New Mexico State University
Computing Research Laboratory

Team members:
Yorick Wilks, David Farwell,
Afzal Ballim, Roger Hartley

CRL's contribution to DARPA's program is to bring to bear on natural language understanding two closely-related belief and context mechanisms: dynamic generation of nested belief structures (ViewGen) and hypotheses for reasoning and problem-solving (MGR).

Cooperation and Planning in a Multiactor System. The ViewGen project investigates theoretical issues in the area of belief systems that pertain to human-computer interaction (communication and cooperative planning). We are using the results to implement a system that reasons and interacts with people in a limited but real domain and incorporates the first perspicuous default algorithm for belief ascription in a concrete domain. Research has shown that effective communication between a computer and a human-in other words, the system and the user-requires modeling of the various beliefs which each has about the topic of conversation. This project is aimed at generating, from the system's own beliefs, the views of people on demand. These views can then be used in reasoning and cooperative processes. This project is the first to offer a theory of dynamic construction of nested belief structures, or viewpoints, and the heuristics associated with such constructions. We developed an initial program called ViewGen that generates nested viewpoints (what some person believes is another person's view of some topic), and the results and insights obtained from this program are being used to develop the larger belief system.

Model Generative Reasoning. Current expert system technology performs effectively on well-defined problems within closed worlds. However, it is brittle when problems are ill defined, data are incomplete, and solutions require integration of knowledge from many different subject domains. These conditions characterize many real-world applications.

The model generative reasoning (MGR) system is a significant advance in existing technology. The MGR algorithm provides a general framework for constructing, comparing, and evaluating hypothetical models of queried events using abductive assembly; that is models are assembled from definitions of general domain concepts to provide alternative explanations for the query and related assumptions.

Explanations are developed progressively through a generate-evaluate cycle. Assumptions are interpreted using concept definitions and then joined to form alternative descriptions (contexts) of the domain structures. Contexts are merged next with procedural information to form programs. These programs can then be run in an environment of facts (observations and necessary truths) to generate models. Last, models are evaluated for their parsimony and explanatory power, providing either a problem solution(s) or the assumptions for the next round of interpretation.

Comparison and Evaluation of Parser Performance. Our proposal for the evaluation of natural language processing systems is an adaptation of the objective procedures for evaluating machine translation systems developed during the 1960's and 1970's. It is

designed to avoid system particular or approach particular bias. We propose to treat individual analysis systems (parsers) as black boxes which cannot be adjusted during experimental trials. We establish (1) a standard vocabulary, (2) a standard generator and (3) a set of structure protocols which are applied to the output of the parsers. Each parser is run on a given text (containing only the standard vocabulary). The resultant representation is "interpreted" by the structure protocols so as to produce an appropriate representation for the standard generator. The generator then produces a second text or set of texts. The texts, input and output, are, in turn, used as a test corpus for various well-established procedures in which, generally, human subjects are asked if they can abstract more information from the input text (having been first provided with an output text) or from an output text (having been first provided with the input text). To the extent that the subjects are able to do so, the black box analysis has failed since if the two texts were perfect paraphrases no such information could possibly be gained or lost.