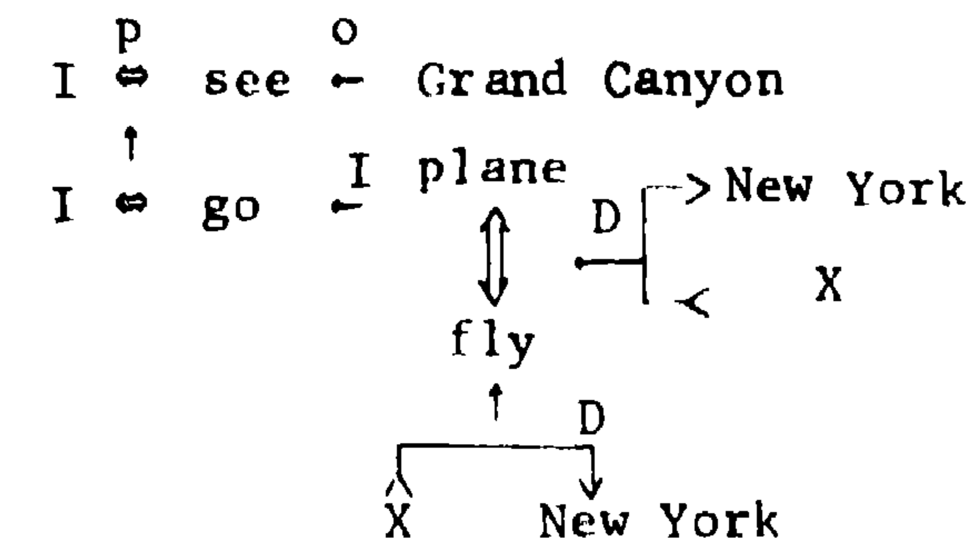
(Here vio denotes a verb that takes a sentential indirect object and vt is a transitive verb. The 'subject' and 'object' refer to sentential expectations, some of which are category names.)

Upon entering the dictionary we look for 'location' as a possible subject of a vio entry for 'fly'. We do not find it, however, and thus must go back to the network and try again.

We are now faced with the problem of what to do with our ACT. Its only possible governor according to the pertnissable dependencies is 'I'. In order to attach the ACT to I we employ an English heuristic which states, 'if you cross back over the = you have created a time T phrase, the actor of which is the previous actor'. We now re-enter the dictionary entry for 'fly' and discover that 'i' is an animal and thus fits under the second conceptual realizate for 'fly'. We place this realizate in the network. When we encounter the 'to N' construction we recognize a



(directive case) construct and our final analysis becomes



(In a dialogue program we might well discover that this was not the intended network at all, and the speaker believed that 'Grand Canyon' had 'flown'. We could then correct our error. However, we certainly would not want to assume that he had intended the latter interpretation.)

The important point here is that the system analyzes this sentence in the same way as a human does insofar as we can tell. That is, we could predict a chuckle on the part of the speaker based upon an attempt to attach two concepts that according to experience could not be attached. In other words, we try to make the 'Grand Canyon fly' but we cannot.

What is occurring here is that we are making use of a stratified linguistic system in order to make predictions. This system has a sentential and a conceptual level. Each level has its own rules for permissable constructions (syntax) and acceptable particular choices within a given construction (semantics). In the analysis operation the sentential syntax allows 'Grand Canyon flying' as do the conceptual dependency rules (or conceptual syntax). But the conceptual semantics disallows the proposed combination and forces the parser to try another dependency instead. In this case, the prediction of the syntax has fooled us but the conceptual predictions have corrected the matter. The conceptual predictions can direct the analysis once an ACT is found. In this case, when 'go' is discovered to be the underlying ACT it is known that a directive case will occur conceptually. The incoming prepositional phrase Is placed in its proper slot by the parser since it 'knows' what it is looking for at this point.

III. Associations

A program is running that will do what has been stated so far for a testing vocabulary of about 250 verbs and for a fairly large range of sentence-types.

But understanding natural language is more than recognizing the conceptual content of a given sentence. One problem that needs to be handled is that of extracting the presupposed information implicit in an utterance.

Consider the sentence 'I like books'. The analysis 'I = like £ books' is not an allowable construction conceptually because the ACT (action) 'like' is of two possible conceptual types, each with its own semantic restriction. As what we call EACT (emotion ACT), 'like' allows conceptual objects as shown above by 'O books' but requires that these objects be of the class 'animal'. The other sense of 'like' is conceptually an SACT (state ACT) which requires an entire conceptualization as object. A conceptualization must have an ACTOR and an ACTION at the least and we are thus faced with the problem of uncovering these in the analysis of the above sentence. We have then: I ⇔ like. We know that 'books' is part

↑
⇔

of this conceptualization and by the heuristics of the conceptual dependency system we know that 'I' is as well. The problem is what arrangement and what ACT is correct.

The ACT 'read' is listed in our system as requiring a 'human' subject and an object that is chosen from the set of objects that have been made by men for exactly the purpose of 'reading' them. That is, while we could list all possible such objects (books, newspapers, etc.) or categorize them in some artificial hierarchical structure, conceptually the object of 'reading' Is 'that which is read'. Specifically this class could Include anything with printing on it or whatever. The point here is that we can call the potential object a member of the class 'read-PP (where PP is the abbreviation for conceptual nominals). Then, in any listing of the elements of the world, their semantic category is the place that they lit in our ACT-based model. 'Book' then, is: 'book: N; read-PP_o;' where 'read-PP_O' denotes that it is the conceptual object of the ACT 'read'. Then our diagram becomes: I «=>> like . The only

read ◆- book

thing missing is the actor, which is 'I' due to a heuristic which governs these situations.

There are conceptual representations for many objects which can be made in the same way as was done for 'book'. For example, consider 'knife'. 'Knife' is an instance of 'cut-PPj', this means that it serves as the object in the instrumental conceptualization in conceptualizations involving 'cut'; and 'banana' is an Instance of 'eat-PP_{0o}'.

There is a second, more complex, type of recognition of implicit information that is a part of understanding. For example, the^ sentence 'I fear bears' is directly related to the harm that a bear might do. That is, a correct analysis might be: I • fear . Here again, this

↑
bears ⇔ do ← I
⇑
I ⇔ hurt

action ('fear' is related to a conceptualization rather than one particular concept.) (In other words, you 'fear' consequences not properties.) What would seem reasonable here is that 'fear' and 'hurt' are directly relatable. Now it is possible to think of this relationship (fear-hurt) as some relatable grouping of ACTIONS. But this is not the case. Conceptually, 'hurt' is a mental ^tate attribute and not an ACT. Mental state attributes (denoted 2PA) are nearly always expressed in English as transitive verbs. The object sententially is often the subject of the attributive statement, when ZPA's are used in a sentence (e.g. 'x hurt y' means 'y is hurt'.) This means that certain ACTs like fear should have consequent 2PA's that they are related to. We can carry this one step further. The reason that 'hurt' is 'feared' is because of another consequence, namely 'death'. Now this may seem a little melodramatic, but it does in fact seem to be the case. In other words, a lot of 'hurt' leads to 'death'. Now 'death* is conceptually the IACT-'die'. So we have here a relationship from SACT (fear) to 2PA (hurt) to IACT (die). In fact, there is one element missing here, namely the 'do' associated with 'bear'. This 'do' may be 'claw', 'eat' or some other physical action that takes instrumental case (PACT). So what we have is the set of relations SACT - PACT - ZPA - IACT (see Weber (13) for detailed explanation of these terms). This relation of conceptualization types always holds for any ACT.

The main point here is analogous to that made in the previous section: Whenever certain concepts are encountered other concepts are actually present in the underlying conceptualization and must be ascertained before a reasonable claim of understanding can be made.

The relationship belween SACT - [variable ACT] - ZPA - IACT essentially states that people do and say things [or reasons or desired eliects. Thus, actions have their *consequences* in *new* mental states for a doer or receiver of this action and these lead to new actual states. In order to talk of actual states it will be necessary to explain the notion of variant levels within a conceptual base. Celce and Schwarcz (1) and Tesler (12) discuss the notion that certain concepts have both mental and physical realities. For example, you read a 'mental book' but lift a 'physical book'. This dichotomy can be broadened to include levels of a social, emotional and spiritual nature as well according to Tesler. Actions, for example, can be seen to have different but related meanings on each level. Consider 'go'. Physically .'go' means to go from one place to another. This has its analogue socially in two ways. On the one hand, you can go to a 'social place' e.g. a convention. On the other you can go to a place within society i.e. social climbing ('He went upwards socially after his election'.). One can 'go[1] emotionally ('After his death, I went to a state of depression[1].). Mentally we have, 'My thoughts went to the days in Tangiers.' And spiritually we have the common 'You will go to heaven.'

The reason for this apparent digression is that certain ACT's relate to certain ZPA's and IACT's according to the variant level of the ACT. Thus, the statement 'I was afraid that the bear would claw me' is a statement of physical dimension where the notion that 'physical clawing leads to physical hurt leads to physical death' holds. Now the fear of death that is implied here does not indicate that the object is necessarily aware of his fear of death.

This same kind of ACT - ZPA - IACT statement can be said to exist on each variant level. Consider the statement 'We are going to take away all your political power in this state'. The 'take' that is being used here is hardly the physical 'take' (take$_{PHYS}$) used in 'He took my toy'. Rather it is a social 'take' (take$_{SOC}$). Now this social 'take' leads to 'impoverishment$_{SOC}$' just as a physical 'take' leads to impoverishment$_{PHYS}$'. That is, when something is taken from one, the consequence is that the 'taker' is richer in some way and the 'taken' is poorer in some way. This is the ZPA of attribute in this case. The last consequence of 'death' holds as well in this case, but here it is 'death$_{SOC}$'. That is to say, the end result of such an action as stated above is that after his political power is taken away, he can be said to have 'died' politically. The end result is 'die$_{SOC}$'.

What we are saying then is that it is possible to get a great deal more information out of one ACT then is readily obvious. On the most apparent level, the notion of expected objects and subjects and other conceptual case dependents can be predicted. But more significantly, we can also make simple implications as a result of the position of the ACT in question with respect to its relation to other conceptual consequences. That is, we can know the way in which an ACT relates to 'living' or 'dying' on a certain level and the range of human mental reactions on these levels to such an ACT.

Consider the following PACT's: a) eat, drink, love, fight, hit; b) hit, cut, attack, The ACT's in list (a) are positive with respect to the subject. Those in list (b) are negative with respect to the object.

When we say that an ACT is positive with respect to the object, we mean that the subject performs this ACT with the intention of having a good result occur on the particular level with which we are dealing.

By the same token, if an ACT is negative with respect to the object, we can assume that this ACT's consequent effect on the object is bad for the object and tends to hurt him on the same level as the type of ACT with which we are dealing. That is, 'rob' is social and leads to 'hurt$_{SOC}$').

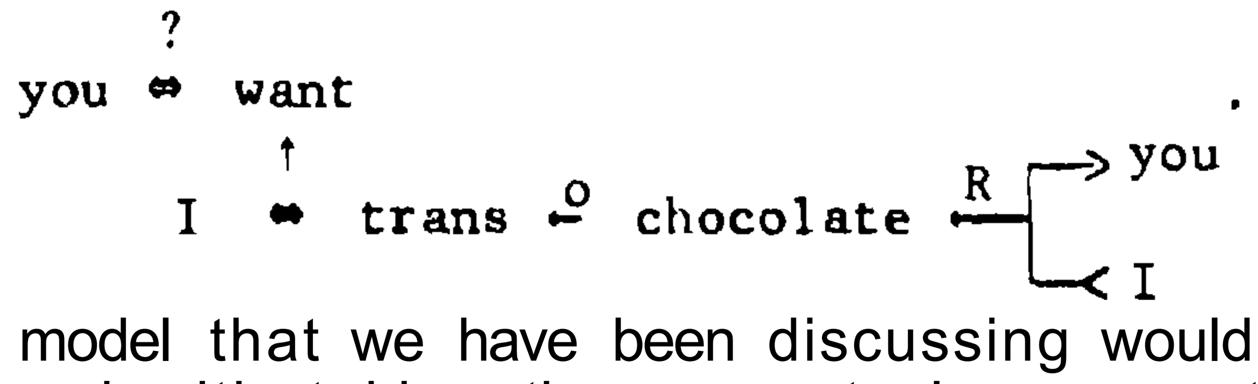
Statements of this kind are assumed to have ordinary circumstances prevailing. While it is possible to envision circumstances under which the supposed implications of an ACT do not hold, for the purpose of making predictions we assume the most likely situation. Often, the likely implications were implicitly part of a given utterance. The problem here is to undo the basically telegraphic nature of natural language utterances, and the possibility that we can make errors in so doing should not surprise us.

IV. Intentions

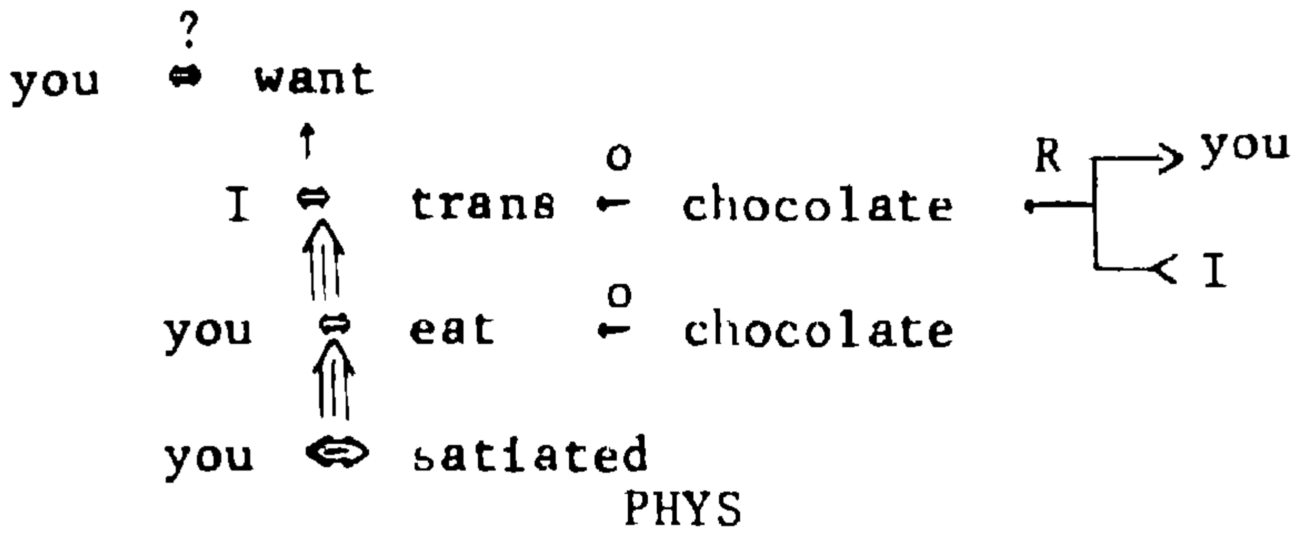
The relationship of conceptualization types and the consistency of variant levels with these types can provide the basis of a schema for discovering the intention of an utterance.

Consider a sequence such as this: Q: Do you want a piece of chocolate? A: I just had an ice cream cone. In a model of natural language understanding, it is unreasonable to claim that the system has understood the utterance unless it is * capable of producing for (A) not only a conceptual diagram of the information just stated, but also something like the answer 'no I don't want a piece of chocolate'. That is what a human could understand in the above situation and it is incumbant upon any so-called understanding model to understand the same. (This was noted by Gardiner, before computers came on the scene, in his definition of meaning (4): 'The meaning of any sentence is what the speaker intends to be understood from it by the listener.')

We can actually do this as follows: The conceptual dependency analysis of (Q) is:

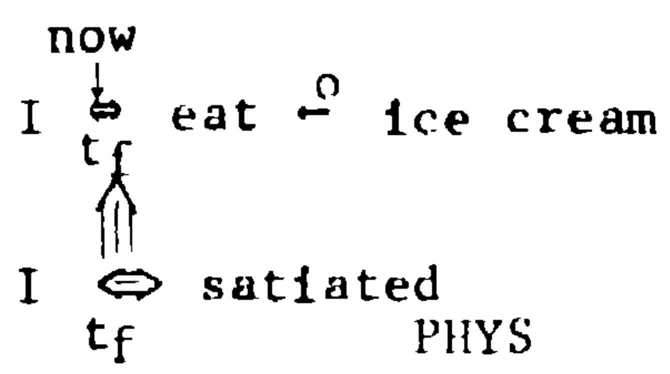


. The model that we have been discussing would be charged with taking the conceptual representation of the input and drawing the necessary implications that can be said to be understood implicitly. In this case, chocolate is discovered in the dictionary to be an 'eat:PP '. The association between 'want' (SACT) and 'eat' (PACT) fits into the SACT - ACT - ZPA - TACT paradigm on the physical level because of the definition of 'eat' and yields the implications that the ZPA 'satiate' is caused by the connection with respect to the subject of 'want'. This gives us:

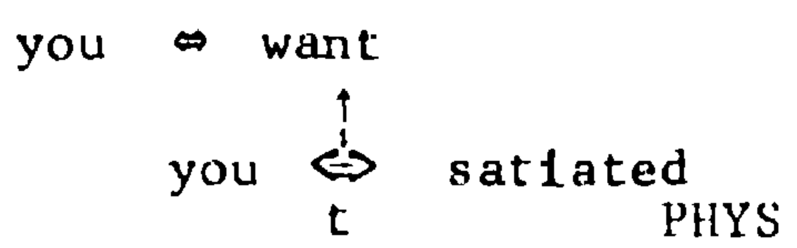


Now it is also true that people eat for reasons other than satiation, particularly for pleasure. So the causal connection ⇑ you ⇔ pleased t_f is also a consequence of the 'eat' conceptualization But this is not necessary here.

Now we are ready to analyze the answer (A). The conceptual diagram associated with the input is: 1 ⇔ t_f eat ← icecream ⇑ cone. This diagram is obtained by treating 'have' as a dummy ACT and finding the ACT associated with 'ice cream',again 'eat' to put in its place. Here again, 'eat' implies the causal for satiation and we have:



Now we can compare the question and the answer. The question can be matched with the answer by looking at:

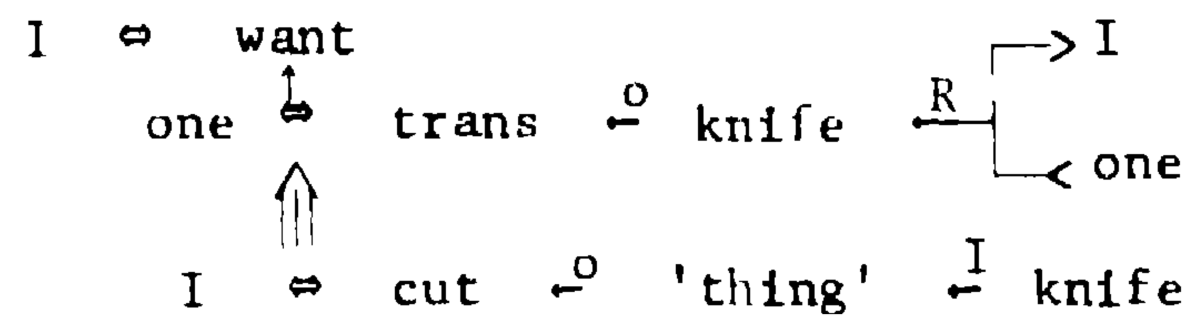


from the question, and: I ⇔ satiated (t_f, PHYS) from the answer. Since 'you' and 'I' represent the same token in memory, the answer to a question about desired transition (t) lias been answered with a statement of completed transition (tf). In other words, we can assume that we have, 'do you want to be satiated?', - 'I have just been satiated'. Thus we have the simple implied negative.

The point here is that the implications that are to be found in this memory model are part and parcel of the understanding process and in fact make little sense without them. We can expect that a natural language analysis system must be continually making these associative implications in order to be able to use them when they are needed. What we are doing is attempting to uncover the reasons that a given sentence was said. In order to correctly respond to an utterance it is necessary to have understood why that utterance was said.

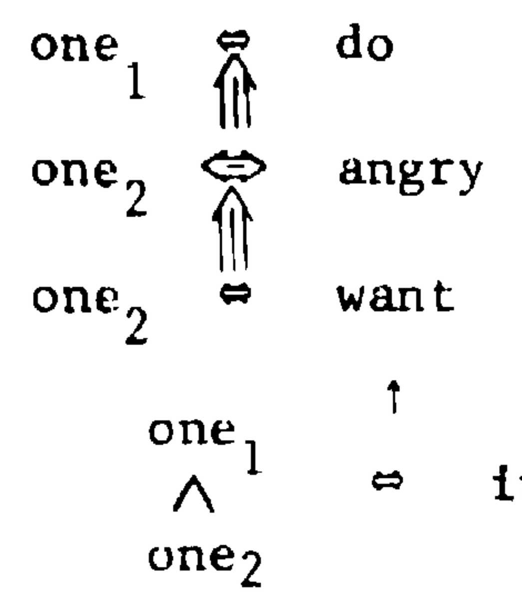
We can now reconsider the levels of expectation with which we were concerned earlier. In the conversation between 'John' and 'Fred' we noted that the context predicts what kinds of conceptualizations are likely to be asserted. We said that what was likely was that John would say 'I think I ought to {kill {Mary / you (Fred)}}' or 'I think I ought to {end my relationship with {Mary / you (Fred)}}. It should be clear that the particular words that would be used here are not at issue, but only their conceptual content. Now, the question is, how do we get a machine to make these predictions?

The problem is one of derivation. That is, where would this information come from? Consider the statement made by John previous to the one under discussion ('I could use a knife right now'). This is represented as:

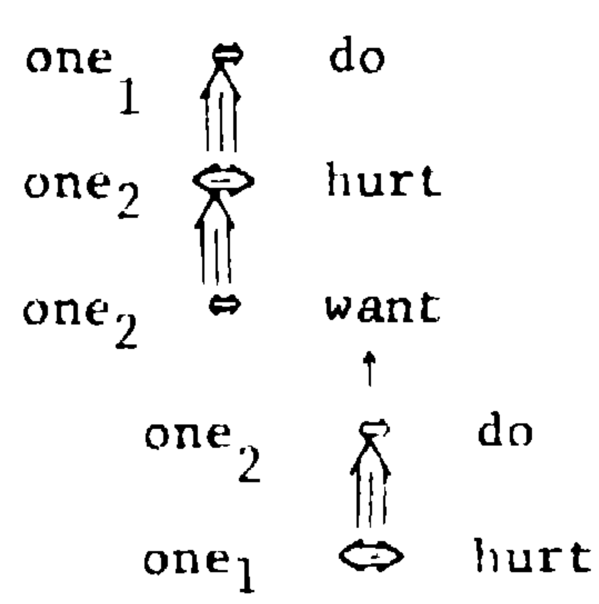


Here the first causal implication comes from the SACT - ACT - fiPA - IACT paradigm, or, in this case - 'want - ACT - ZPA - live'. Now, we can say that we have a conceptualization in the short-term memory that will affect the context. That Is, John » want ↑ John «* cut *- (o) •- (T) knife

In order to make accurate use of this information, it is necessary to have at the system's disposal a belief that could be characterized as part of the world view expectation. This belief is of the general order:

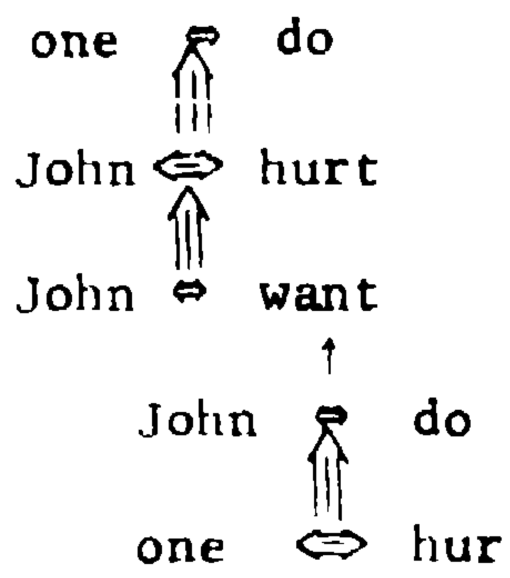


That is, this rule explains that if one is angry at someone that means that one doesn't want to interact with that person. We also need a rule that says:

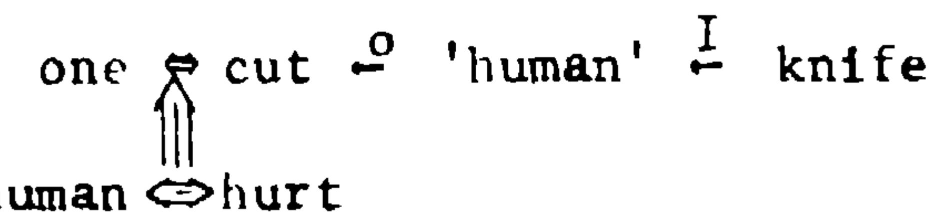


In other words, if one is hurt one wants to retaliate. Now oi course, this rule is not always true for every individual. We would like to note the conditionality of this rule by placing a 'c' over 'one$_2$ = want' and then using the rule if it is the case that in our memory of the individual to whom we are talking we have some information about previous actions of a 'cause to die' nature. That is, if we know that John already killed for some reason, we might guess that John will retaliate again. On the other hand we might have the information that John frequently talks about harming people, but doesn't do it.

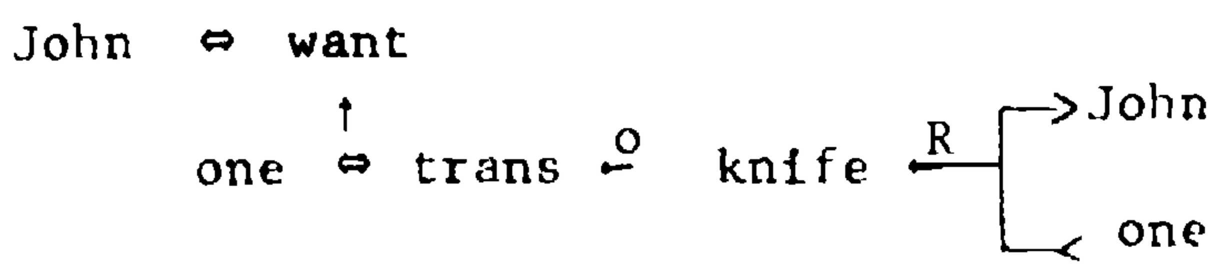
The point is that if we can decide that John believes that:



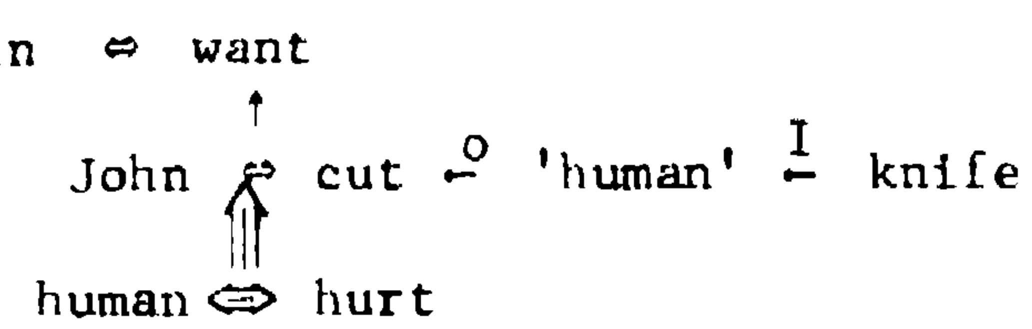
and we know that:



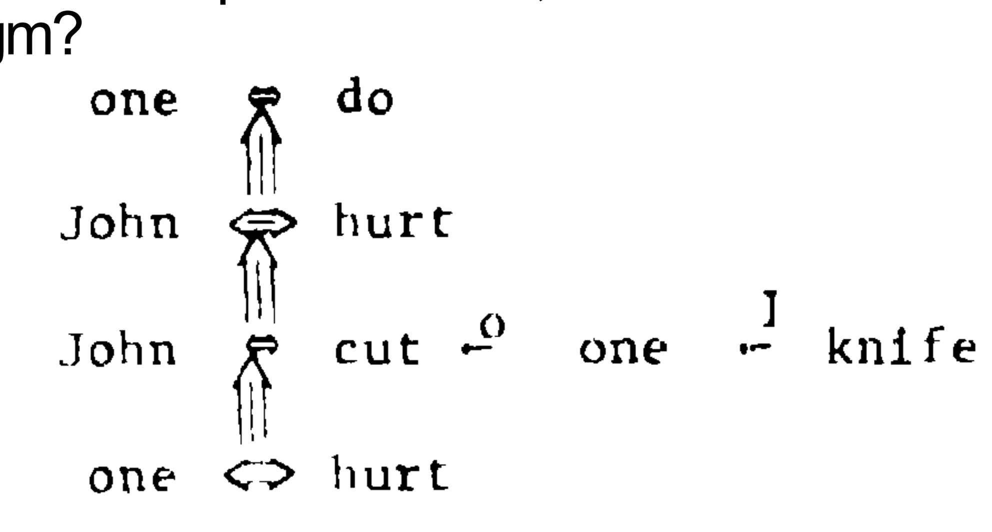
and, we know that:



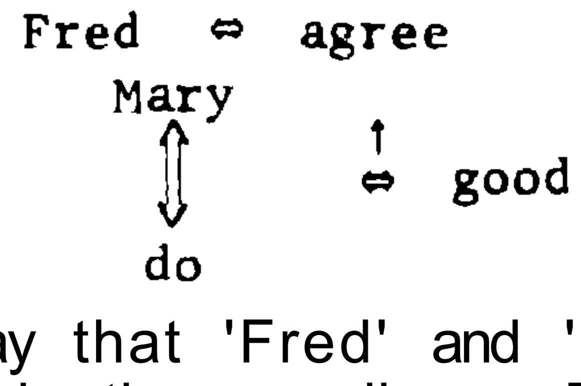
then we can conclude that:



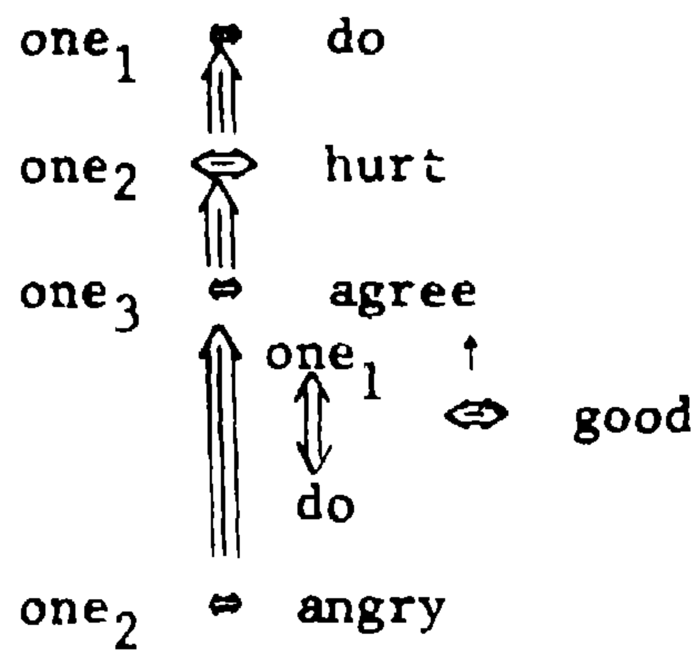
Now the question is, who fits the following paradigm?



Since, John has said that Mary angered him, she fits the paradigm by definition of 'angered'. Since, Fred has just convinced John that:



we can say that 'Fred' and 'Mary' are in the same situation in the paradigm. This is done by yet another belief that says:



That is, if one sides with one's enemy then one is angry at the enemy's compatriot also. Thus we can say that John is likely to say that he will kill either Fred or Mary. Also, we can say that the context of the knife aside, he is likely to say that he doesn't want to interact with either Fred or Mary.

The important point here is that it is possible to make contextual predictions as to the content of expected conceptualizations, but that this process of prediction is based on primitive beliefs that include generalized rules for operating in the world, and idiosyncratic beliefs about the behavior of an individual in the world based upon one's view of people and the particular person under discussion.

Although it may seem so, the number of the primitive beliefs necessary to handle tasks such as this is not large. Colby (2) and Morris (7) have estimated the core beliefs of a human as under 50.

V. Conclusion

In order to enable computers to use natural language it may be a good idea to understand how it is that people do these things that we would like our machines to do. In order to achieve this goal, I claim that it is not possible to separate language from the rest of the intelligence mechanisms of the human mind. Language simply does not work in isolation. It is a nice idea that one should in principle be able to fully describe and characterize language by itself as most linguists are trying to do, but in fact it is as absurd an idea as trying to understand the workings of the human mind by cutting off a man's head and taking a look inside. No doubt It is possible to find out some things that way, but the separation is artificial, it destroys the very process that we would be trying to investigate-. So it is with language. The ability oi linguists to Ignore this while trying to separate language into neat formal rules has caused an unbelievable number of unrealistic studies to take place under the banner of linguistics. People neither randomly generate sentences nor do they attempt to assign syntactic markers to input discourse. It is certainly true that humans may perform some of the subtasks that are needed in order to have a formal model do these things, but the overriding question is one of purpose. Since we are trying to enable computers to communicate with people we must deal with problems of communication

Certain pairs of people find it harder to communicate than other pairs. This is indicative of a lack of certain common memory structures and inference relations. We cannot understand somebody whose initial assumptions and cultural background are radically different from our own, even if we share a common language. That is, understanding language is a misnomer or at least is only a small part of the problem. Understanding what one has heard is a complex process that necessitates connecting words with certain conceptual constructions that exist in one's memory. The entire linguistic process uses the output of such understanding and interpreting mechanisms in order to produce reasonable replies (verbal or not). What constitutes a reasonable reply is an intrinsic part of the linguistic process, but yet is still a conceptual process and is therefore 1 suppose out of the domain of modern linguistics. Yet it is unreasonable for it to remain in that scientific no-man's-land. A computer model must respond as well as understand. Of course, its response must be connected to a powerful responding mechanism that is in fact the point of the entire computer program, that is, why the program was written in the first place. These then are the problems of computer understanding of natural language.

Now, why should it be necessary to make all these different predictions that have been outlined here? The answer is that in a complete automatic linguistic system the responses that are generated will be dependent on the corroboration of the predicted input as compared to the actual input and the memory structure. That is, we respond differently to different people saying the same things, and differently to the same people saying the same things in different contexts. These contexts include, physical, conversational and time contexts. fn other words, no person is really the same at any given point in time as he was at some other time with respect to the viewer's own memory model of that person. So, 1n some sense, the context is always different and the responses should always be potentially different according to the time of the conversation. It is precisely the predictive ability that permits this difference in response. And, the difference in response is caused by the difference in analysis. That is, in order to effectively analyze a given linguistic input, it is necessary to make predictions as to what that input might look like, compare the actual input to the expected input and coordinate both with the memory model. Understanding is, therefore, a complicated process which cannot be reasonably isolated into linguistic and memory components but must be a combined effort of both.

The remaining question is, will the suggestions made here for understanding natural language actually work? The answer is that we can't really know that until we are through. The structures that must be built are large and the number of beliefs and implication rules arc also large. But the basic elements of the process should be not much larger than has been described here.

It should in principle, be possible to use the suggestions made here for a beginning to attempt to truly understand input utterances. We are beginning to extend our conceptual analyzer to incorporate these ideas.

REFERENCES

1. Celce, M. and Schwarcz, R. "A Note on Two Basic Semantic Distinctions", System Development Corporation, Santa Monica, California, April 1969.

2. Colby, K., Tesler, L., and Enea, II. "Experiments with a Search Algorithm for the Data Base of a Human Belief Structure", in Proceedings of the International Joint Conlerence on AI. Walker and Norton (eds.) Washington, D.C., May 1969.

3. Fillmore, C. "The Case for Case", in Bach and Harms (eds.), Universals in Linguistic Theory, Holt, Rinehart and Winston, New York, 1968.

A. Gardiner, A. "The Definition ol the Word and the Sentence", British Journal of Psychology, V12, 1922.

5. Kuno, S., and Oettinger, A. "Multiple Path Syntactic Analyzer", in Information Processing 1962, North Holland, Amsterdam, 1963.

6. Lakoff, G. "Presuppositions and Relative Grammaticality", in Automatic Translation and Mathematical Linguistics, NSF-24, Harvard Computation Center, Cambridge, Massachusetts, 1970.

7. Morris, C. Signification and Signjficance, MIT Press, Cambridge, Massachusetts, 1964.

8. Schank, R., and Tesler, 1,. "A Conceptual Parser for Natural Language", in Proceedings of the International Joint Conference on AI. Walker and Norton (eds.), Washington, D.C., May 1969.

9. Schank, R. "A Conceptual Dependency Representation for a Computer-Oriented Semantics Stanford A.I. Memo 83, Computer Science Department, Stanford University, Stanford, California, March 1969.

10. Schank, R., Tesler, L., and Wrber, S. "Spinoza II: Conceptual Case-Based Natural Language Analysis", Stanford A.I Memo 109, Computer Science Department, Stanford University, Stanford, California, lanuary 1970.

11. Schank, R. "Semantics in Conceptual Analysis", Stanford A.I. Memo 122, Computer Science

Drpartment, Stanford University, Stanford California, May 1970.

12. Tesler, L. "New Approaches to Conceptual Dependency Analysis", in (10).

13. Weber, S. "Conceptual ACT Categories", in (10).