

Book Reviews

Intentions in Communication

Robert P. Goldman

Intentions in Communication, Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, eds., System Development Foundation Benchmark Series, MIT Press, Cambridge, Mass., 1990, 508 pages, \$29.95, ISBN 0-262-03150-7.

Intentions in Communication is the outgrowth of an interdisciplinary workshop on the role of intention in theories of communication. Attending the workshop were researchers in computer science, linguistics, philosophy, and psychology. The resulting book contains edited versions of 14 papers (13 of which were presented at the workshop), commentaries, and an introduction. The topics of these papers range from philosophical analyses of the concept of intention to algorithms for recognizing plans, from logical formalizations of speech acts to analyses of intonational contours in discourse.

The idea of relating intentions to the use of language is an outgrowth of speech act theory. The foundation of this theory is largely owed to the philosopher J. L. Austin (1962), with substantial contributions from John Searle (1969) and Paul Grice (1975). Until J. L. Austin, the philosophy of language was largely concerned with the meaning of utterances, where the meaning of an utterance was taken to be what states of the world would make this utterance true or false. Austin's central (and simple) insight was that many common utterances are not simply true or false representations of the state of the world but are actions that are done to change the state. For example, there is no sense in which the sentence "I promise to review *Intentions in Communication*" is either true or false. The sentence doesn't describe the world; it alters it by causing a promise to exist.

Concern for speech act issues has come to AI natural language process-

ing primarily through the work of Philip Cohen, C. Raymond Perrault, and James Allen (Cohen and Perrault 1979; Allen and Perrault 1980; Perrault and Allen 1980). In their work, plan recognition was applied to the processing of dialogues (particularly those between a customer and a ticket clerk at a railway station). Programs were developed that reasoned from the explicit utterances of the customer and from facts about the situation to the goals of the customer and then to the formulation of an appropriate response. This reasoning could lead to better responses, as in the following case:

Patron: When does the Montreal train leave?

Clerk: 3:15 at gate 7.

Here, the clerk adds the information about the gate number to the response explicitly requested (Allen and Perrault 1980).

This review is organized around three of the themes that are sounded in *Intentions in Communication*: (1) foundational work on intention and its relation to speech act theory, (2) the problems of group intentions, and (3) intentions and plan recognition. The last theme is the topic closest to my heart.

My biases and the constraints of space dictate some omissions in this review. One theme sounded by the collection that is skirted is the relationship of truth-conditional semantics to speech act theories that rely on notions of intention. Two other papers that have to go without critical comment are Janet Pierrehumbert and Julia Hirschberg's paper on intonation and Herbert Clark and Deanna Wilkes-Gibbs's paper on reference. I am not qualified to evaluate either of these papers.

Intention and Speech Act Theory

Michael Bratman's paper, "What Is Intention?" is the perfect way to start off the volume. Bratman does an excellent job of introducing the reader

to the philosophical problems of the concept of intention. Primary among these problems are the side-effects of actions freely chosen. Do we believe that these side-effects are also equally intended? Consider the case of someone who does something abhorrent to end a war as opposed to someone who does something abhorrent to cause pain. The essence of Bratman's proposal is that one can distinguish between intended and unintended actions by asking whether the actor would still want the action to occur if there was some other way to meet the primary goal. To return to our example, the difference between the two persons (Bratman calls them "strategic bomber" and "terror bomber") is that the former would not commit the abhorrent act if there were some other way to achieve the goal of winning the war.

Bratman's definition of intention is the jumping-off point for Cohen and Levesque's two papers: "Persistence, Intention, and Commitment" and "Rational Interaction as the Basis of Communication." In the first of these papers, Cohen and Levesque attempt to develop Bratman's concepts into a formal theory of rational action. In isolation, this paper is difficult and not particularly interesting because it requires the reader to plow through a large amount of dense notation without a clear sense of his/her objective. The paper acquires its interest in context because it shows a formalization of Bratman's philosophical ideas, which is then used to give an account of the problem of speech act planning.

In their second paper, Cohen and Levesque propose an intriguing new analysis of speech acts. This treatment is unusual in that it handles speech acts no differently than other intentional acts. In a conventional speech act theory, if one asks a question like "Can you pass the salt?" with the intent of getting the hearer to pass the salt, two acts are said to be performed: A question act is indirectly used as a way of making a

request. The theory labors to explain under what circumstances a question is used as a request. According to Cohen and Levesque, however, this utterance is both a question and a request rather than a question acting as a request. Furthermore, it is to be so recognized based on a general theory of rational action, from which the specific theory of speech act interpretation falls out as a special case. This paper is interesting, but it would be nice to know if this new account would have any ramifications for natural language processing. For example, it would be nice to have some discussion of the way Cohen and Levesque would change Perrault and Allen's work on speech act interpretation based on their theory, if, in fact, they would change it at all.

It seems appropriate at this point to interject a quibble about the organization of the volume: Splitting the two papers by Cohen and Levesque serves to make each one less comprehensible. Also, in separating the two papers, the editors placed Cohen and Levesque's speech act paper after Perrault's, which is defined in reaction to Cohen and Levesque's.

Perrault argues that Reiter's (1980) default logic gives a better treatment than the modal logics of knowledge previously used. The argument for default logic is based on the fact that the frame problem arises in speech act theory. Perrault asserts that conventional treatments of speech act theory are of the form "when a speaker carries out a speech act of a type x , the result is some change, y , in the hearer's set of beliefs." The problem, of course, is that other events could prevent the expected changes from occurring and that these events are too many and too unpredictable to be enumerated. Perrault proposes that we use default logic to specify that the consequences will occur unless prevented. The reviewer is left with two concerns about this project. First is the use of the default logic formalism as such in light of Hanks and McDermott's (1986) work on anomalous extensions. An *extension* is a set of propositions that results when one uses default rules to extend a set of propositions. For example, given the statement that "Tweety is a bird" and the default rule that "all birds fly," the set of statements "Tweety is a bird; Tweety flies" is an extension. In their famous Yale shooting problem, Hanks and McDermott showed that

even in simple domains, the presence of conflicting default rules can lead to multiple extensions in ways that are difficult to predict.

The second concern is about the status of the speech act theory itself. Most existing formalisms, including most of those in this book, are normative theories. They are intended to specify what the effects of a speech act should be rather than what actually occurs in all cases. However, what is the status of a normative theory that uses default logic? Isn't it akin to saying "when x occurs, y should occur if the conditions are right" instead of simply "when x occurs, y should occur"? Does the phrase "if the conditions are right" add anything that isn't covered by the "should"? Clearly, it would if the conditions were spelled out. However, the purpose of default logic is to evade the need for spelling out these conditions.



Group Intentions

Three of the papers in this volume are concerned with the

problems that arise in considering the intentions of groups of cooperating agents. From most to least abstract, these papers are by John Searle, Richmond Thomason, and Barbara Grosz and Candace Sidner. Searle's paper strikes the keynote, attempting to demonstrate the necessity for special group intentional states, or *we-intentions*. Grosz and Sidner's paper is much the same project but translated into the language of AI. They define structures of shared plans and apply these structures to the analysis of collaborative dialogues (the same domain as in Diane Litman and James Allen's paper).

Richmond Thomason's paper suggests a new way of looking at speech act interpretation. Earlier approaches to this problem have seen speech acts as attempts by the speaker to directly manipulate the hearer's beliefs, intentions, and so on. Thomason argues for the existence of a shared conversational record, mediating between the beliefs of participants in a dialogue. He argues that participants cooperate to manipulate this conversational record. Such an account predicts that discourse participants will carry out some acts solely to maintain the conversational record and keep the dia-

logue functioning. They will collaborate in this way even when their goals in the dialogue conflict (for example, when they are arguing).

Intention and Plan Recognition

Chapters 5, 6, and 17 are concerned with the theme of *plan recognition*, the problem of recognizing an agent's plans from observations of his/her actions. Martha Pollack's paper is a discussion of the relation between plans (as they are understood in AI) and intentions. Pollack argues that AI needs terminology that allows us to distinguish between plans as recipes, algorithms, or procedure descriptions and as objects of intentions. For example, the United States possessed (and no doubt possesses) plans as recipes to attack the Soviet Union with nuclear weapons. However, the United States does not plan to do so.

Pollack's argument is somewhat difficult to follow. Although her chatty, informal style is easy to read, it gets in the way of a clear statement of the research objectives. The primary motivation for this theory is to build a foundation that will allow us to build systems that can reason about the plans of other agents even when these plans are faulty. Previous plan-recognition systems have all relied heavily on the assumption that the system and the planner share a library of correct plans. However, this motivation is buried in Pollack's paper instead of brought to the reader's attention at the outset.

For me, Henry Kautz's plan-recognition paper is one of the high points of this volume. In work reported at the Fifth National Conference on Artificial Intelligence, Kautz and Allen (1986) developed a formal model for the problem of plan recognition (based on McCarthy's [1980] circumscription). It is no exaggeration to say that this paper has set the standard for all later work in this area. Unfortunately, to the best of my knowledge, Kautz's follow-up work, in which he developed a set of graph-covering algorithms as tractable approximations to the declarative model of plan recognition, has not been published except in his thesis (Kautz 1987). It is a great pleasure to see some of these results published here.

Last Words

Different readers will have different reactions to this book. It will be of great interest to researchers con-

cerned with the foundations of natural language processing, but AI practitioners building systems today will find it of little appeal. Because it assumes so much previous knowledge, the book will not be useful to the casual reader. One would be at a disadvantage without a reasonable familiarity with predicate calculus and modal logic, AI planning formalisms, and the work of Perrault and Allen on interpreting speech acts (for example, Allen and Perrault [1980]; Perrault and Allen [1980]). Accordingly, the reader of this review should be warned that my point of view is that of a researcher (specifically, an academic researcher) rather than a system builder; your mileage might vary.

No review of this book would be complete without some mention of the commentaries, critical pieces written by other workshop participants that follow groups of related papers. Each commentator did an excellent job. The inclusion of these well-considered short pieces helps focus the reader's attention on important features of the related papers, giving him/her a feeling of participation in a fascinating discussion. The editors deserve congratulations for their fine work in editing and arranging. *Intentions in Communication* is one of the best-edited collections I have had the privilege to read. This point is particularly laudable in light of the book's origin in a workshop, which often makes for slapdash publications.

In fact, my only serious argument with the book is that it might have been improved by the editors being even more of a presence. A more substantial introductory chapter that gave more background and, perhaps, even a glossary would have opened the book to more readers. This need is particularly true with the casual reader, who must grapple with a dialect that consists of the jargons of philosophy, AI, and linguistics together. An expanded introduction and the inclusion of a glossary might also have spared the reader from wading through four or five fragmentary introductions to the work of Austin and Searle.

I borrow the words of Abraham Lincoln to conclude: "People who like this sort of thing will find this is the sort of thing they like." Not for all readers, *Intentions in Communication* is essential for those interested in foundational issues of natural language processing. Certainly, no

research library should be without it.

References

- Allen, J. F., and Perrault, C. R. 1980. Analyzing Intention in Utterances. *Artificial Intelligence* 15(3): 143–178.
- Austin, J. L. 1962. *How to Do Things with Words*. London: Oxford University Press.
- Cohen, P. R., and Perrault, C. R. 1979. Elements of a Plan-Based Theory of Speech Acts. *Cognitive Science* 3:177–212.
- Grice, P. 1975. Logic and Conversation. In *Syntax and Semantics*, eds. P. Cole and J. Morgan, 41–58. New York: Academic.
- Hanks, S., and McDermott, D. 1986. Default Reasoning, Nonmonotonic Logics, and the Frame Problem. In Proceedings of the Fifth National Conference on Artificial Intelligence, 328–333. Menlo Park, Calif.: American Association for Artificial Intelligence.
- Kautz, H. 1987. A Formal Theory of Plan Recognition, Technical Report, TR 215, Dept. of Computer Science, Univ. of Rochester.
- Kautz, H., and Allen, J. 1986. Generalized Plan Recognition. In Proceedings of the Fifth National Conference on Artificial Intelligence, 32–38. Menlo Park, Calif.: American Association for Artificial Intelligence.
- McCarthy, J. 1980. Circumscription—A Form of Nonmonotonic Reasoning. *Artificial Intelligence* 13:27–39, 171–172.
- Perrault, C. R., and Allen, J. 1980. A Plan-Based Analysis of Indirect Speech Acts. *American Journal of Computational Linguistics* 6(3–4): 167–182.
- Reiter, R. 1980. A Logic for Default Reasoning. *Artificial Intelligence* 13:81–132.
- Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language*. London: Cambridge University Press.

Robert P. Goldman is an assistant professor of computer science at Tulane University. He received his Ph.D. in computer science from Brown University for work on the application of probability theory to natural language processing.

On Being a Machine

Gabriel Valiente

On Being a Machine, Volume 1: Formal Aspects of Artificial Intelligence, A. Narayanan, Ellis Horwood Limited, Chichester, England, 1988, 200 pages, hardback, US\$39.95, ISBN 0-85312-957-6.

A great debate concerning the possibility for machine intelligence began with the advent of computing.



Roughly stated, in one corner, we find AI researchers and practitioners developing computational models that exhibit an ever-increasing

degree of intelligence, and in the opposite corner, we find formal theoreticians, philosophers, and psychologists arguing about the fundamental capabilities and limitations of machines. This debate seems to have no end because although there are strong arguments that shake the conceptual foundations of AI, counterarguments are as strong, and none presents decisive, irrefutable evidence of the basic capabilities or limitations of machines.

Ajit Narayanan's book *On Being a Machine, Volume 1: Formal Aspects of Artificial Intelligence* sheds new light on these issues by providing a formal analysis of the main arguments and counterarguments of AI proponents and critics. The book gives an introduction to this subject from a formal basis that is suited for a wide audience, including computer scientists, AI researchers and practitioners, formal theoreticians, philosophers, and psychologists. A description of the main formal aspects of AI is also provided, and the possibility for machine intelligence is analyzed from this formal standpoint.

The essence of the author's perspective is that "AI, despite considerable advances in its techniques, tools and applications, has not developed significantly as far as its theoretical and philosophical aspects are concerned, because from the very beginning AI has been miscategorized by theoreticians, philosophers, and even AI researchers" (p. 9). In addition, the cause of the miscategorization is rooted in the acceptance that all formal limitations that apply to computer science and philosophy apply to AI as well.

The basis for analyzing these formal limitations is Turing's imitation game and the objections to the game that Turing himself formulated. Although other arguments are covered as well (Minsky's proof of unsolvability of the halting problem, Searle's Chinese room argument, and Lucas's version of the mathematical objection), many important arguments to the possibility for machine intelligence are not mentioned at all, among them computational complexity limits (that is, time and space