

Philip R. Cohen
 Center for the Study of Reading
 University of Illinois, &
 Bolt, Beranek and Newman, Inc.
 Cambridge, Mass.

This panel was asked to consider how various "problem contexts" (e.g., cooperatively assembling a pump, or Socratically teaching law) influence the use of language. As a starting point, I shall regard the problem context as establishing a set of expectations and assumptions about the shared beliefs, goals, and social roles of those participants. Just how people negotiate that they are in a given problem context and what they know about those contexts are interesting questions, but not ones I shall address here. Rather, I shall outline a theory of language use that is sensitive to those beliefs, goals, and expectations.

The theory is being applied to characterize actual dialogues occurring in the familiar task-oriented situation [10], in which an expert instructs a novice to do something, in our case to assemble a toy water pump. In such circumstances, the dialogue participants can be viewed as performing speech acts planned, primarily, to achieve goals set by the task. Other contexts undoubtedly emphasize the instrumental uses of language (e.g., [11]) but those problem contexts will not be considered here. The application of a model of speech act use to actual dialogue stresses the need for sources of evidence to substantiate predictions. The purpose of this paper is to point to one such source -- speaker-reference [3].

The natural candidate for a theory of instrumental use of speech acts is an account of rational action [12] -- what is typically termed "planning". However, contrary to the assumption of most planning systems, we are interested in the planning of (usually) cooperative agents who attempt to recognize and facilitate the plans of their partners* [1,4,5,16,20]. Such helpful behavior is independent of the use of language, but is the source of much conversational coherence.

A plan based theory of speech acts specifies that plan recognition is the basis for inferring the illocutionary force(s) of an utterance. The goal of such a theory is to formalize the set of possible plans underlying the use of particular speech acts to achieve a given set of goals. In light of the independent motivation for plan generation and recognition, such a formalism should treat communicative and non-communicative acts uniformly, by stating the communicative nature of an illocutionary act as part of that act's definition. A reasoning system, be it human or computer, would then not have to employ special knowledge about communicative acts; it would simply attempt to achieve or recognize goals.

The components of speech act planning and recognition systems developed so far include: a formal language for describing mental states and states of the physical and social worlds, operators for describing changes of state, associations of utterance features (e.g., mood) with certain operators, and a set of plan construction and recognition inferences. Illocutionary acts are defined as operators that primarily affect the mental states of speakers and hearers [3,8,13,17].

To be more specific, in the most fully developed attempt at such a theory, Perrault and Allen [13] show how plan recognition can "reason out" a class of indirect speech acts. Briefly, they define "surface" speech act operators, which depend on an utterance's mood, and operators for illocutionary acts such as requesting. Plan recognition involves inferences of the form "the agent intended to perform action X because he intended to achieve its effect in order to enable him to do some other action Y". Such inferences are applied to surface speech act operators (characterizing, for instance, "Is the salt near you?") to yield illocutionary operators such as

* For this brief paper, I shall have to curtail discussion of the planning/plan recognition literature.

requests to pass the salt.

The remainder of this paper attempts to illustrate the kinds of predictions made by the theory, and the use of anaphora to support one such prediction.* Consider the following dialogue fragment (transmitted over teletype) in the water pump context described earlier:

Expert: 1). "We need a clear bent tube for the bottom hole."
 Novice: 2). "OK, it's done."
 Expert: 3). "OK, now, start pumping"

The example is constructed to illustrate my point, but it does not "feel" artificial. Experiments we are conducting show analogous phenomena in telephone and teletype modes.

The theory predicts two inference paths for utterance 1 -- "helpful" and "intended". In the former case, the novice observes the surface-inform speech act indicated by a declarative utterance, and interprets it simply as an inform act that communicates a joint need. Then, because the novice is helpful, she continues to recognize the plan behind the expert's utterance and attempts to further it by performing the action of putting the spout over the hole. The novice, therefore, is acting on her own, evaluating the reasonableness of the plan inferred for the expert using private beliefs about the expert's beliefs and intentions. Alternatively, she could infer that the expert intended for it to be mutually believed that he intended her to put on the tube. Thus, the novice would be acting because she thinks the expert intended for her to do so. Later, she could summarize the expert's utterance and intentions as a request [7]. Perrault and Allen supply heuristics that would predict the preferred inference route to be the "intended" path since it is mutually believed that putting the tube on is the relevant act, and his intending that she perform pump-related acts is an expected goal in this problem context. To use Perrault and Allen's model for analyzing conversation, such predictions must be validated against evidence of the novice's interpretation of the expert's intent.

Signalling Interpretation of Intent

For this problem context and communication modality, the novice and expert shared knowledge that the expert will attempt to get the novice to achieve each subgoal of the physical task, and the novice must indicate successful completion of those subtasks. However, not all communicative acts achieving the goal of indicating successful completion provide evidence of the novice's interpretation of intent. For instance, the novice might say "I've put the bent tube on" simply to keep the expert informed of the situation. Such an informative act could arise if the problem context and prior conversation did not make the salience of putting the tube on mutually known. To supply evidence of the novice's interpretation of intent, her response must pragmatically presuppose that interpretation.

In our example, the novice has used "it" to refer to the action she has performed. It has been proposed that definite and pronominal/pro-verbal reference requires mutual belief that the object in question is in focus [10,15] and satisfies the "description" [6,14]. Assuming that the inferring of mutually believed goals places them in focus [15], the shared knowledge needed to refer using "it" is supplied by only one of the above interpretations -- the one summarizable as an indirect request.

* Robinson [15] has identified this problem of reference to actions and has implemented a system to resolve them. In this paper, I stress the importance of that work to theories of speech act use.

Other signals of the interpretation of intent need to be identified to explain how the expert's "OK, now start pumping" communicates that he thinks she has interpreted him correctly -- mutual signalling of intent and its interpretation is central to conversational success.

A formal theory that could capture the belief, intention, and focus conditions for speaker-reference is thus clearly needed to validate models of speech act use. A plan-based theory might accomodate such an analysis via a decomposition of currently primitive surface speech acts to include reference acts [2,18]. By planning reference acts to facilitate the hearers' plans (cf. [4]), a system could perhaps also answer questions cooperatively without resorting to Gricean maxims or "room theories" [19].

I have given a bare bones outline of how a description of speaker-reference can serve as a source of empirical support to a theory of speech acts. However, much more research must take place to flesh out the theoretical connections. I have also deliberately avoided problems of computation here, but hope the panel will discuss these issues, especially the utility of computational models to ethnographers of conversation.

Acknowledgements:

I would like to thank Chip Bruce, Scott Fertig, and Sharon Oviatt for comments on an earlier draft.

References:

1. Allen, J. A plan-based approach to speech act recognition (Tech. Rep. No. 131/79). Toronto: University of Toronto, Department of Computer Science, January 1979.
2. Appelt, D. Problem-solving applied to language generation. (This volume).
3. Bruce, B. Belief systems and language understanding (BBN Report No. 2973). Cambridge, Mass.: Bolt, Beranek and Newman, January 1975.
4. Bruce, B., & Newman, D. Interacting plans. Cognitive Science, 1978, 2, 195-233.
5. Carbonell, J. G. Jr. POLITICS: Automated ideological reasoning. Cognitive Science, 1978, 2, 27-51.
6. Clark, H. H., & Marshall, C. Definite reference and mutual knowledge. In A. K. Joshi, I. A. Sag, & B. L. Webber (Eds.), Proceedings of the Workshop on Computational Aspects of Linguistic Structure and Discourse Setting. New York: Cambridge University Press, in press.
7. Cohen, P. R., & Levesque, H. L. Speech acts and the recognition of shared plans. In Proceedings: Annual meeting of the Canadian Society for the Computational Study of Intelligence, Victoria, B.C., 1980.
8. Cohen, P. R., & Perrault, C. R. Elements of a plan-based theory of speech acts. Cognitive Science, 1979, 3, 177-212.
9. Donnellan, K. Speaker references, descriptions, and anaphora. In P. Cole (Ed.), Syntax and semantics (Vol. 9): Pragmatics. New York: Academic Press, 1978.
10. Grosz, B. The representation and use of focus in dialogue understanding (Technical Note 151). Menlo Park, Calif.: Stanford Research Institute, Artificial Intelligence Center, July 1977.
11. Hobbs, J. R., & Evans, D. E. Conversation as planned behavior (Technical Note 203). Menlo Park, Calif.: Stanford Research Institute, Artificial Intelligence Center, 1979.
12. Morgan, J. L. Toward a rational model of discourse comprehension. In D. Waltz (Ed.), Proceedings: Theoretical Issues in Natural Language Understanding. Urbana: University of Illinois, Coordinated Science Laboratory, 1978.
13. Perrault, C. R., & Allen, J. F. A plan-based analysis of indirect speech acts. In submission.
14. Perrault, C. R., & Cohen, P. R. Inaccurate reference. In A. K. Joshi, I. A. Sag, & B. L. Webber (Eds.), Proceedings of the Workshop on Computational Aspects of Linguistic Structure and Discourse Setting. New York: Cambridge University Press, in press.
15. Robinson, A. E. The interpretation of verb phrases in dialogs (Technical Note 206). Menlo Park, Calif.: Stanford Research Institute, Artificial Intelligence Center, 1980.
16. Schank, R., & Abelson, R. Scripts, plans, goals, and understanding. Hillsdale, N.J.: Erlbaum, 1977.
17. Schmidt, C. F. Understanding human action. In Proceedings of the conference on Theoretical Issues in Natural Language Processing. Cambridge, Mass., 1975.
18. Searle, J. R. Speech acts: An essay in the philosophy of language. Cambridge: Cambridge University Press, 1969.
19. Shannon, B. Where-questions. In Proceedings of the Seventeenth Annual Meeting of the ACL, San Diego, 1979. Pp. 73-75.
20. Wilensky, R. Understanding goal-based stories (Research Rep. No. 140). New Haven, Conn.: Yale University, Department of Computer Science, September 1978.