AD-A070 011 YALE UNIV NEW HAVEN CONN DEPT OF COMPUTER SCIENCE F/G 5/7
THE USE OF STEREOTYPE INFORMATION IN THE COMPREHENSION OF NOUN --ETC(U)
MAY 79 R C SCHANK, M LEBOWITZ N00014-75-C-1111
UNCLASSIFIED RR-144 NL

| OF |
AD
A070011

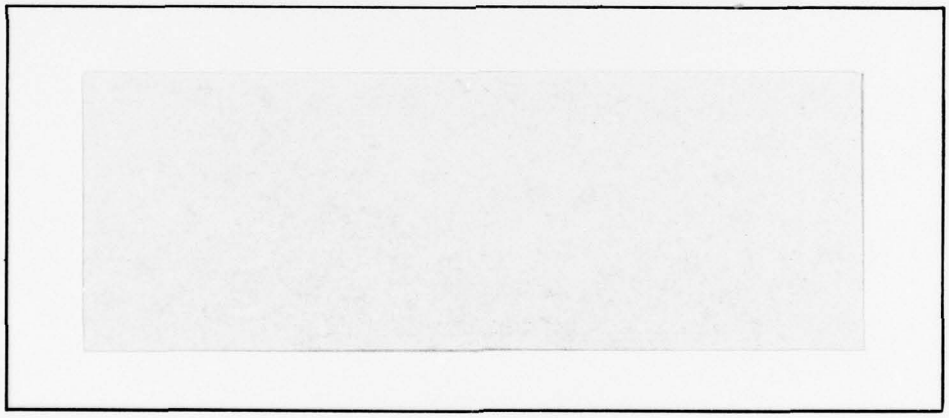
END
DATE
FILMED
7-79
DDC

LEVEL

12

AD A070011

DDC
RECEIVED
JUN 19 1979
C



DDC FILE COPY

This document has been approved for public release and sale; its distribution is unlimited.

YALE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

The Use of Stereotype Information in the Comprehension of Noun Phrases

by

Roger C. Schank and Michael Lebowitz

Research Report #144
1979

The research described here was done at the Yale Artificial Intelligence Project and is funded in part by the Advanced Research Projects Agency of the Department of Defense and monitored under the Office of Naval Research under contract N00014-75-C-1111.

This document has been approved for public release and sale; its distribution is unlimited.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: #144
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): The Use of Stereotype Information in the Comprehension of Noun Phrases.

5. TYPE OF REPORT & PERIOD COVERED: Technical Report

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): Roger C. Schank, Michael Lebowitz

8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-1111

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Yale University - Department of Computer Science
10 Hillhouse Avenue
New Haven, Connecticut 06520

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 30 p.

11. CONTROLLING OFFICE NAME AND ADDRESS:
Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

12. REPORT DATE: May 79

13. NUMBER OF PAGES: 24

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
Office of Naval Research
Information Systems Program
Arlington, Virginia 22217

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this report is unlimited.

RR-144

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Natural Language Understanding
Semantic Memory
Characterization
Stereotypes
Predictive Processing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The importance of stereotypical information in natural language understanding is discussed. Two Classes of words - those which create them - are presented. Combinations of such words, such as "ambitious hippie" are examined for the clues they give about what should be part of the knowledge structure used to hold information about a person. The use of characterization in generating predictions to facilitate understanding is discussed. A program designed to investigate the problems in processing stereotypes is

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

407051

described, and examples of its operation are shown.

Accession For
NTIS GRA&I
DDC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or special
A

-- OFFICIAL DISTIRUBTION LIST --

Defense Documentation Center Cameron Station Alexandria, Virginia 22314	12 copies
Office of Naval Research Information Systems Program Code 437 Arlington, Virginia 22217	2 copies
Advanced Research Projects Agency Cybernetics Technology Office 1400 Wilson Boulevard Arlington, Virginia 22209	3 copies
Office of Naval Research Branch Office - Boston 495 Summer Street Boston, Massachusetts 02210	1 copy
Office of Naval Research Branch Office - Chicago 536 South Clark Street Chicago, Illinois 60615	1 copy
Office of Naval Research Branch Office - Pasadena 1030 East Green Street Pasadena, California 91106	1 copy
Mr. Steven Wong Administrative Contracting Officer New York Area Office 715 Broadway - 5th Floor New York, New York 10003	1 copy
Naval Research Laboratory Technical Information Division Code 2627 Washington, D.C. 20375	6 copies
Dr. A.L. Slafkosky Scientific Advisor Commandant of the Marine Corps Code RD-1 Washington, D.C. 20380	1 copy
Office of Naval Research Code 455 Arlington, Virginia 22217	1 copy

Office of Naval Research — 1 copy
Code 458
Arlington, Virginia 22217

Naval Electronics Laboratory Center — 1 copy
Advanced Software Technology Division
Code 5200
San Diego, California 92152

Mr. E.H. Gleissner — 1 copy
Naval Ship Research and Development
Computation and Mathematics Department
Bethesda, Maryland 20084

Captain Grace M. Hopper — 1 copy
NAICOM/MIS Planning Board
Office of the Chief of Naval Operations
Washington, D.C. 20350

Mr. Kin B. Thompson — 1 copy
Technical Director
Information Systems Division
OP-91T
Office of the Chief of Naval Operations
Washington, D.C. 20350

Advanced Research Project Agency — 1 copy
Information Processing Techniques
1400 Wilson Boulevard
Arlington, Virginia 22209

Professor Omar Wing — 1 copy
Columbia University in the City of New York
Department of Electrical Engineering and
Computer Science
New York, New York 10027

Office of Naval Research — 1 copy
Assistant Chief for Technology
Code 200
Arlington, Virginia 22217

Captain Richard L. Martin, USN — 1 copy
Commanding Officer
USS Francis Marion (LPA-249)
FPO New York 09501

Major J.P. Pennell — 1 copy
Headquarters, Marine Corp.
(Attn: Code CCA-40)
Washington, D.C. 20380

The Use of Stereotype Information in the Comprehension of Noun Phrases

by

Roger C. Schank and Michael Lebowitz

Introduction

Does a hippie own a hairdrier? Does a professor own a Mercedes? Does a business executive wear bermuda shorts on Wednesdays? Would a jetsetter be interested in meeting an AI researcher?

The answer to questions such as these is dependent on the use of stereotypical information. Others (e.g. Bobrow and Winograd, 1977, Rosch, 1975) have commented on the need for such information in the understanding task. Indeed scripts (Schank and Abelson, 1977) and frames (Minsky, 1975) are really versions of more elaborate stereotypes. In our work on story understanding, we have come across the need to handle a class of information that relies on stereotypical characterizations of people for which we had no existing no theoretical or programmed apparatus. The answer to the above questions would certainly be nice to have in any understanding system. But, far from being an added frill, it is imperative to understand that words that characterize people and their behavior do a great deal more than just

---

This work was supported in part by the Advanced Research Projects Agency of the Department of Defense and monitored by the Office of Naval Research under contract N00014-75-C-1111.

create unstructured tokens. These words serve to foster a great many predictions. Since most of our understanding systems rely heavily on their ability to predict future inputs (e.g. the programs of Riesbeck, 1975, Carbonell, 1979, Wilensky, 1978 and DeJong, 1977 are all top-down predictors of information) personal characterizations can strongly affect the understanding process.

This problem was first worked on by us (Schank and Abelson, 1977) within the context of role themes. In that work we tried to explain how a lawyer picking up a street's garbage differed from a garbage man doing the same thing. A new variation of the problem has caused us to further consider the issue of stereotypes, namely the problem of trying to combine certain adjectives with stereotypical descriptions.

What is an "ambitious hippie" or a "permissive parent" or a "kind professor" or a "morbid rabbi" or an "aggressive cheerleader"? These combinations create images in our minds of the person being discussed. The creation of such images in a computer, in some form or another, is an important goal. To a great extent, understanding means being able to form such images. The precise nature of images, will not be addressed here. The key issue is to establish exactly what each of these words might mean. Words like "ambitious" and "aggressive" are particularly difficult to represent in any of the methods that we have used over the years. The reason

these words are difficult is that they function as general predictors of behavior. Somebody who is ambitious can be said to want more things than normal. But what do words such as "things" and "normal" mean? They really don't mean anything until a stereotypical noun of the type mentioned above is found. But then to understand the entire noun group we have to understand the goals of a normative hippie or rabbi. This implies that the concept underlying such words must contain goal and plan type information associated with the stereotype. To take a concrete example, we are saying that when psychological experiments such as those done by Anderson and Bower (1973) use sentences like "In the park the hippie touched the debutante," it is necessary to realize that there is a lot more going on than just recognizing the syntactic and semantic relations that hold between the words. Part of understanding such a sentence is realizing that the debutante was likely to have been horrified or at least upset. Knowing this is a part of understanding in the fullest sense what these words mean.

We are discussing then, two classes of concepts. One which we shall call <u>characterization stereotypes</u> build the default normality frame for a concept of a particular type of individual. Many words have as their definition a characterization stereotype. These words are frequently used to introduce a new person or character. The second concept is the <u>stereotype modifier</u> which alters a stereotype. There is also a class of words with stereotype

modifiers as definitions.

Characterization Stereotypes

Characterization stereotype words (CSW's) are the words which create a normality frame. They appeal to the stereotypes we have for numerous classes of people. Stereotypes give an understander access to vast amounts of information. A typical stereotype will include information about all aspects of the person's life. These aspects include his appearance - a basketball player is tall, a hippie is scruffy; his job - a doctor sees patients and doesn't make house calls; his background - a lawyer probably came from a middle class neighborhood and is upwardly mobile; his goals - a business executive wants to make lots of money, an AI researcher wants to publish landmark papers; and the interpersonal relations he is part of - how an assistant professor treats his students or a company vice-president acts toward his wife. It is information such as this that allows an understander to answer questions such as those we asked at the beginning of this paper. Information in a stereotype can also serve as the basis for explanation. If a story understander is trying to explain why a character is going to the library at 8:30 in the morning, knowing that he is a librarian would be an adequate explanation. Knowing that he is a student who has an exam later that day is adequate as well of course. It is important for a program to not only understand that

something is reasonable, but to know why it is reasonable.

The information contributed to a characterization by a CSW varies in certainty. Some characteristics must hold for the person to be accurately described by that CSW. A lawyer must have a law degree or he is not a lawyer. Other pieces of information we are quite sure of, and we would be surprised if we were wrong, but there is no logical guarantee of correctness. For example, the stereotypical rabbi has a beard, and the stereotypical doctor has a fancy car. Finally, there is information which we belive, but will disregard the minute better information comes around. The stereotypical professor at Yale may teach English, but we will believe he teaches History without the slightest concern. One of the most striking points about stereotypes is that they produce values for virtually every conceivable part of a characterization. Some of the information (such as car type) may be highly speculative, but it is there.

The common CSW's that we find in English tend to describe the following characteristics of a person: his profession or his position within that profession (i.e., politician vs state senator); a familial relationship (mother-in-law); his life style (hippie, jetsetter); his habitual activities (smoker, kvetch); his intellectual affiliations (existentialist, socialist); his social affiliations (Chamber of Commerce member, Chicano); his attitudes (megalomaniac, fundamentalist). There are more

than just these types of CSW's, but these will suffice for our purposes here.

Each of these words builds up a stereotype in the hearer's mind. This stereotype can be modified by the class of words we call stereotype modifiers (SM's). What is most interesting about SM's is that they point out to us the kinds of information that must be present in a frame representing our knowledge of a person. Certainly, we cannot modify what is not there. In setting up a standard frame for a person then, we must pay careful attention to the various kinds of SM's, so as to know what belongs in the default person frame.

Stereotype Modifiers include most of the adjectives normally used to describe people. These words span an enormous range, being able to modify any part of a characterization. The words range in complexity from "tall" and "blue eyed," which have fairly obvious meanings, to "ambitious" and "successful," where understanding the total affect on a characterization requires considerable thought. An important fact is that the meanings of these words often depend upon the context of the stereotype being modified. "Ambitious" does not have the same meaning in the phrases "ambitious lawyer" and "ambitious graduate student," as it would if we defined "ambitious" in a static fashion. This implies that some SM's function like programs that alter in a regular manner the frames set up by CSW's but do not have

explicit meanings in isolation.

Stereotype Modifiers can be used in two ways - in conjunction with CSW's ("John is an ambitious lawyer") or by themselves ("John is ambitious"). These two usages are quite similar, since in the latter case we are modifying the stereotype we use for people we know nothing about, in exactly the same way the first example modifies a lawyer stereotype. As a research strategy we have concentrated on SM's in combination with CSW's, since the changes and contradictions caused by the modifiers seem more dramatic and easier to study. Examining why the SM "ambitious" does not fit well with the CSW "hippie" tells us a great deal about the meanings of both words.

An understander can have one of two reactions to the modification of a characterization by an SM. Either the change simply updates memory creating a new person frame, or the new information from the SM clashes with the old, and the understander finds the formation of a coherent image in his mind difficult. The surprise can result from a combination of attributes which is impossible to comprehend, but more frequently it comes from an attempt to modify something we are reluctant to give up. An understander hearing about a "vicious priest," immediately recognizes that information in the priest stereotype that the person is gentle and compassionate is quite wrong.

Another way an understander can be surprised by an SM is if it tries to modify a part of a characterization which the current stereotype does not contain. The concept of a "greedy cheerleader," for instance, is very hard to comprehend, since "greedy" applies to situations which involve acquiring things, none of which appear in the cheerleader stereotype. Does it mean that a "greedy cheerleader" wants to have as many pom-poms as possible? The implications of this for what belongs is a person frame are large. To accommodate words such as greedy, is it necessary to know what kinds of objects a given stereotyped person might want to possess? The answer is that we believe it is necessary. However the alarm we might feel about having to put almost everything in the universe in a person frame is mitigated by the realization that such information is a natural part of what we know about the goals of an individual. The fact that goal information needs to be present should hardly be surprising.

Normally, an initial stereotype defines the basic range of a given property for a person, and the SM narrows it. For example, our stereotype for hippie includes a low value for ambition. We assume a "lazy hippie" (pronounced with the correct intonation) is even less ambitious than the stereotypical hippie. (With the phrase said as a phrase with uniform intonation, "lazy hippie" is simply a reaffirmation of one of the slots in the stereotype.) On the other hand, an "ambitious hippie" is ambitious relative to

the stereotypical hippie, and will have a few active goals. But what are an "ambitious hippie's" active goals? The fact that this is a more difficult question than for an "ambitious politician" or for an "ambitious professor" indicates that the default goal information for hippie is of a different nature than that for professor or politician.

We can see then that the class of SM's that exist must be dependent on the kinds of structures built up by CSW's. Words like "deceitful" or "ruthless" for example, refer to the planboxes (see Schank and Abelson, 1977) a person is likely to use in planning. This implies that CSW's may specify typical planbox choices (a debutante will use EXPRESS DESIRE when she wants something, a "thug" will use OVERPOWER) and that there are default planbox choices. Similarly, the scripts a person chooses to engage in can be set up or modified. Words such as "boring" or "iconoclastic" will modify default script information. For example, how does a "hooligan" get on a subway? Clearly he doesn't pay. This information cannot be stored for every script. Rather, a slot in the person frame is set so that when scripts involving societal norms are present, we can expect the hooligan to get around the norms, while we expect a solid citizen to conform to them. "Getting around" norms then becomes an important part of understanding. Programs that model this will have to identify each data point in a script and understand planful ways around them (like ducking under the turnstile in a subway).

These problems are very difficult. For now, we will concentrate on solving some of the simpler problems in combinations of SM's and CSW's, keeping in mind the issue of identifying what information belongs in a person's frame.

Consider the follow two stories:

(A) John was a compassionate professor. One of his students came in to complain about his grade on a term paper. After a brief discussion John told him ..."

(B) John was a coldhearted professor. One of his students came in to complain about his grade on a term paper. After a brief discussion John told him ..."

The two stories are the same except for the characterization of John. However, that one difference causes us to have very different expectations about how the two stories will conclude. In (A), from the meaning of the SM "compassionate," we know John will be concerned with other people in his professor activities, and will frequently give into their desires. So in (A) we expect John to raise his student's grade. In (B), on the other hand, we make exactly the opposite prediction. Since John is described as "coldhearted," he will ignore the problems of the student, and turn him down cold. Predictions of this nature are exactly the kind we need to understand stories of this sort, and would be impossible to make without the proper treatment of stereotypes.

As another example, we will consider how an understander would process the sentence from the Anderson and Bower experiment, "In the park, the hippie touched the debutante." Both "hippie" and "debutante" are excellent examples of CSW's, and each brings forth a very complex stereotype. When our understander reached "hippie," it would create a characterization with all the facts we commonly associate with such people (such as they haven't existed since 1969). It will also make predictions about what is likely to come next. Specifically we expect to hear about his appearance, his philosophy or some aspect of his carrying out a "free" lifestyle. But instead the sentence has him touching someone, the debutante. This certainly does not fulfill any expectations, and is very hard to understand. Probably an understander would try and fit the action to a prediction by assuming the touching satisfies some "hippie-type" goal, offending the establishment, perhaps. In any case, the processing would certainly be more complex than if the sentence had been, "In the park, the hippie was smoking pot."

Understanding of this story is further complicated by the CSW "debutante." When our understander reaches it, it will create another detailed characterization, and make still more predictions. It will also have to determine how the debutante will react to the event already mentioned, being touched by the hippie. From our debutante stereotype, we infer she will be appalled and predict she will scream

and run off. This processing is very different from what we would have gone through had the sentence been, "In the park the hippie touched the policeman." In this case top-down processing would have required totally different predictions. It is not clear whether they would be easier or more difficult to make (probably easier, since a policeman fits the context of a park much better) but they certainly would be different. As we have shown, a human processing sentences of this type must deal with stereotypes to have any hope of properly understanding them. As a result, any experiment which assumes words like "hippie" and "debutante" to be simple tokens, can easily be confounded by the complex understanding process we have just described.

CLAUDE

CLAUDE (Characterization-based Language Understanding Device), is a small program designed to help investigate the problems inherent in understanding stereotypes. The user gives CLAUDE a CSW and some number of Stereotype Modifiers, and the program replies with a few actions typical of the person.

CLAUDE does not deal with fully developed characterizations. Instead, its definitions for CSW's include only those pieces of information which are relevant to the Stereotype Modifiers it knows about. Specifically, CLAUDE is concerned with what plans a person will use for the goals he pursues, and the interpersonal relations he is

likely to be involved with. As we will see, even these parts of the frame are simplified.

A characterization in CLAUDE only contains information about two aspects of a person - goals and interpersonal relations. CLAUDE's representation of goals is oriented towards the plans used to achieve them. It knows little about the semantic interpretation of the goals. The representation of a goal in CLAUDE consists of three parts - the name of the goal, four descriptor scales which indicate in general terms how the person is apt go about achieving the goal, and a set of specific plans which the subject can use. Associated with each plan are ranges on each of the four descriptor scales indicating when the plan is likely to be used, which allow CLAUDE to make specific inferences about what actions of the subject it is likely to hear about.

In the goal segment of the characterization, it is the four descriptor scales for each goal which will be changed by Stereotype Modifiers. Changing them will in turn change the plans the person is likely to use. The four scales used are 1) Energy : How hard will the person try to achieve the goal. 2) Power : How strongly do the person's actions influence his attainment of the goal. 3) Ability : How effectively will the person's efforts be towards achieving the goal. 4) Compassion : How strongly does the person consider other people in his attempts to achieve the goal.

While these scales may not be definitive in describing what plans a person is likely to use, they do cover a wide range, and the effect of many Stereotype Modifiers can be well represented as changes to these scales, for one or more goals of the subject. There is one set of scales for each goal, and the values can differ widely. For instance, an undergraduate premed student will have as a goal getting into medical school. In trying to achieve this goal, he will be willing to expend large amounts of Energy, his Power is marginal, his Ability can assume a large range of values, and his Compassion is probably limited. On the other hand, he will also have some additional, possibly conflicting goals, such as having fun.

In the descriptor scales under each goal, as well as in the other scales CLAUDE uses, which we will see shortly, values can take one of three forms. They can be unknown. This means CLAUDE knows nothing about that scale, and it will totally ignored for all processing. The value can be a default. This means the current number will be used in processing, but it can be easily displaced. Defaults are the most common type of information in a characterization. Finally, the number can be certain. In this case it is used in processing, and it is modified, instead of being displaced if a Stereotype Modifier wants to change it.

As with goals, CLAUDE's representation of plans is rather simple, again avoiding the semantics behind the plans. Each plan consists of a name and a set of four numerical ranges, one for each of the descriptor scales of the goal it is in service of. CLAUDE will assume that a plan is likely to be used, if all four descriptor scales fall within the specified ranges associated with the plan. Either end of each range can be unspecified, in which case that boundary is ignored. If both are unspecified, the scale is ignored, meaning any value is considered in range. A sample plan under a premed's goal of getting into medical school might be:

```
Plan - negatively affect competition
   Energy      : 1  U
   Power       : U  5
   Ability     : U  U
   Compassion  : U -8
```

This means that for this plan to be used the subject descriptor scales under the get into medical school goal must have Energy greater than 1, Power less than 5 (or he would do something else), and Compassion less than -8 (since this plan is a crummy thing to do). Ability is not relevant. We know that any premed could potentially use this plan, but only the characterizations of a very few would be such that we actually expect the plan to be used.

The general plan "negatively affect competition" may take many forms in its actual implementation. Similarly, many SM words can call this plan to mind. Thus an "unscrupulous premed" might be expected to use this plan.

Its manifestation for him would be to examine the things that are necessary for his competition to succeed, and to sabotage them. To understand this fully, requires knowing many of the details of a premed's life. In the absence of knowing these details, we would expect an understander of this phrase to expect generally negative affects.

Interpersonal relations are represented in CLAUDE in a way similar to goals and plans. The characterization may contain many interpersonal relations. Each one is represented by a description of the other party of the relation, three scales and a set of actions which are likely to occur between the subject and the other party. The actions have associated ranges, which are used in a manner similar to the ranges associated with plans.

The three scales used to characterize interpersonal relations are those described in Schank and Abelson (1977) - positive/negative, dominant/submissive, and intimate/distant. The actions associated with each interpersonal relation are given a name and a set of three ranges, which are constructed like those of plans. As with plans, each interpersonal relation has its own set of scales. As an example of the types of actions CLAUDE might have, in an employee-boss interpersonal relation, two actions would be "ask for a raise," and "demand a raise," (and threaten to quit). Which of these actions we would expect would depend upon the values in the positive/negative

and dominant/submissive scales for this relation.

CSW's in CLAUDE consist of definitions of goals (with their plans) and interpersonal themes (with their actions). Along with the definitions of the goals and interpersonal relations are initial values for the attached scales. As an example of what makes up a CSW, consider an assistant professor's goal to get tenure, as understood by CLAUDE.

An assistant professor might be expected to have a great many goals. As part of our knowledge of assistant professors, we can expect one of these to be "get tenure." The available plans for these are known to us if we have detailed knowledge of this profession, or else can be just inferred from rules about "getting higher-ups to like you," "working hard" and so forth, that are attached to what we know about "promotion" in general.

A word modifying a position CSW tends to specify aspects of that position. Thus, "ambitious assistant professor," or "lazy assistant professor" tend to alter the default values on certain goals, rather than just combining the defaults for the SM and CSW in general. Someone who is both "lazy" and an "assistant professor" may not be lazy in his goals related to his profession. But, the combination, when stated as a phrase, tends to imply just that, and it does not imply that he will necessarily be lazy in other aspects of his life.

"Lazy" then, can be seen to set the energy value for the CSW. The question is, "lazy for what?" As before, we must examine the goals and plans of the CSW and find what the person must do to get what he wants. One possibility, if we have the information, is to find the tenure goal, establish that the normal plans are "write papers" and "go to conferences," and recognize that with low energy an assistant professor is not highly likely to do either of those things.

Similarly, a "compassionate assistant professor" forces us to find a person for him to be compassionate to (the default is any known underling - a student probably) and then find out specifically what his relationship to that person is. For this relationship the initial values we use are:

```
Interpersonal relation - assistant-professor/student
          Positive/Negative   : -1 default
          Intimate/Distant    :  3
          Dominant/Submissive :  7
```

CLAUDE's processing of CSW's is straightforward. For each goal or interpersonal theme in the definition, it creates a token to hold information. To that token it attaches the initial values of the appropriate scales, as indicated by the definition. From then on it makes all modifications to the tokens, and looks at them when it wants to produce a description of the person.

Stereotype Modifiers can alter either segment of one of CLAUDE's characterizations (goals or interpersonal relations), or both. If it modifies goals, it has a list of values for the four descriptor scales. (Some of which may be unspecified.) Similarly, if it modifies interpersonal relations it has values for the three scales used in their description. The definition can also include restrictions of the goals and interpersonal relations it should apply to. CLAUDE provides the testing mechanism to determine how well various definitions work.

A typical CLAUDE Stereotype Modifier definition is the following one for compassionate.

```
Stereotype Modifier - compassionate
 GOAL MODIFICATION
        (no restriction)
        Energy      : U
        Power       : U
        Ability     : U
        Compassion : 7
 INTERPERSONAL RELATION MODIFICATION
        (restriction value on DOMINANT/SUBMISSIVE scale > 0)
        POSITIVE/NEGATIVE    : 5
        INTIMATE/DISTANT     : 5
        DOMINANT/SUBMISSIVE : U
```

This definition tells CLAUDE that a compassionate person is (naturally enough) more compassionate in achieving all his goals. The other goal scales are not affected. In Interpersonal Relations, he will be more postive and intimate in all relations he is dominant in. (The restriction is, by the way, why the phrase "compassionate serf" is odd; there are no dominant relations of interest created by "serf" for "compassionate" to modify.)

CLAUDE processes Stereotype Modifiers by checking to see if the word modifies goals and/or interpersonal themes. It then tries all appropriate tokens, and for each that passes any restriction the Stereotype Modifier might have, it combines the old and new values for scale according to the following rules.

1) If the new value is unspecified, use the old one.

2) If the old value is unknown, use the new one.

3) If the old value is a default, use the new one.

4) Otherwise, merge the two values.

   a) If the values are on the same side of 0,
      move the old value away from zero.

   b) If the values are on the opposite sides of 0,
      move the old value towards zero.

The first three rules are rather intuitive. The fourth is the most interesting case. 4a is when the modifier is being used for emphasis, as in "compassionate priest," and 4b is contrast, "compassionate landlord." The exact numeric manipulations done are not important, but it is significant that simple functions such as average cannot be used. Instead, if we apply a modifier with a positive dimension of compassion to a characterization where the person is already thought to be compassionate, we should make the person more compassionate, whether or not the new value is greater or less than the old.

CLAUDE generates a description by cycling through all the goals and interpersonal relations it knows about. For each goal, it checks to see if the current descriptor slots are all within the range of any plans. If so it generates a brief statement indicating which plans will be used for that goal. The English generation should not be taken seriously. What CLAUDE chooses to generate should be. After it finishes with goals, interpersonal relations are processed in the same way, looking to see if the current scale values are in range of any actions. What follows is a computer run of CLAUDE on a few examples.

```
[PHOTO:  Recording initiated  Wed 17-Jan-79 8:16PM]

 TOPS-20 Command processor 3(414)

YALE/RUTGERS/UCI LISP - 29 July 77

*(CLAUDE)

Stereotype...? *assistant-professor

  IN PURSUIT OF HIS GOAL OF get-tenure HE WILL PROBABLY
  write-papers

  IN HIS DEALINGS WITH students, HE WILL PROBABLY
  do-recommendations

Stereotype...? *lazy assistant-professor

  IN PURSUIT OF HIS GOAL OF get-tenure HE WILL PROBABLY
  not-worry-about-it

  IN HIS DEALINGS WITH students, HE WILL PROBABLY
  do-recommendations

Stereotype...? *ruthless assistant-professor

  IN PURSUIT OF HIS GOAL OF get-tenure HE WILL PROBABLY
  write-papers

  IN HIS DEALINGS WITH students, HE WILL PROBABLY
  nail-to-wall

Stereotype...? *premed

  IN PURSUIT OF HIS GOAL OF get-into-med-school
  HE WILL PROBABLY  study

Stereotype...? *ruthless premed

  IN PURSUIT OF HIS GOAL OF get-into-med-school
  HE WILL PROBABLY sabotage-other-predmeds

Stereotype...? *done

BYE

[PHOTO:  Recording terminated  Wed 17-Jan-79 8:17PM]
```

There are several interesting points about this session. The first example shows what CLAUDE believes an assistant professor will do, if it has no additional information work with. Its comment on getting tenure is from the goal segment of the characterization, and on dealings with students is from the interpersonal relation segment. The next two examples show how the assistant professor stereotype can be modified, first by "lazy," then by "ruthless." The responses CLAUDE gives seem to correspond nicely with how a person might describe a lazy or ruthless assistant professor. The next example shows an unembellished premed. CLAUDE's final description is of a "ruthless premed." This is interesting because it shows how the use of scales allows the same modifier, "ruthless," in this case, to have different effects depending upon what is already in the stereotype. As we would hope, CLAUDE does not believe a "ruthless premed" behaves in the same way as a "ruthless assistant professor."

Conclusion

In this paper we have shown that characterization must be handled properly if we wish to develop a reasonable predictive processing system. We can certainly not treat words like "lawyer" and "professor" as creating simple, unstructured tokens. Stereotypes give us access to enormous amounts of information, and serve as a vital part of predictive processing.

Finally, to end the suspense the reader has been in from the outset of this paper, no, a hippie does not own a hairdrier, a professor only owns a Mercedes if he has outside business interests, a business executive wears bermuda shorts only on Saturdays and on Wednesday, if he has a golf date and most assuredly, a jetsetter has absolutely no interest in meeting an AI researcher.

References

Anderson, J. and Bower, G. (1973) Human Associative Memory. Winston-Wiley, Washington, D. C.

Bobrow, D. G. and Winograd, T. (1977) An Overview of KRL, a Knowledge Representation Language. Cognitive Science, 1977, Vol. 1, pps. 3-46.

Carbonell, J. G. Jr. (1979) Subjective Understanding: Computer Models of Belief Systems. Research Report 150, Department of Computer Science, Yale University.

DeJong, G. F (1977) Skimming Newspaper Stories by Computer. Research Report 104, Department of Computer Science, Yale University.

Minsky, M. (1975) A Framework for Representing Knowledge, in Winston, P. H. (ed.), The Psychology of Computer Vision, McGraw-Hill, New York.

Riesbeck, C. K. (1975) Conceptual Analysis. Schank, R. C. (ed.), Conceptual Information Processing, North-Holland, Amsterdam.

Rosch, E. (1975) Cognitive representations of semantic categories. Journal of Experimental Psychology: General, Vol. 105, pps. 192-305.

Schank, R. C. and Abelson, R. P (1977) Scripts, Plans, Goals, and Understanding, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Wilensky, R. (1978) Understanding Goal-Based Stories. Research Report 140, Department of Computer Science, Yale University.