PATHFINDER: INVESTIGATING THE ACQUISITION OF COMMUNICATIVE CONVENTIONS

Robert Cummins
The University of Wisconsin--Milwaukee
Eric Dietrich
Martin Marietta Corporation

This work was supported in part by a grant from the National Science Foundation, and by the Institute of Cognitive Science, University of Colorado (Institute of Cognitive Science publication no.).

ABSTRACT

PATHFINDER is a system that solves coordination problems that require acquisition of a convention governing the intended meaning of a symbol. LEADER blazes a trail through a maze by leaving symbols in the various paths, and FOLLOWER must find LEADER by discovering the intended meanings of these blazes. PATHFINDER is the first step in a project to design a system that can solve a variety of coordination problems of the sort implicated in language acquisition. Solving certain coordination problems is communicating. Since coordination problem solution can become conventional (as David Lewis has shown), communication can become conventional, and that is language in its most general form. As conventions are acquired, more sophisitcated coordination problems can be solved, and more sophisticated conventions can be acquired. Eventually, it should be possible to acquire conventions governing identifiers and general terms, and this will enable use of a first order language via a recursive procedure adapted from Tarski by Cummins.

PATHFINDER: INVESTIGATING THE ACQUISITION OF COMMUNICATIVE CONVENTIONS

The PATHFINDER project is a study of the acquisition of the capacity to communicate by means of convention-governed symbols, and of the knowledge structures required for such communication. The project revolves around a series of PATHFINDER programs, each of which contains two programs--LEADER and FOLLOWER--which together solve coordination problems in a way that requires acquisition of conventions governing the meaning of a symbol. We begin by sketching the theoretical background, then turn to PATHFINDER itself.

In 1973, Jonathan Bennett (Bennett, 1973, 1976) outlined a two phase account of language acquisition based on the pioneering work of Grice on meaning (1957, 1969) and Lewis on conventions (1969). In phase one, he explains along Grician lines what we shall call pre-conventional communication: cases in which a speaker S performs some action and thereby communicates with an audience A in a way that doesn't depend on the prior existence of any shared rules or conventions. In phase two, he imports Lewis' account of conventions to show how pre-conventional cases could lead to the establishment of a convention between S and A with the result that S's act-type comes to have a conventional meaning. Since Bennett's work in this area has not received the attention it deserves outside of philosophy, (especially in AI and cognitive psychology) we begin with a brief review of his two-phase account.

Phase One: Pre-conventional Communication. Bennett takes from Grice the following conditional.

> (GC) If S utters E, intending thereby to get A to believe that p, and relies for the achievement of this upon the Grician Mechanism (GM), then S means by E that p.

Here is what we shall understand by the Grician Mechanism.

> (GM) A recognizes S's intention to get A to believe that p, and is led by that recognition, through trust in S, to believe that p.

This is a simplified version of Grice's more recent accounts, but we require only a rather crude sufficient condition at this stage of the account. Bennett claims that (GC) could be satisfied by pre-linguistic S and A, i.e., by S and A who share no conventional means of communication. We agree with this assessment for reasons that will emerge later. For now we shall simply assume that pre-linguistic S and A could satisfy (GC)--though perhaps only rarely and in rather special circumstances--and that (GC) does in fact formulate a sufficient condition for communication between S and A.

Phase Two: Conventionalization. The second phase of Bennett's account imports Lewis' treatment of conventions to show how a convention could emerge between S and A governing S's communicative actions. For present purposes, the crucial feature of Lewis's theory is this.

> (L) When a group achieves coordination in a certain situation by acting in a certain way, and they act that way because (i) they wish to achieve coordination, and (ii) each actor knows, and knows the others know, that that is how coordination has been achieved in the past, then the group has a convention governing that situation.

(L) applies to cases involving coordination of action, whereas our problem involves coordination between S's action and A's beliefs. But (L) is easily extended to accomodate this fact because the sorts of reasons A can have for adopting a belief so as to coordinate with S are the same sorts of reasons A will typically have for acting so as to coordinate with S. In particular, A can have as a reason for adopting the belief that S intends A to believe that p in uttering E the fact that A knows, and knows that S knows, that in the past S's intention in uttering E has been to get S to believe that p. If A is then led, through trust in S, to believe that p, we have a case that satisfies (GC) because S's utterance of E is governed by a convention existing between S and A. This yields the following account of conventional meaning.

> (CM) Utterance-type E conventionally means that p when uttered by S to audience A if (a) in the past, S has uttered tokens of E to A only when S meant that p, and (b) this fact is mutually known to S and A, and (c) because of this mutual knowledge it continues to be the case that when S utters tokens of E, S means, and is understood by A to mean, that p.

We can put the pre-conventional case and the conventional case together in an obvious way. Suppose S intends to get A to believe that a coconut is about to fall on A, and S goes through a certain performance that results in A recognizing S's intention and, via trust in S, adopting the belief that A is about to be hit by a coconut. Here we have a pre-conventional case in which communication occurs only because conditions are especially propitious, and because S's performance has a certain natural suggestiveness. Next time, however, the mechanism of convention will set in, and, as repetitions occur, the special conditions favoring the original success will no longer be necessary. S's performance can be streamlined by a process akin to stimulus substitution to the point where it need have no special features beyond the fact that A and S perceive it to be of the same type as its predecessors. Thus, the account allows for the fact that a sign may, so far as its physical characteristics go, have any meaning whatever.

Extending the Account. As it stands, the account just sketched hasn't a chance of being a full-scale theory of communicative conventions, for it begins and ends with sentence meanings--meanings have the form "that p" where p is a proposition. Since there cannot be infinitely many meaning conventions, it follows that the account just rehearsed runs afoul of the fact that a natural language contains infinitely many non-compound sentences having distinct meanings.

This defect has been repaired in Cummins (1978), by introducing Grician meanings for identifiers and general terms. Here are the relevant conditions.

> (ST) There is a convention whereby N refers to x in S's language if (a) in the past S has uttered N only when intending to identify x, and (b) this fact is mutually known by S and S's audience, and (c), because of this mutual knowledge it continues to happen that when S utters N S identifies x.
>
> (P) There is a convention whereby G means yellow in S's language if (a) in the past S has uttered G only when he/she/it meant yellow, and (b) this fact is mutually known to S and S's audience, and (c), because of this mutual knowledge it continues to happen that when S utters G, S means, and is understood to mean, yellow.

We can now state a relation between these meanings and satisfaction conditions, and import the standard recursion on the latter, to generate conventional meanings (though not meaning conventions) for an infinity of non-compound sentences.

> (S) 'The i-th member of the sequence f is red' gives the satisfaction condition for a token consisting of the general term G applied to the i-th variable iff the (or a) conventional maning of G is 'red'.

(S) allows us to go back and forth between satisfaction conditions and conventional meanings. If we start with cases for which conventions exist for the primitive general terms, we get satisfaction conditions for those terms by moving from the meaning to the satisfaction part. We can then use the standard recursion to get a satisfaction condition for any first order combination of the primitive general terms. Then, moving from the satisfaction part of (S) to the meaning part, we get conventional meanings, though not meaning conventions, for complex general terms. It is well-known that this sufficies to fix the truth-conditions for each sentence in a first-order language.

Investigating Convention Acquisition. The acquisition and use of communicative conventions has not been very extensively investigated by researchers in artificial intelligence or cognitive psychology, presumably because the requisite theoretical background has seemed lacking. However, putting Grice's account of communication together with Lewis' account of conventions yields a powerful theory of the acquisition of communicative conventions. Extending the account to apply to acquisition of conventions governing identifiers and general terms makes it possible to use the recursive apparatus of Tarski's theory of truth definitions to generate meaning conventions for every sentence in a first order language having a finite number of semantically primitive terms. The upshot is a theory of language acquisition for first order languages. This theory, however, is incomplete or vague at several critical points. (1) The theory tells us what it is to be a party to a communicative convention governing a symbol with a propositional meaning, but it does not tell us how humans can or do actually solve primitive communicative convention acquisition problems. (2) The theory tells us what it is to be a party to a communicative convention governing an identifier or general term, but it does not tell us how humans can or do acquire such conventions on the basis of simpler shared communicative conventions, viz., conventions governing symbols with propositional meanings.

We propose to meet point (1) by adding the hypothesis (i) that primitive communication problems can be solved, and appropriate conventions acquired, in the course of solving simple coordination problems that contain the communicative problems as sub-problems. The problem analyzed by PATHFINDER is just such a containing problem. We propose to meet point (2) by adding two hypotheses: (ii) that the power of a group of agents to solve coordination problems increases as that group acquires communicative conventions; (iii) that solving relatively more complex containing coordination problems enables agents to acquire relatively more sophisticated communicative conventions. It is these three hypotheses that the PATHFINDER PROJECT is primarily designed to investigate.

PATHFINDER: Embedding Communication Problems in other Coordination Problems. Pre-linguistic communication problems are difficult to solve in part because propositional attitudes are hidden. It is difficult for a speaker-audience pair to determine whether or not they have succeeded. This difficulty can be overcome by embedding primitive communication problems in other non-communicative coordination problems that are more tractable. If S and A are engaged in some cooperative activity, the success or failure of their efforts to communicate will be more or less obviously reflected in the success or failure of that activity.

In PATHFINDER, LEADER and FOLLOWER must solve such an embedded coordination problem. LEADER blazes a trail through a maze by leaving symbols in the various paths, and FOLLOWER must find LEADER by discovering the intended meanings of these blazes: LEADER must enable FOLLOWER to find LEADER. In the process, they must solve a primitive communication problem. For example, in the level-one version of PATHFINDER, FOLLOWER may learn that when LEADER marked a path "Y", LEADER meant that that path is to be avoided. Suppose FOLLOWER locates LEADER by avoiding paths marked "Y". Then LEADER and FOLLOWER will have solved their main coordination problem, and they will have solved a primitive communication

problem as well. Most importently, however, they will have solved a primitive convention acquisition problem: both know that "Y" means "avoid this path". This convention can be used in the solution of other related coordination problems, thereby increasing the power of LEADER and FOLLOWER to solve such problems, and hence increasing their power to acquire other conventions. For example, it is evidently easier for FOLLOWER to grasp an identifer in the context of an already understood instruction. "Avoid Y at zz," links use of the identifier to solving the embedding coordination problem (find LEADER), thereby making it possible for LEADER and FOLLOWER to recognize successful communication, and hence to acquire a convention governing use of the identifier. Conventions are a special kind of knowledge that increase capacity to solve coordination problems far more effectively than other types of shared knowledge. Advanced LEADER-FOLLOWER pairs will come to share conventions governing such things as the identifiers, general terms, and syntactic rules of a relatively sophisticated language.

Preliminary research has suggested a list of parameters of two types, intrinsic and contextual, the values of which define a relative level of sophistication. The coordination problems analyzed by PATHFINDER are significantly different from each other depending on the type of maze FOLLOWER faces (intrinsic parameters) and the amount and type of knowledge, including conventions, shared by LEADER and FOLLOWER (contextual parameters). This is especially significant given the hpypothesis that the capacity of two parties (LEADER and FOLLOWER, SPEAKER and AUDIENCE) to solve coordination problems should increase as simple problems are solved and conventions are acquired for future use.

Intrinsic Parameters. FOLLOWER will eventually have to face mazes that vary in at least the following ways: (i) number of branches per node; (ii) number of symbols per branch (including blanks); (iii) complexity of symbols--e.g., context sensitivity and reference to other parts of the maze; (iv) noise--e.g., symbol-like objects in the maze not left by leader.

Contextual parameters. To solve the coordination problem set by a relatively general maze, LEADER and FOLLOWER will have to share some knowledge. The amount and type of shared knowledge are contextual parameters of the coordination problem, for they specify the cognitive context in which the coordination problem is attacked. These include: (i) previously acquired conventions, if any, (ii) mutual knowledge of capacities--e.g., can LEADER cut down a tree, and does LEADER know FOLLOWER knows this? (iii) mutual knowledge of what is likely to be a natural rather than an artefactual feature--e.g., that pine cones are noise in a forest, but possible blazes in a building; (iv) mutual antecedent knowledge of the terrritory; (v) mutual knowledge of behavioral and cognitive tendencies. These parameters are best thought of as "passed" to LEADER and FOLLOWER from containing systems that specify the goals (blaze trail; find LEADER), contain records of mutual knowledge, and handle general reasoning and decision making, including when to give up, or to give up trying hard and just "try something" (a common strategy in communication).

The level-one version of PATHFINDER (which has already been implemented), involves a maze in which all branching is binary, there is at most one symbol per branch, and noise is limited by the assumption that only the symbols encountered at the first node are significant. In a level-one maze, FOLLOWER faces a relatively simple but non-trivial task. A maze that is general along all four dimensions specified above will evidently require a highly "experienced" LEADER-FOLLOWER team, a team that, we suspect, will have to share several powerful conventions to be effective.

Summary. The PATHFINDER project is designed to investigate the following strategy for language acquisition. S and A, given some shared knowledge and goals, but no shared conventional means of communication, solve a coordination problem such as that faced by LEADER and FOLLOWER. Several successes produce a shared convention. Now that S and A share a convention, they can solve more diffecult coordination problems, hence acquire more sophisticated conventions. Eventually, S and A will be able to acquire conventions governing identifiers and general terms, and hence, by a recursive process, a first order language. Since solving certain coordination problems is communicating, and coordination problem solution can become conventional, communication can become conventional, and that is language. Standard approaches to the problem of symbolic communication have emphasized acquisition of knowledge of a language. Yet it seems clear that learning a language is neither necessary nor sufficient for communication. Knowledge of a language is a means to understanding a speaker, or communicating with an audience. Language use and understanding is not likely to be properly understood if it is studied independently of the cognitive task that motivates it. The present project, in emphasizing the acquisition of communicative conventions, focuses on the cognitive task which language learning subserves and thereby avoids studying language acquisition "out of context".

REFERENCES

Bennett, Jonathan (1973). "The Meaning-Nominalist Strategy." Foundations of Language 10, 141-168.

Bennett, Jonathan (1976). Linguistic Behavior. Cambridge: Cambridge University Press.

Cummins, Robert (1978). "Intention, Meaning and Truth-Conditions." Philosophical Studies 35, 345-360.

Grice, H. P. (1957). "Meaning." The Philosophical Review 66, 377-388.

Grice, H. P. (1969). "Utterer's Meaning and Intentions." The Philosophical Review 78, 147-177.

Lewis, David (1969). Convention. Cambridge: Cambridge University Press.