

## GENERATION AS A SOCIAL ACTION

Bertram C. Bruce  
BBN

On his first visit to kindergarten, while mother was still with him, Bruce, age five, looked over the paintings on the wall and asked loudly, "Who made these ugly pictures?"

Mother was embarrassed. She looked at her son disapprovingly, and hastened to tell him, "It's not nice to call the pictures ugly when they are so pretty."

The teacher, who understood the meaning of the question, smiled and said, "In here you don't have to paint pretty pictures. You can paint mean pictures if you feel like it." A big smile appeared on Bruce's face, for now he had the answer to his hidden question: "What happens to a boy who doesn't paint so well?"

-from Between Parent and Child

Haim Ginott, 1961

### I. INTRODUCTION

This paper is about the "why" and the "how" of natural language generation. Specifically, why does a person choose to communicate one idea rather than another (or none at all), and how does this choice get translated into a particular utterance? I want to present a few of the major issues and then suggest some ways of viewing the problem.

In the example above, the child's question is understandable in terms of his wants and fears and beliefs about the world. In order to explain why he asked the question he did, we must view his utterance as an action, rather than just a string of words. A string of words, per se, is not associated with any plan or goal. But an action is; in fact, the full representation of an action seems to require a representation of both its actual and its intended effects, its actual and its assumed preconditions.

This viewpoint further justifies the study of what Searle, Austin and others have called the "speech act." The production of a natural language utterance is best understood as an action which alters the state of the world, rather than as a mapping from a given meaning representation into a surface structure. Indeed, many of the problems of deep structure representation, such as focus and presupposition, are more profitably attacked in terms of action and plan structures.

Generation can then be seen as a two stage process. First, a plan is formulated which requires a communication with others. Second, the communication is made in terms of conventions which allow complex intentions to be expressed easily. These

stages are inextricably linked since the language conventions for expressing intentions make implicit references to plans of both the speaker and hearer. At the same time knowledge of the conventions can be used in plan formation.

### II. GENERATION AS A SOCIAL ACTION

Before discussing some of the conventions used to express intentions, it will be useful to consider the notion of a social action. A social action is one whose definition refers to beliefs, wants, fears, or intentions. There may be non-social actions. For example, eating can be described without reference to the diner's beliefs about the substance he puts in his mouth. But the description itself implies beliefs about the action which may or may not be shared. Another observer might say that the alleged diner is just pretending, or picking at the food, or wolfing it down.

A description of eating is a special case of INFORMING. It implies that the speaker believes that the diner is eating, and thus, that he believes that the diner intends to chew, swallow, and whatever else constitutes the physical act of eating. In other words, the act of describing encodes beliefs about plans. Speech is a social action because the definition of a speech act requires reference to the beliefs of both the speaker and the hearer. Implicit in each speech act is the goal of the speaker in making the utterance.

There are actions other than speech acts for which the notion of belief and intention are necessary. For example, HELPING is a social action in which the actor does something which furthers a plan inferred for someone else. I believe that generation can best be understood as a social action, i.e., the problem is first to understand how to represent and process any action defined in terms of beliefs and intentions, then to further specify that understanding for language generation.

### III. HOW TO REPRESENT A SOCIAL ACTION

The problem with a social action is that it cannot be represented by any predetermined, finite structure. Instead, its representation requires elements such as beliefs about plans, where plans are themselves defined in terms of beliefs and other plans. Rather than a structure itself, the definition of a social action is best thought of as a set of operations to be performed on a belief system. For example, a REQUEST requires the modification to a belief system such that (equivalently, is valid if) there is a plan of the speaker which has as either a goal or a subgoal a condition which the speaker believes would result from the action requested. Such an operation may require formulation of a plan or modification of existing plans.

The recognition and representation of plans is a complex and by no means fully

understood process. It requires a variety of types of knowledge to be applied. For instance, motivation rules must be used to determine whether a goal is appropriate for a person. Normative rules are used to account for behavior done under a sense of obligation. Wants are properties of persons which are used in inferring their goals in a given situation. These concepts and others are discussed in the references [1-3]. In the remainder of this paper I want to focus on one particular aspect of generation as a social action, namely, "How is intention encoded?"

How does a speaker indicate the purpose of his utterance to his listeners? In the example in the beginning of this paper, the child is only partially successful in making his intention known, and to the extent that he fails, he also fails to achieve his goals. In this case at least one of his goals seems to be reassurance that even if he doesn't paint well he won't be rejected.

Note that in recognizing his intention, Bruce's teacher also makes other inferences. He/she probably assumes that Bruce believes that he doesn't paint well; that he fears that he may be punished for painting "mean pictures;" and that he believes that the painter of the "ugly picture" stands in the same social relationship to the teacher as does Bruce. These inferences are both consequences and determinants of the perceived intention.

#### IV. HOW TO ENCODE INTENTION

Presuppositions. What can a speaker do to ensure that his purpose, and consequently, its associated inferences, are communicated to his listeners? One way is to establish, in the discourse previous to the utterance, the presuppositions for the purpose. For example, the purpose of REQUESTING INFORMATION has the presupposition that the speaker does not know the information. In the kindergarten example, Bruce is making a different REQUEST, in this case, for reassurance. There is a different set of presuppositions which needs to be established. We can assume that the teacher's familiarity with children entering kindergarten makes it easier for him/her to establish such presuppositions as that Bruce fears that the teacher may do bad things to him. If the listener fails to establish the presuppositions, as Bruce's mother does, then the communication fails. Bruce would need to emphasize his fears to his mother in order to have his utterance understood.

Linguistic Conventions. A second way that intentions are encoded is through use of linguistic conventions. For example, to indicate a REQUEST of any kind, the question form is typically used. A request often has a rising intonation, future tense or a modal, inverted word order, or a special word like "please." Many intentions, such as REQUEST, have a special associated verb, e.g., "I request that you..."

Discourse Structure. A third way to encode intentions is to take advantage of higher order linguistic conventions about discourse structure. There are places in a discourse where questions make sense, others where explanations are expected. Knowledge of typical discourse structures allows persons to condense and simplify utterances, avoiding the explicit establishment of presuppositions or explicit use of words like "promise."

While there is probably not a "discourse grammar" which would define "well-formedness of discourses," it is useful to have a model of how social actions typically fit together, and thus a model of discourse structure. Such a model can be viewed as a heuristic which suggests likely action sequences. By focusing the search involved in recognizing intentions it facilitates generation and subsequent understanding.

I have used the term "social action paradigm" (SAP) [1] for such a model of the flow of social actions. A SAP is a pattern of behavior (its body) with constraints (its header) on the applicability of the body. The header checks conditions on the situation in which the body is to be applied. At the same time, it binds variables in the SAP body to elements (people, times, locations, things) of the situation. A typical SAP body is shown in the attached figure. In the figure, REQUEST, SUGGEST, PROMISE, etc. are social actions; A and R are persons; and X is an action.  $F_1(X)$  is an alternative to X;  $F_2(X)$  is information which relates to the doing of X; and  $F_3(X)$  is a reason for not doing X.

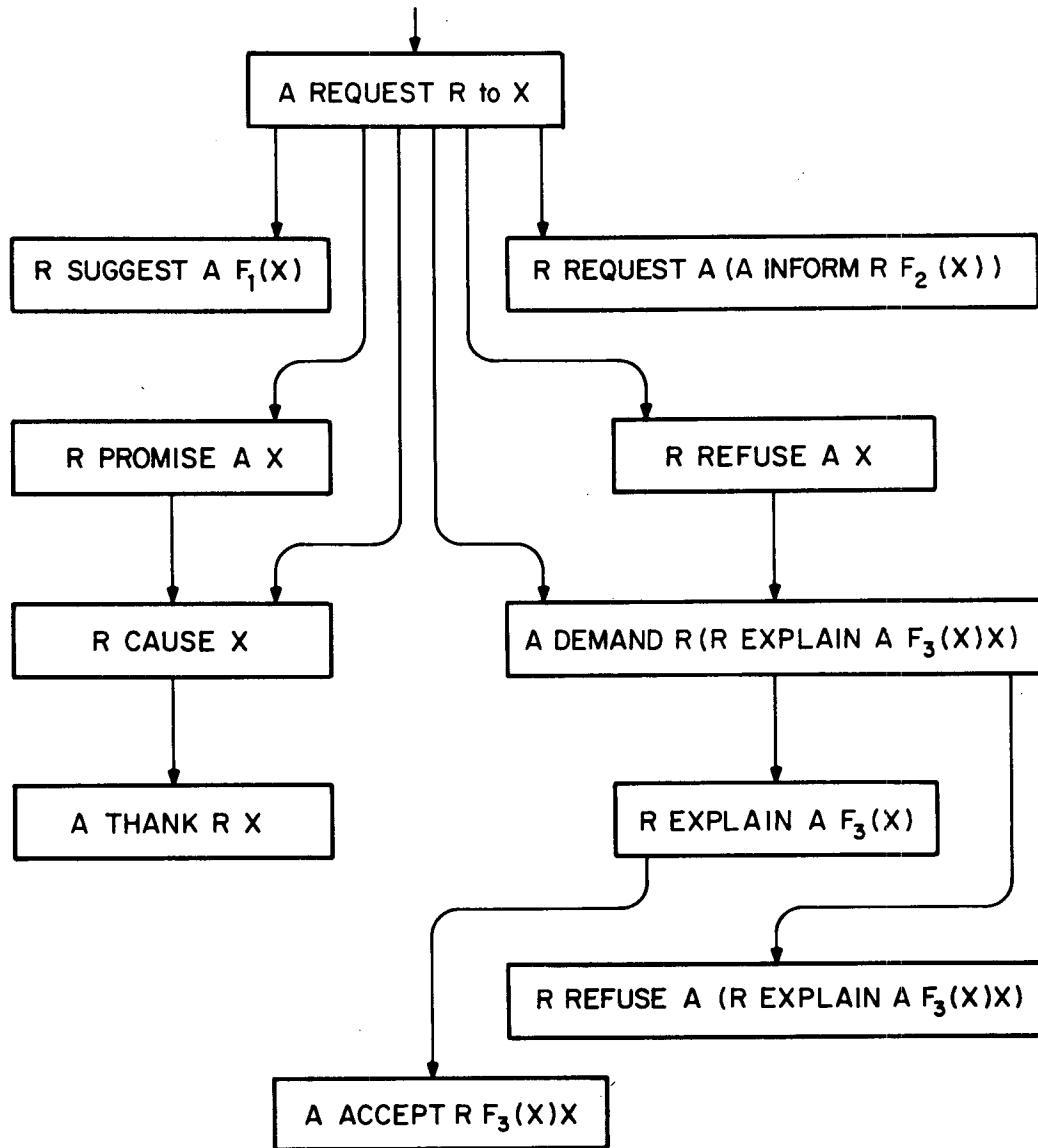
The SAP body says that A can REQUEST that R do X. Following the REQUEST, R may SUGGEST an alternative to X, may PROMISE to do X, may do X, may REFUSE to do X, or may REQUEST additional information. Following R's REFUSAL or inaction, A may DEMAND that R EXPLAIN, and so on.

Both speakers in a discourse can be expected to know various SAP's. For instance, knowing that a SUGGESTION often follows a REQUEST it is not necessary to encode the SUGGESTION explicitly. A person does not have to say, "I suggest instead that you..."

#### V. HOW TO FIT IT ALL TOGETHER

The principal point of this paper is that generation needs to be understood as an action in a social context. Let us examine such a context to see how a person's plan is carried out by encoding his intentions.

The context: Bill and Catherine are growing a vegetable garden. They have planted the seeds and have seen the first plants appear. The rhubarb is being attacked by small insects which have eaten holes in the leaves. Catherine notices the holes.



A Social Action Paradigm

Catherine's plan: Catherine's goal of having rhubarb to eat is threatened by the insects. In this case let's assume that she formulates a plan to poison the insects with Bill's assistance. She assumes that Bill has the same goals as she does with regard to the garden. Furthermore, let's assume that she doesn't know what poison to use but believes that Bill does, and that Bill doesn't know about the holes. Thus Catherine's plan involves INFORMING Bill about the holes so that he will be motivated either to put poison on the rhubarb or to tell her what poison to use. Another appropriate action for this plan is a REQUEST to Bill to do something about the holes (and the insects).

Encoding Catherine's INFORM/REQUEST: Catherine needs to do two things. One is to give information to Bill which she believes he does not have. This is called an INFORM. The other is to ask Bill to do something on the basis of his new knowledge. This is called a REQUEST. She can do these things with two utterances. However, if she believes that Bill doesn't want the holes, that he will infer that either she needs to know what poison to use or that he must apply the poison himself, and that he believes that she believes these things, then one utterance may be sufficient. Thus Catherine may just say,

"Bill, the rhubarb's got holes."

In that case she is relying on shared presuppositions about her utterance to carry the information about intention. On the other hand she could use explicit linguistic conventions as in,

"I inform you, Bill, that the rhubarb's got holes. I request that you either tell me which poison to use or apply poison to it yourself."

Bill's plan: While Bill has basically the same goals as Catherine let's assume he doesn't know much about plants, particularly rhubarb. When Catherine tells him that the rhubarb has holes he fails to make the inference that insects are eating the plant. Without that inference her last utterance might appear as an INFORM but not a REQUEST. Thus he has no reason to modify his plans about the garden. However he might well wonder why she said such a thing and formulate a plan to satisfy his curiosity. An action for his plan could be to REQUEST Catherine to explain her last utterance.

Encoding Bill's REQUEST: Following an INFORM a common action is a REQUEST to the first speaker to EXPLAIN his INFORM. This fact is expressed in the SAP's which include INFORM's. The general expectation of such a REQUEST coupled with a commonly used linguistic convention makes it possible for Bill to express his REQUEST succinctly:

"So?"

Catherine's plan: Realizing that Bill misses the point of her INFORM/REQUEST Catherine also realizes that her plan needs further action. She has to infer from Bill's REQUEST that he is not making the appropriate inferences himself and needs to be told directly that there are insects on the rhubarb which need to be poisoned.

Encoding Catherine's second INFORM/REQUEST: Catherine still believes that Bill will recognize her implicit REQUEST and that the problem with her first utterance was that facts were left out which Bill needed. Thus she says,

"It's covered with little bugs!"

## VI. CONCLUSION

The little dialogue introduced above could be continued in any of several directions. One likely continuation might be:

"I guess we oughta dust it then."

"I don't know what to use."

"How about the rose bush powder?"

"On rhubarb?"

"Sure, a bug's a bug."

"OK. But you do it. I don't know how much to use."

For each utterance in this sequence there is an associated plan and set of beliefs. At the same time there is heavy use of presuppositions, linguistic conventions, and SAP's to improve the speed and ease of communication.

These comments provide only a partial answer to the question of why Catherine says, "The rhubarb's got holes," or why Bruce says, "who made these ugly pictures?". They give but a sketch of how words are selected to encode intentions. I hope, though, that the comments have supported a consideration of generation as a social action occurring in the context of the speaker's and listener's intentions and beliefs.

## REFERENCES

- [1] Bruce, Bertram C., "Belief Systems and Language Understanding", Report No. 2973, Bolt Beranek and Newman Inc., Cambridge, Mass., January 1975.
- [2] Bruce, Bertram C., and C.F. Schmidt, "Episode Understanding and Belief Guided Parsing", Computer Science Department, Rutgers, 1974, NIH Report CBM-TR-32.
- [3] Schmidt, C.F., "Modeling of Belief Systems, Section 3". Second Annual Report of the Rutgers Special Research Resource on Computers in Biomedicine, 1973. Computer Science Department,