

CONVERSATION AS PLANNED BEHAVIOR

Jerry R. Hobbs
SRI International
Menlo Park, California

In this paper, planning models developed in artificial intelligence are applied to the kind of planning that must be carried out by participants in a conversation. A planning mechanism is defined, and a short fragment of a free-flowing videotaped conversation is described. The bulk of the paper is then devoted to an attempt to understand the conversation in terms of the planning mechanism. This microanalysis suggests ways in which the planning mechanism must be augmented, and reveals several important conversational phenomena that deserve further investigation.

1. Introduction

Perhaps the most promising working hypothesis for the study of conversation is that the participants make use of planning mechanisms much like those developed in artificial intelligence. The purpose of this paper is to explore some issues that arise in attempting to use this working hypothesis as the basis for a theory of conversation.

An empirical theory is an elegant description of a selected body of data. This characterization indicates three steps, if we are to construct such a theory:

1. to specify the data;
2. to define the language of description;
3. to describe.

These simply stated steps hide subtleties, some of which are explored below. Specifically, in section 2, we take up the question of just what the data of a theory of conversation can be. In section 3, the language of description, viz. a language for the operations of a planning mechanism, is presented.

The bulk of the paper, sections 4 and 5, is then devoted to an exercise in description -- the microanalysis of an actual conversation. This is a way of investigating two mutually recursive problems. The first is -- given an adequate planning mechanism with access to adequate conversational strategies and world knowledge, show that such data could be produced. This, in a sense, is a hand simulation. The second problem is -- assuming the data was produced by a planning mechanism, what sort of planning mechanism would it have to be? That is, the data is used to put pressure on the formalism. The computational mechanisms suggested by this exercise are discussed in section 6. In section 7, some general properties of conversation are discussed in the form of questions for future research.

None of this can tell us what went on in the actual production of the conversation, only what could have gone on. But in this, it is no different from other explanations in cognitive science. The best we can do is to tell a story that contradicts nothing we know.

2. What is the Data?

A theory of conversation concerns itself with utterances which are appropriate in particular contexts to particular conversational goals. Among the goals are goals external to the conversation, such as a task jointly engaged in (Grosz 1977), the speaker's "coherence goals" to structure the conversation in a way that will ease comprehension (Hobbs 1978), and "image goals", the speaker's desire to project a favorable image (cf. Goffman 1974, Labov and Fanshel 1977). But appropriate utterances do not exist as data independent of a competent observer's judgments as to which utterances are appropriate. We require competent judgments just to determine what could count as data.

Since the data is not exhaustively presented, the theory must make predictions to verify that it covers the data. But we need to be precise about what we can expect our theory to predict. We cannot expect predictions of the utterances, given only the context and the speaker's goals. A goal can be realized in many ways, and the mystery of human choice intervenes. The most we can hope for is to predict the set of possible utterances. But this set exists as data only in the form of its characteristic function, the appropriateness judgments of a competent observer. It is this that the theory should predict.

The best observer is someone with the greatest possible access to the context of utterance and the speaker's conversational goals. But this is just the speaker herself. In studying real conversation, we may assume utterances to be appropriate in context unless

there is strong evidence to the contrary. The fact that a competent speaker uttered a sentence and did not retract it is generally the best appropriateness judgment we have.

To predict an utterance, we would have to show why a derivation of the utterance from the conversational goals was chosen over derivations of all other possible utterances. To predict the appropriateness judgment, we need only show that some derivation exists. Thus in the microanalysis of section 5, we assume the utterances were appropriate, insofar as the speaker could tell, and seek to show that they could have been produced by a planning mechanism of the sort described in the next section.

3. The Planning Mechanism

Our working hypothesis is that an utterance is appropriate in a context if and only if it could be generated by an adequate planning mechanism operating in that context. Hence, our language of description is simply the language for describing the operations of the planning mechanism, a language of goals, plans, and beliefs. In order to begin our construction of a theory, we have to assume an initial version of such a planning mechanism.

We assume a planner of the sort familiar in robot planning (Sacerdoti 1977, Sussman 1975). It has goals and has access to general knowledge, including certain causal axioms encoding knowledge about what causes or enables what. Among these, for conversation, must be axioms that encode knowledge of conversational strategies, such as the facts that humor generally causes the listener to have a favorable image of the speaker and that descriptions of mishaps are frequently humorous.

The planner is able to decompose a goal into subgoals by using its knowledge of what actions will bring about the goal and what prerequisites enable these actions. It generates a conversational plan by decomposing its top-level goals into subgoals, and these subgoals into further subgoals, until the subgoals are directly realizable as utterances or gestures. When actions are executed, it uses its causal axioms to recompute the state of the world, by noting new conditions caused by the action and old conditions eliminated.

There has been much important work applying similar planners to the planning of speech acts (Searle 1969, Cohen 1978, Allen and Perrault 1978, Levy 1978, Moore 1978), and a theory of conversation will have to incorporate this work as a subpart. But it has not gone beyond the planning of single utterances. When we do, we can expect to encounter plans that require a

number of steps to execute and that can go awry at any point. To deal with these problems, we need to borrow and modify two further mechanisms from research in robot planning -- a monitor and a debugger.

In a natural language processor, a major component for interpreting utterances is an operation which finds a place for the utterance in an overall ongoing plan. Such an operation has been described elsewhere (A. Robinson 1978, Hobbs and J. Robinson 1978), and is similar to work on plan recognition (e.g. Schank and Abelson 1977, Wilensky 1979, Bruce 1978, Beaugrande 1979, Genesereth 1978). This operation serves to monitor the execution of a participant's plan. In the conversation analyzed, Y's reaction to X's moves (D5) and (D12) are interesting examples of monitoring.

The debugging situation in conversation is unlike debugging programs where we change the code and start again. It is more like a single run where we break at critical points, check certain conditions, patch up our errors, and go on. We have to live with the mistakes we've made. In our planning, we are using causal axioms that are at best only plausible, and sometimes actions don't cause what they are expected to cause. If the monitor has learned new information that contradicts what was expected, the debugger must attempt to determine which of the causal axioms happened not to be true, to account for this by searching deeper into the knowledge base for factors not previously considered, and to call on the planner to generate a repair and a new plan.

As the microanalysis of section 5 reveals, what has been sketched here is inadequate in at least two significant respects. These are discussed in section 6.

The planning metaphor provides an attractive vocabulary for describing conversation, for it seems to accord with the way we feel about our conscious moves and with what we are willing to attribute to our unconscious moves. The non-determinism of the planning process allows room for our sense of free choice. Unlike more rigid formalisms, e.g. flow charts, behavior outside the norm is not outside the system; rather it is a result of a less frequent path being taken by the planner. Unlike the rule systems proposed by ethnomethodologists (e.g. Sacks, Schegloff, and Jefferson 1974), the planning metaphor allows us to be explicit about the motives that lie behind the strategies we use. Among the various mechanistic metaphors of cognitive psychology, this one seems to detract the least from our humanity.

4. The Data to be Analyzed

For the rest of the paper, we will be analyzing a fragment of a free-flowing conversation, to exercise the planning mechanism. The fragment comes from the beginning of a videotaped conversation between a man X and a woman Y. The man enters the room first and sits down. Several minutes later the woman enters carrying a manuscript that happens to be her dissertation and four large manila envelopes. She sits down and they begin the conversation shown below.

- (D1) [Y displays dissertation.]
(D2) [Y displays four bulky envelopes.]
(D3) X: What's all this mail?
(D4) Y: My child is entering a Q-tips art contest.
You see, you haaa--
(D5) [X grins.]
(D6) Y: You don't have any children, obviously. You must...
(D7) You have to either draw or make things with the little Q-tips.
(D8) So she thinks she's going to win an \$8000 first prize.
(D9) So I have to send in this trash for her.
[One second pause.]
(D10) All these nice things made out of Q-tips.
(D11) And of course all the Q-tips will fall off.
[One second pause.]
and... in the mail....
(D12) X: And it's all to be sent to Blair Nebraska, huh?
(D13) Y: Yeah. This sounds really flakey though.
(D14) I... I never heard of Blair Nebraska
(D15) and you send it to a P.O. box.
So what happens too if I
(D16) What happens if you have dishonest mailmen
[X leans back in chair and crosses legs.]
and they see all these things going to an art contest, so they open it up and change it so that it's being sent from them?
[Y leaning forward. Three and a half second pause.]
(D17) X: How would they change it?
Y: Well... Instead of...
(D18) Instead of the return address being my address they would put down their address, so they would win, you see.
(D19) [Two and a half second pause. Y picks up envelopes, revealing dissertation for the first time since (D1).]

I am indebted to Will Leben and Dave Evans for making this videotape available to me.

- (D20) Y: Not that my poor child is going to win.
(D21) But anyway.
(D22) X: I don't think anybody, except for a child, would want to enter a Q-tips art contest.
(D23) [Both laugh. Y picks up dissertation and begins to leaf through it. Leans back. Shoulders relax.]
(D24) Y: Well, maybe the postman has children.
(D25) You never can tell.
(D26) This is my dissertation. It's just been approved.

Like all transcripts of everyday conversation, this appears incoherent at first. However, when we examine it closely in terms of what X and Y are trying to accomplish, we discover an intricate structure.

The purpose of the conversation was just to talk. It therefore provides an excellent minimal example of how conversation gets planned and of the structure that results, with the least intrusion from an accompanying task or event. The two people have met each other only briefly before, and this is their first lengthy conversation, so in our analysis we need not worry about shared knowledge that we lack access to.

5. Planning Topics of Conversation

Since most of the action in this conversation is Y's, we concentrate on her plans. We may assume she has two principal goals -- To project a favorable image and to cooperate with the experimental setup by maintaining the conversation.

The conversation can be divided into four episodes, each characterized by a different problem that faces Y. In the first, (D1)-(D2), Y attempts to introduce first her dissertation, then the mail, as a topic. In the second, (D3)-(D9), Y elaborates on the Q-tips. In (D10)-(D18), she tries to continue the conversation by fishing for a productive subtopic, finally hitting on the dishonest mailmen. In (D19)-(D26), due to the failure of this subtopic, she attempts to close the topic and again tries to introduce the dissertation.

Each of these episodes forms what we may call a "subtopic", a segment with a high degree of internal coherence. Each provides a different example of the speaker's ability to manipulate the topic of conversation.

5.1. Attempting to Introduce Topics

Y's initial problem is to introduce a topic that will cast her in a favorable light. She has the material for it: Her dissertation has

just been approved, and if they could talk about that, X would conclude she was at least intelligent enough to earn a Ph.D. degree.

Here's where the Catch 22 comes in, however. If X believes Y is intelligent, then X will think favorably of Y. But if X believes Y has uttered something with the intention of causing X to believe Y is intelligent, the utterance will be interpreted as boasting, and will make X think unfavorably of Y. Thus for Y to introduce a topic that will lead to a positive image too directly is dangerous.

However, if she can get X to introduce the dissertation, she will have achieved the goal of talking about it without the side effect of boasting about it. To get X to introduce it, she can display it prominently, and at last, we have arrived at an action which is executable. Y waves the dissertation about a bit. X does not pick up on it, and the plan fails.

Another way to convey a favorable image is to project the image of a good mother, and the good work of one's child is one way to do this. The problem as before is to introduce the topic, and the same hitch as before presents itself -- how to avoid boasting. The solution is the same as before. Y displays the envelopes, and the plan works as X asks, "What's all this mail?"

5.2. An Answer Perturbed

At first glance, Y's answer seems somewhat incoherent. But let's examine it more closely. To describe mail, one should describe its contents and destination, so an answer might be

- (1) CONTENTS: The envelopes contain Q-tips designs.
- (2) DESTINATION: I'm sending them in to a contest.

In fact, (2) appears, almost as is, in (D9). But (1) is a bit unusual; the Q-tips designs require some explanation. Y must tell of the situation that gave rise to them -- the Q-tips art contest. Since this is also unusual, she has to elaborate on the nature of the contest and might, among other things, specify what the contestant must make or do (the entry) and something about the prize structure:

- (3) My child is entering a Q-tips art contest.

ELABORATION:

- (4) ENTRY: You have to draw or make things with Q-tips.

- (5) PRIZE: There is an \$8000 first prize.

Y begins this orderly answer. She says (3) and then begins her elaboration (4). But she is interrupted, in a way that changes the rest of her answer significantly.

While just beginning (4) she looks up, the smile that X has been trying to suppress breaks into a grin, and they both laugh. His reaction to the notion of a Q-tips art contest is a negative evaluation of sorts. Y must therefore justify her involvement if she is going to maintain a favorable image. She does so by saying

- (D6) You don't have any children, obviously.

The implicit line of reasoning is -- if X had children, then X would understand Y's situation, and hence not evaluate negatively. So (D6) is an accusation of ignorance as a defense against the negative evaluation.

The next utterance (D7) is unaffected by the interruption, since it was entirely planned out before. There are several indications of this in her gestures. The rest of her answer, however, does seem affected in subtle ways.

The next utterance

- (D8) So she thinks she's going to win an \$8000 first prize,

is quite problematic. It does convey the information in (5). But (5) is not really an essential part of the background information for the answer to X's question, for it does not explain anything that is out of the ordinary.

One possible explanation for Y saying (D8) is that the daughter's high expectations provide a very strong motivation for Y to take the trouble to mail the entries. One does not like to shatter one's child's dreams. For this reason, (D8) functions as a further retort to X's negative evaluation.

The next utterance, "So I have to send in this trash for her," completes the answer. But it also defends against X's evaluation. We will examine how in section 6.1.

5.3. Searching for Something to Say

It is now X's turn to talk, but he doesn't, so Y must continue in a way that coheres with what has just been said. Her first attempt involves an inappropriate elaboration (D10), uttered in a forceless, off-hand manner. But it is also coherent to say "what happens next." Sending in the Q-tips designs provides the

occasion for them to fall off. Hence, (D11) continues coherently.

She has now tapped into a productive topic, so she thinks -- possible mishaps to the designs on their way to contest headquarters. At this point X interposes with the remark that it is all going to Blair Nebraska. It is likely that this is no more than a follow-up to his question about the nature of the mail. But Y, rather than deducing the place of this remark in X's conversational plan, such as it is, incorporates it into her own. She takes Blair Nebraska to be an example of sending the packages out into the unknown, and states the generalization of which Blair Nebraska is one example, namely, that the situation is flakey. Then she gives a further example, that the destination is a post office box. At this point, she pops up to the general topic of mishaps, of which the Q-tips falling off and the strange destination are two examples, and gives her third example, which she apparently believes will turn out to be a productive subtopic -- the dishonest mailmen.

Figure 1 illustrates this development.

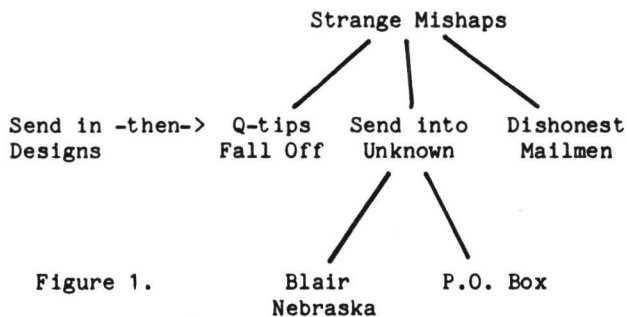


Figure 1.

Y confronts X in (D16), demanding a response with her direct "what if" question. There is a pause of 3 1/2 seconds. This is very long for a conversation like this, and it has a humorous effect on most viewers of the videotape. Finally, X does about the worst thing he could do with this topic -- he takes it seriously, and consequently dismisses it as a possibility. Y responds by trying to construct a serious means by which the disaster could happen, and finally it becomes apparent that the topic has failed on all counts -- it isn't productive of further conversation and isn't making her look good. She decides to cut bait and introduce a new topic, her original choice, the dissertation.

5.5. Escaping from a Failed Topic

She now faces the final topic-manipulation problem that we will examine -- how to escape

the current topic. When asked a question, she must answer if she is to cohere. If what is said to her reflects unfavorably on her, then she must retort. Finally, it is incoherent to suddenly switch topics -- or insofar as it is coherent, it is an admission of the failure of the previous topic.

Y is thus faced with three subgoals in pursuit of maintaining the conversation in a way that will make her look good -- she must salvage the current topic by arguing for the scenario's plausibility, close the current topic, and introduce the dissertation as a new topic. These three goals interweave in her next sequence of utterances and actions. She first displays the thesis for the first time since (D1), by removing the envelopes from on top of it. She has already defended plausibility in (D18), so she is free to close the topic. One way to do this is to deny the relevance of the topic to practical affairs, which she does with "Not that my poor child is going to win." At this point however, X won't let go. He responds to the whole idea with "I don't think anybody, except for a child, would want to enter a Q-tips art contest." This challenge puts Y back in the position of having to retort and then to close again before introducing the new topic. She retorts with "Maybe the postman has children", thereby denying the relevance of his argument, and then says "You never can tell," indicating that it is beyond their means at present to settle the question. She has thereby closed the topic again. In (D23) she has already picked up the dissertation and started to leaf through it, making the introduction of it as a topic less of a break with ongoing events. Then, in what represents a triumph of sorts, she says (D26), "This is my dissertation", thus succeeding in her original goal and closing the portion of the conversation that we torment with our microanalysis.

6. Computational Mechanisms

6.1. Multiple Acts in Single Utterances

In contrast to robot planning, where a single goal is realized by a sequence of actions, in conversation a single utterance frequently effects multiple goals. For example, utterance (D9) simultaneously answers a question, explains the motivation for an action, disavows the same action, and is humorous. We may approach this problem by viewing an utterance not as a single act but as a composite of several acts, each of which can realize separate goals. (D9) is a good example.

There seem to be at least three goals operating at this point. First, if Y is to cohere, she must complete her ANSWER of the question (D3) by conveying the information in (2). Moreover, the goal of defending against the negative evaluation remains, leading to two subgoals: She wants to show that her involvement results from some inexorable external circumstances (call this MOTIVATE), and to DISTANCE herself from the events by indicating that they are not a serious concern of hers.

Realizing all these goals in a single utterance requires something like the following mechanism: View a sentence as a conveyor belt with slots for each of its elements. The goals compete to fill these slots with lexical or syntactic material that will aid their own realization. Filling a slot is a matter of finding a match between what the sentence requires and what resources a goal has available. (A more pedestrian description would speak of looping through the slots, and for each slot, looping through the goals, and so on.)

Figure 2 illustrates this process:

DECISIONS:	UNMARKED:	ACTUAL:	SOURCE:
Conjunction	--	so	MOTIVATE
Subject	I	I	ANSWER
Aspect	Prog	have to	MOTIVATE
Verb	send	send	ANSWER
Destination	in	in	ANSWER
Object	these	this trash	DISTANCE
Beneficiary	--	for her	MOTIVATE

Figure 2

The goal ANSWER has determined the unmarked propositional content of the utterance at what Thompson (1977) has called the strategic level. At Thompson's tactical level, certain unmarked lexical choices are displaced with material provided by other goals. Thus, MOTIVATE supplies the conjunction "so" to indicate that Y's sending in the envelopes is due to her child's high expectations. MOTIVATE also replaces the present progressive tense with "have to"; Y's obligation excuses her for carrying around an armful of Q-tips. "For her" in the beneficiary slot indicates the circumstances behind the obligation.

Since those things one takes seriously, one necessarily values, one way to create DISTANCE is to evaluate the Q-tips designs negatively. Noun choice is a rich resource for such evaluations. The word "trash" means material with no value, and fits the bill perfectly.

6.2. Bidirectional Planning for a Next Utterance

Notice something about the segment from (D4) to (D16). First the daughter makes something to put into the envelopes (D4, D7), then Y sends them in (D9), they are in the mail (D11), they arrive in Blair Nebraska (D14), and are put into a P.O. box (D15) by mailmen (D16). What we have is the most mundane story imaginable of mail going to its destination. But Y has infused humor into this framework at every point, transforming dull raw material into the stuff of a good conversation. This is very suggestive about how utterances get planned.

A bidirectional search for a plan works not only from the goal, but also from whatever moves are currently possible. The tendency, once a schema is tapped, to follow the natural flow from one event to the next may constrain the possible next moves enough to make a bidirectional search feasible.

For example, we can imagine the initial stages of planning utterance (D11), "And of course all the Q-tips will fall off ... in the mail," going as follows: The next step in the mail scenario after the sending is that the envelopes are "in the mail" for a while. The design's fragility was a concern in packing, so associated with this step in the scenario is the possibility that the Q-tips might pop off. This is recognized as a mishap, and possible mishaps are a source of humor. Since the beginning, Y has had the goal of being humorous as a way of projecting a favorable image.

Once tapped into, the "mishaps" strategy gives a productive way of transforming further steps in the mundane scenario into conversational material. Bottom-up, we ask what happens next; top-down, we ask how this can be made interesting.

7. Some Questions for the Study of Conversation

Among the most useful ideas in our microanalysis was the notion of a subtopic as something that has structure and can be manipulated. Each subtopic served some overall purpose for the speaker and had some internal structure, and they flowed together naturally. This leads to three areas we might pinpoint for further investigation:

1. Can the notion of a subtopic be made precise, and does it represent a useful intermediate level between a person's top-level goals and what he ultimately says. In planning conversations, is there a useful distinction between deciding upon a subtopic and developing the subtopic?

2. What internal structures can a subtopic have? We have seen three examples in the microanalysis. The Answer (D3)-(D9) showed a structure of Background + Primary Information. The Strange Mishaps (D10)-(D18) proceeded through parallel examples. In the Escape (D19)-(D26), there was an intermixing of utterances aimed at three different goals. What other organizing principles are there for subtopics?

3. Adjacent subtopics flow together smoothly. We have seen two interesting transition devices that accomplish this: The link between the Answer and Strange Mishaps was the causal link between (D9) and (D11). In what was spoken, the move from the mail to the dissertation in (D26) was abrupt, but accompanying actions prepared the way. What other transition devices are there?

Starting with the hypothesis that conversation is planned behavior, in a sense that artificial intelligence has made precise, we have begun to see both what sort of planning it has to be and what kinds of plans result.

ACKNOWLEDGMENTS

Much of this work was done with Dave Evans. He is responsible for many of the insights and none of the errors. The work was supported by the National Science Foundation under Grants No. MCS-76-22004 and MCS-78-07121 and by the Advance Research Projects Agency under Contract No. N00039-79-C-0118.

REFERENCES

- Allen, J. & R. Perrault 1978. Participating in dialogues: Understanding via plan deduction. Proceedings, Second National Conference, Canadian Society for Computational Studies of Intelligence, Toronto.
- Beaugrande, R. de. 1980. The pragmatics of discourse planning. To appear in Journal of Pragmatics, Vol. 3. 1980.
- Bruce, B. and D. Newman. 1978. Interacting plans. Cognitive Science, Vol. 2, 195-233.
- Cohen, P. 1978. On knowing what to say: planning speech acts. Technical Report No. 118, Department of Computer Science, University of Toronto. January 1978.
- Genesereth, M. 1978. Automated consultation for complex computer systems. Ph.D. thesis, Harvard University. September 1978.

- Goffman, E. 1974. Frame analysis. New York: Harper.
- Grosz, B. 1977. The representation and use of focus in dialogue understanding. Stanford Research Institute Technical Note 151, Stanford Research Institute, Menlo Park, Cal. July 1977.
- Hobbs, J. 1979. Why is discourse coherent? To appear in F. Neubauer (Ed.) Coherence in natural language texts.
- Hobbs, J. & J. Robinson 1978. Why Ask? SRI Technical Note 169, SRI International, Menlo Park, California. October 1978.
- Labov, W. and D. Fanshel. 1977. Therapeutic discourse. New York: Academic Press.
- Levy, D. 1978. Communicative goals and strategies: Between discourse and syntax. To appear in T. Givon (Ed.), Syntax and semantics, Vol. 12. New York: Academic Press.
- Moore, J. 1978. Modelling communication as a problem solving activity. Unpublished manuscript.
- Robinson, A. 1978. Investigating the process of natural-language communication: A status report. SRI Technical Note 165. SRI International, Menlo Park, California. July 1978.
- Sacerdoti, E. 1977, A structure for plans and behavior. New York: Elsevier.
- Sacks, H., E. Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. Language, Vol. 50, No. 4, 696-735.
- Schank, R. and R. Abelson. 1977. Scripts, plans, goals, and understanding. Hillsdale N.J.: Laurence Erlbaum Associates.
- Searle, J. 1969. Speech acts: An essay in the philosophy of language. Cambridge, England: Cambridge University Press.
- Sussman, G. 1975, A computer model of skill acquisition. New York: Elsevier.
- Thompson, H. 1977. Strategy and tactics: A model for language production. Papers from the Thirteenth Regional Meeting, Chicago Linguistic Society, 651-668.