RECOGNIZING HUMOR IN NEWSPAPER CARTOONS BY RESOLVING AMBIGUITIES THROUGH PRAGMATICS

Lawrence Mazlack
Noemi M. Paz

ABSTRACT

Newspaper cartoons can graphically display the results of ambiguity in human speech. The result can be unexpected and funny. Captioned cartoons derive their humor from a sudden incongruity which can be made to follow by a human being who can automatically use stored world knowledge to resolve the ambiguous situation.

Likewise, computer analysis of natural language statements also needs to successfully resolve ambiguous situations. Computerized understanding of dialogue that takes place between humans must not only include syntactical and semantical analysis, but also pragmatical analysis. Pragmatics consists of an understanding of the speaker's intentions, the context of the utterance, and social implications of polite human communication.

Computer techniques have already developed been use restricted world knowledge in resolving ambiguous language use. This paper illustrates how these techniques can be used in resolving ambiguous situations arising in cartoons.

1. THE GENERAL ROLE OF PRAGMATICS IN NATURAL LANGUAGE UNDERSTANDING

Within linguistic theory, the study of language use can be called pragmatics. One defintion of pragmatics developed by Charles Morris (1946) is that pragmatics can be characterized by the relationship between signs and their human users. Signs fall into three classes: icons, indices, and symbols. Pragmatics relates directly to signs that are indices because indices can only be understood when they are actually used.

The meanings of indices can be found by describing rules for relating the sign to a context. These are pragmatic rules which are in essence "action" rules for "finding" relationships. The set of structures developed for describing these rules are called pragmatic-semantic "trees" and divide into three catesories:

1. Performatives - which describe the speaker's intention or goal in using a sentence as a question, a command, etc.

2. Presuppositions - assumptions about the context that are necessary to make that sentence verifiable, or appropriate, or both.

3. Conversational Postulates - a class of presuppositions concerning the nature of human dialogue which can be referred to as discourse codes of conduct. (Bates,1976)

For semantic-pragmatic structures to operate, it is not enough to determine the meanings of individual words. Other types of information must be accessed. In order to select between competing meanings, knowledge is required about the grammatical functions represented by particular word orders in the natural language sentence. Also, knowledge about the "real world" (presuppositions) is needed; i.e., the context in which the utterance took place.

Along with contextual knowledge, a semantic-pragmatic structure needs to account for "speech acts" (performatives). Speech acts demonstrate the speaker's goal. They can be a command, a question, a statement, etc. In other words, the associated meaning and the implied action must be understood. The theory must account for the fact that the listener or reader of the statement understands this double "meaning". (Bates, 1976)

Next, the semantic-pragmatic structure must explain the speaker's ability to understand sequences of language which should mean one thing but clearly mean another (conversational postulates). It is assumed that normal human beings who enter into a conversation have agreed to be cooperative. This means speakers will tell each other the truth, that they will only offer information assumed to be new and relevant to the listener, and will only request information which they sincerely want. This represents a set of standard rules. Deviations from this "code of conduct" will be seen as violations.

2. SUPPLYING PRAGMATICS FOR COMPUTER ANALYSIS

There are several question answering systems that make use of various techniques for including pragmatic analysis in understanding a natural language. As these techniques are described, a cartoon will be analyzed to illustrate how pragmatic analysis could be used to disambiguate the situation. The characters that correspond to the computer and those that correspond to a human in a man-machine dialogue will be identified. These cartoons are found in the Appendix.

2.1 COOPERATIVE DIALOGUE

The CO-OP System (Kaplan, 1979) is a question answering data base system that follows the "codes of conduct" presented earlier. Its objective is to provide cooperative responses from a natural language data base query. Some examples from this system follow.

CO-OP is able to determine from a

question not only what information is required, i.e. the direct, literal, and correct response, but also that the questioner is unaware of highly pertinent facts not explicitly requested in the question. A heuristic used by the system is knowledge of such facts frequently makes asking the question unnecessary, because they entail an answer. The system action is to ignore the question and provide the pertinent fact. For example:

> Question: How many students failed CSE110 in Spring, 77?
>
> System's answer is: CSE110 was not given in Spring '77.

The answer of "zero" would not have been cooperative.

The user who posed the above question presumed that the CSE110 class was taught in the Spring of 1977. The system on finding that this presumption was false responds with a "corrective indirect response" by supplying the negated presumption.

Cartoons often lead to funny results when their statements are ambiguous. In the cartoon TIGER (appendix, fig. 1) there is an example of a cooperative response in the answer to the question: "Did he catch him?". Prior to this question the human in the dialogue only knows there is a chase going on and that there are two participants: Stripe and Mrs. Parker's cat. The situation is ambiguous because we do not know who is chasing whom. Here the common human presumption is dogs usually chase cats and therefore Stripe must be a dog. The computer system on finding this presumption to be false, could respond with a corrective indirect response: "Nope, Stripe got away" rather than with the direct answer of "no".

Another type of cooperative response the CO-OP system replies with is a suggestive indirect response. Following the "codes of conduct", it is appropriate for an answer to contain relevant information likely to be requested in a follow-up question. A heuristic used here is to change the focus of the original question and to respond with a direct answer to the original question, but with the focus changed. The focus is that aspect of the question which is most likely to shift in a follow-up question.

The BC cartoon (appendix, fig. 2) could have used a suggestive indirect response. In this cartoon, a census taker (the human) is asking questions of a subject (the computer). A common presumption here is the first question the census taker will ask is the subject's name. The computer needs to realize the ambiguous situation created by the question: "Name, please". If a change of focus is analyzed the question could be seen as two questions:

> 1) What is your (subject's) name, please?
>
> 2) What is your dog's name, please?

The computer on realizing the possible ambiguous situation could then respond:

> 1) Fido is my dog's name.
> OR
> 2) "Computer" is my name.

2.2 RESTRICTED DOMAIN OF DISCOURSE

Another data base system, ROBOT (Harris, 1978), uses the data base itself to find the use of words in the question, to build expectations, and to resolve ambiguities. The system interprets input based on what makes sense within its limited world model, the data base.

In processing ambiguous statements, several interpretations may arise. A heuristic used is unintentional interpretations of input questions are usually not false for the specific domain, but have a vacuous response (Coles, 1972). To use this heuristic these interpretations are posed to the data base as queries. If all interpretations fail to find a response to the question, then the answer is "there aren't any". This negative answer assumes that the dialogue will only be about information contained in the data base. If more than one interpretation can be answered successfully, then the system enters into a clarification dialogue, just as humans would have to do when faced with an ambiguous question. If exactly one interpretation that is found, then the system responds using this interpretation for the question.

In the cartoon GARFIELD (appendix, fig. 3) the human speaker is asking Garfield to play with Nermal. The following ambiguous situation is created:

> 1) "play with Nermal" means that Nermal is a toy
>
> OR
>
> 2) "play with Nermal" means that Nermal and Garfield should play together.

The computer system could resolve this ambiguous situation by searching Toy and Friend domains for the entry Nermal. On not finding Nermal in the Toy domain but in the Friend domain the ambiguous situation is resolved.

The LIFER System (Hendrix, 1978) has capabilities for extending the natural language subset that is understood by the system. Users may employ easy-to-understand notions such as synonyms and paraphrases to extend the language. The users can then ask questions about information contained in the data base using their own natural language "style". In this way "utterances" by the speaker (user) can be understood by the listener (computer).

In the cartoon WIZARD OF ID (appendix, fig. 4) the human is stating to the Sire (computer) that "our records show there is a dip in unemployment" and is perhaps implying the question "what do we do next?". The ambiguous situation here is

that "a dip" could describe either a foolish person or a downward trend in a statistic or figure. To resolve this ambiguous situation the human could have entered the paraphrase:

> "Dip in unemployment" is a paraphrase of "temporary decline in the unemployment statistic".

2.3 WORLD KNOWLEDGE WITH FRAMES OR SCRIPTS

Many other systems have included world knowledge and information related to the dialogue with a user in a frame or script. Frames, simply put, are just highly structured sets for keeping pragmatics. When an action is carried out, some canonical description is stored in the frame that will permit the program to reconstruct the context in which the event took place. Frames carry over to the subsequent statements and conventions that mark anaphora or presupposition link the program to slots in the current or past frames that will resolve the reference.

Frames not only include syntactic information, eg. subject, object, prepositional phrases, but also semantic and pragmatic facts which provide various reasons, motivations and purposes not explicitly stated.

Scripts are like frames in that they also have empty slots that are filled with the context from a dialogue or text. However, scripts provide world knowledge about common experiences or situations in terms of Schank's conceptual dependency primitives. A text is understood by mapping sentences into actions or primitive acts as described in the script. Unstated facts described in a script but not in the sentence are assumed to be true. This provides a "background" or world knowledge for understanding and reasoning. (Schank 1975).

The BEETLE (appendix, fig. 5) cartoon can be used to illustrate the frame and the script concepts. Here the relative pronoun must be resolved in the phrase "that gun". The two possibilities are:

> 1) shoot at that gun, i.e. use that gun as a target
>
> 2) use that gun to shoot with and shoot at the original target.

The ambiguous situation can be resolved by using world knowledge about target practice. For example, it is helpful to know that targets should not be expensive, useful things, i.e. a gun, and that a target is not located near a human being. This type of information can be provided by demons in the case of frames or by the reason or goals statement in the case of scripts.

2.4 RECOGNITION OF HUMOR

Humor due to ambiguous statements requires the reader to recognize that an ambiguous interpretation has occurred. The humorous interpretation is the unexpected one. Through the use of pragmatic analysis, humorous interpretations can be recognized as well as generated.

SUMMARY

Cartoons can graphically represent the humor due to ambiguities in human speech. It may be possible to recognize humorous ambiguities using already existing techniques. Three levels of the use of pragmatics have been described. The first is to resolve double meanings by restricting the domain of discourse, eliminating the occurrence of double meanings. This device is employed by LIFER in restricting the language and by ROBOT by restricting the objects in the domain to only those that appear in the data base.

The second level of the use of pragmatics involves making the domain larger rather than restricting it. Frames and scripts are used to involve more world knowledge in the natural language analysis so as to disambiguate based on what is common occurrences in a given situation. The third level involves an ability to deduce from complex frames and scripts a purpose and then acting in agreement with that purpose. CO-OP is capable of determing the questioner's motive, and if necessary, posing for itself a question more in keeping with the questioner's motive than the original question.

Pragmatic devices used include proscribing context, enlarging context, and deducing motivation from context.

REFERENCES

Barr, A. 1980. "Natural Language Understanding", AI Magazine, Vol. 1, No. 1, Spring.

Bates, E. 1976. Language and Context: The Acquisition of Pragmatics, Academic Press, New York.

Intelligence: Can Computer Think?, Boyd & Fraser Publishing Co. San Francisco.

Coles, L.S. 1972. "Syntax Directed Interpretation of Natural Language", in Representation and Meaning: Experiments with Information Processing Systems. H.A. Simon and L. Siklossy (eds.) 44-66.

Elgin, S.H. 1979. What Is Linguistics?, 2nd ed., Prentice-Hall, Inc., Englewood Cliffs, NJ.

Goldstein, I. and Papert S. 1976. "Artificial Intelligence, Language and the Study of Knowledge", MIT, AI Memo. 337, March.

Grosz, B. 1980. "Utterance and Objective: Issues in Natural Language Communication", AI Magazine, Vol. 1 No. 1, Spring.

Harris, L.R. 1978. "Using the Data Base Itself as a Semantic Component to Aid in the Parsing of Natural Language Data Base Queries", Dartmouth College, TR 77-2, October.

Hendrix, G., Sacerdoti, E., Sagalowicz, D., and Slocum, J. 1978. Developing a natural language interface to complex data. ACM transactions of database

systems 3: 105-147.
Kaplan, S. 1979. cooperative responses from a portable natural language data base query system. doctoral dissertation dept of computer and information science, university of pennsylvania.
Mazlack, L.J. 1979. "Some Considerations in Mapping Natural Languages Queries onto Data Base Systems". Working Paper, September.
Schank, R.G. 1975. "The Structure of Episodes in Memory". Representation and Understanding Studies in Cognitive Science, D.G. Bobrow and A. Collins (eds.), Academic Press, New York.
Siklossy,L. and Simon, H.A. 1972. "Some Semantic Methods for Language in Representation and Meaning: Experiments with Information Processing Systems. H.A. Simon and L. Siklossy (eds.) 44-66.
Siklossy, L. 1978. impertinent question-answering systems: justification and theory. ACM proceedings annual conference Washing, d.c. vol 1 39-44.
Waltz, d July, 1978. an english language question answering system for a large relational database. CACM vol 21 no 7, 526-539.

APPENDIX

FIGURE 1

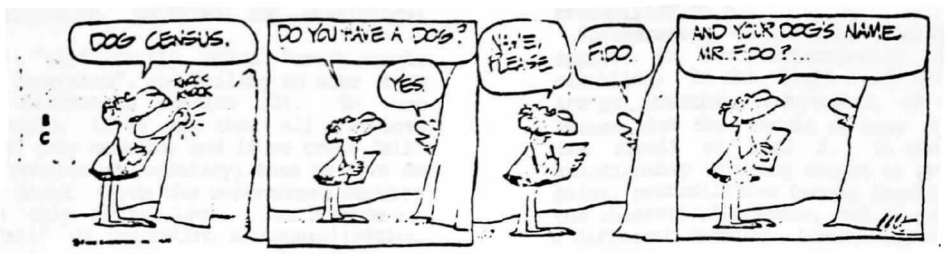

FIGURE 2

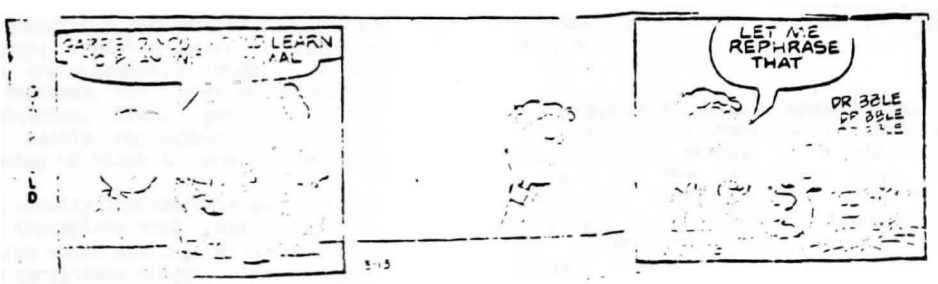

FIGURE 3

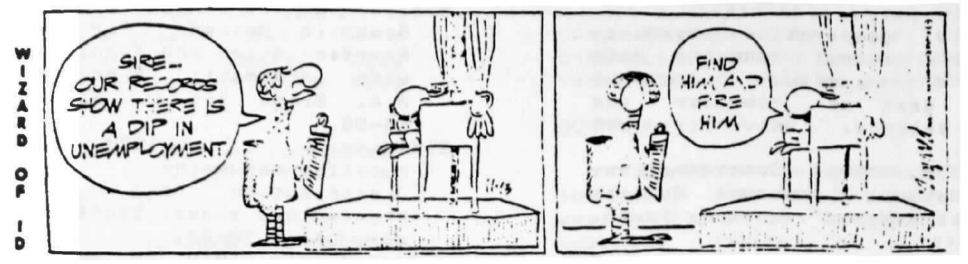

FIGURE 4

FIGURE 5