COGNITION, COMPUTING, AND COOPERATION

edited by

Scott P. Robertson

Wayne Zachary

John B. Black

COGNITION, COMPUTING, AND COOPERATION

edited by

Scott P. Robertson

Wayne Zachary

John B. Black

ABLEX PUBLISHING CORPORATION
NORWOOD, NEW JERSEY

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data

Cognition, computing, and cooperation / [edited by] Scott P. Robertson, Wayne Zachary, John B. Black.
p. cm.
Some papers were first presented at symposia during the 150th annual meeting of the American Association for the Advancement of Science in New York, 5/84, sponsored by the AAAS Psychology Section and various other sections of AAAS.
Bibliography: p.
Includes index.
ISBN 0-89381-536-X
1. Interactive computer systems—Congresses. 2. Human-computer interaction—Congresses. I. Robertson, Scott P. II. Zachary, Wayne. III. Black, John B. (John Benjamin), 1947– IV. American Association for the Advancement of Science. Section J—Psychology.
QA76.9.I58C62 1989 89-14895
004′.33—dc20 CIP

Ablex Publishing Corporation
355 Chestnut Street
Norwood, New Jersey 07648

Preface

The chapters in this volume represent work from a broad sample of disciplines including anthropology, computer science, information science, and psychology. In this sense we hope that the book is an exercise in *cognitive science*, yielding something more interesting than a collection of papers from within any one of these disciplines alone. Because many disciplines are involved, different methodologies are strikingly apparent. Unfortunately, cooperation is not taken for granted in interdisciplinary endeavors, but we have tried to craft the section introductions and concluding remarks so that common features of the papers are highlighted. It is a working hypothesis that important parts of a theory of cooperation will emerge from comparison of these different approaches.

This book grew out of a day-long symposium, under the general title *Cognition, Computing, and Interaction*, presented at the 150th annual meeting of the American Association for the Advancement of Science that took place in New York in May, 1984. Scott Robertson and John Black organized a half-day program on *Human-Computer Systems*, and Wayne Zachary organized a half-day program on *Models of Cooperation*. The original contributors and participants in that symposium, John Anderson, Richard Frankel, Arthur Graesser, John Goodson, Carl Hewitt, Jerry Hobbs, Robin Jeffries, Thomas Landauer, Victor Lesser, Robert Mack, Kelly Murray, Denis Newman, Mary Beth Rosson, Charles Schmidt, Lucy Suchman, and Bonnie Lynn Webber, provided us with an exciting, interdisciplinary array of research on the structure and behavior of multi-agent systems. Seven of the chapters are based on work that was presented at that symposium, now much elaborated and extended. Other chapters represent work that has been done since or that we have added to round out the contributions and present a broader view of cooperative systems.

We are grateful to the sponsors of the original AAAS meeting, specifically the AAAS Psychology section and AAAS Information, Computing, and Communication section as main sponsors; and the AAAS Mathematics section, AAAS Education section, and AAAS Office of Science and Technology Education as cosponsors. Arthur Herschman, head of the Meetings and Publications Center for AAAS at the time, was responsible for putting together our two, initially separate symposia.

We owe a tremendous debt of gratitude to the authors who stayed with us throughout the editing process. We also owe considerable thanks to the authors who joined us later. Their contributions together provide an exciting perspective on cognition, computation, and cooperation.

Table of Contents

1
Introduction to Cognition, Computation, and Cooperation

Wayne W. Zachary

CHI Systems Incorporated
Spring House, PA

Scott P. Robertson

Department of Psychology
Rutgers University
New Brunswick, NJ

INTRODUCTION

New computer technologies have, in rapid succession, opened new vistas for interactive computing, graphical interaction, and personal computing. A major consequence of these technological solutions is a new kind of technological problem—the need to engineer intelligent interactions between people and computers, sometimes even between computers and other computers. The simple fact that so many human users of computers express displeasure with their human–computer interfaces makes it clear that this is indeed a problem. Until the very recent past, intelligent interaction was studied only as a human phenomenon. It was observed, analyzed, and theorized about by behavioral scientists, but always from the perspective of explanation, not that of design. Now, computer interface designers are looking to the behavioral sciences for theories, models, and tools to understand intelligent interaction, just as the behavioral sciences are beginning to see human–computer interaction as an important context in which to collect data and test theories. The opportunity and need to design interaction thus seems to be leading both the design sciences and the human sciences in new directions. This volume is an attempt to sort out some of these new directions.

In this book we have brought together several studies that are concerned with intelligent interaction, regardless of whether this interaction occurs between humans, machines, or some combination of the two. This selection principle has resulted in papers that cut across some conventional distinctions. Some are studies of interaction among intelligent machines, originating from the work in distributed artificial intelligence. Other papers focus only on intelligent interaction among humans. These address concerns that have deep roots in the behavioral sciences but are important applications to interface design and engineering. And still other papers deal directly with human factors engineering of human–computer interaction, a hybrid field that is at the same time both behavioral science and engineering.

All the papers, however, do have three points in common. The first point has to do with *cognition*. Interaction has often been approached from a strictly behavioral perspective, particularly in the social sciences and some branches of psychology. The last 30 years of research into human problem solving, however, overwhelmingly suggests that a cognitive perspective which postulates unobservable internal information processing is needed to adequately model intelligent activity. This perspective seems particularly appropriate for intelligent interactions, which require the interacting agents to understand each others' goals and motives (or at least behave as if they do) and figure out how to integrate these goals with their own. Such a cognitive perspective focuses on the information processing that gives rise to the interactive behavior—the problem solving, the knowledge about the relationships of actions to goals and subgoals to superordinate goals, and the knowledge about the behavior of distributed systems. We understand that humans, as *cognitive* systems, exhibit such capabilities, making the study of human cognitive abilities a primary path to the understanding and successful designing of cooperative systems. One goal of this book is to present studies of human cognition in situations that involve cooperation, especially situations involving human–computer interaction. Another related goal is to present research on how models of human cooperation can be applied to designing cooperative computer systems.

The second point of commonality among the papers in the book is that they all have to do with *computing*. There are two ways in which computing is relevant to the studies of interaction presented here—as a model for interaction and as a context for interaction. Abstract computation has, beginning with Chomsky (1957), become the primary metaphor for studying human cognition (Pylyshyn, 1984, recounts this process elegantly). In this framework, theories and models of cognitive phenomena are expressed in abstract computational terms (i.e., Turing computable frameworks), and often simulated via computer programs. Journals, research communities, and whole professional societies have formed around this computational metaphor of cognition. Although the overwhelming majority of these computation and cognition studies have considered individual behavior, we believe that this computational metaphor may be equally valid in the context of cooperation. Several papers included here frame models of interaction in explicitly computational terms.

Just as abstract computation can provide a language for modeling interaction, instrumental use of computing forms a concrete and crucial example of interactive systems. Computer systems are coming to be used more like consultants or even collaborators, and less like tools. In these contexts, people come to expect interaction with computer systems to have the same properties as interaction with other human beings. At a minimum, users expect computers to make available certain options at certain times and suppress others, just as a person would. In order to do this, a system must embody a task description of some kind and be able to employ it when making decisions. In more sophisticated systems, the computer is expected to take on parts of the task. An air traffic control computer, for example, should do more than monitor the objects within an airspace; it should also warn controllers about potentially dangerous situations. This requires information, not only about the task, but also knowledge about the consequences of actions in the task domain. Several contributions here deal with the design of cooperative computer systems. These papers argue that *cooperative* computer systems require an understanding of the problem-solving strategies employed by the person in the application domain. Additionally, these papers suggest, a computer system that is cooperative must make ongoing decisions about the plans of its user(s), other agents, or other systems that affect the domain in which it is acting.

Perhaps the most important contribution that such studies can make is an empirical one. When building an intelligent interactive system, there is ultimately feedback on the efficacy of the design—the system either works or it doesn't. Actually, the results are usually somewhere in between the two. The program exhibits some expected behaviors, fails to exhibit others, and may even produce some that are totally unexpected. Because these successes and failures can be traced back to elements of the original design, the relationship between theory and behavior can be examined empirically. This is never possible in purely human studies because the original design for human intelligence seems to have been lost somewhere along the way.

Finally, all of the papers here are concerned in one way or another with the concept of *cooperation*. As we use it, *cooperation* is a special form of interaction in which the cooperating agents share goals and act in concert with them over an extended period of time. It differs from other computational studies of interaction, such as most of those in economics, or evolution (e.g., Axelrod, 1981), which examine competitive rather than cooperative behavior. Every chapter in this book deals in one way or another with cooperation situations and cooperative interaction. Each contribution provides a different view of cooperation and a different requirement of participants in cooperative situations. These views are based on the kind of cooperative system studied, as well as on the disciplinary vantage point of the authors. Our primary goal in this volume is to find a common thread or concept that underlies cooperative behavior in the varied contexts in which it is examined—in human interaction, through human–computer interfaces, or in the design and analysis of abstract cooperative systems. In conclusion, we attempt to draw out this common thread as a preliminary theory of cooperation and cooperative systems.

BACKGROUND TO THE STUDY OF COOPERATIVE SYSTEMS

The chapters in this volume represent scholars from anthropology, computer science, information science, linguistics, human factors, and psychology. Because divergent disciplines are involved, different methodologies, assumptions, and forms of data are apparent. We have deliberately chosen to not offer any discussion of the history of research in this area, or any detailed review of the literature, as is normally the case in the introduction to an edited volume. Cross-disciplinary research is something that is always discussed and encouraged (and to that extent is nothing new), yet only rarely makes a lasting impact on major, established fields of study. Bibliometric research on cross-disciplinary citation (Chubin, Rossinni, Porter, & Connelly, 1986; Neeley, 1981; Rigney & Barnes, 1980; Small & Crane; 1979) has shown that mainstream fields are associated with well-defined research journals, and that publications in those primary journals show low rates of cross-disciplinary citation that do not vary much through time. While citation is not the only (or the best) measure of the *interdisciplinary-ness* of a given piece of research, it is nonetheless a reasonable one, and one which leads to the conclusion that most of the time, most fields of study tend to look inward toward resolving well-defined problems that have (local) theoretical significance. Even the study of cognitive science, with its recent origins as a multi-disciplinary field, may have already begun to grow introspective and focus on problems that are well-defined and important only in terms of cognitive science.

We therefore want to view this volume as a new endeavor. The study of cognition, computation, and cooperative interaction does not appear to be anything remotely resembling a well-defined field. Researchers from many different disciplines are getting involved, often with an original intent to study something else. This has resulted in some cases of the wheel being re-invented. It has also led to other cases where researchers are clearly unaware of important and relevant results in related fields. For example, Suchman's chapter argues that most cognitive scientists working on interaction are woefully unaware of the highly relevant social science work in ethnomethodology. The point is that we are not arguing that the convergence of interests contained in this book is any sort of historical inevitability or theory-driven breakthrough. Instead, we see it as an interesting accident that arose from the need to design intelligent interactions. We hope that this accident might produce results of interest to many.

A PRELIMINARY DEFINITION OF COOPERATION

What is meant by cooperation? In this book many types of cooperative situations will be discussed. Before plunging into the various contributions it is reasonable to ask just what constitutes cooperative behavior. One approach to the problem of cooperation is to compare example scenarios and discuss whether they should be categorized as demonstrating *cooperative* behavior. Here we will present several scenarios and follow the tradition from linguistics by marking an asterisk by those which we feel are not members of the class of cooperative behaviors. By comparing cooperative and non-cooperative acts, distinguishing features should begin to emerge.

The position that cooperation is a special form of goal-directed behavior among several agents will be set forth in more formal analyses of the examples. This position is shared by all of the contributers to this book. In fact, it could be argued that only by analyzing the interaction of multiple agents' goals can we understand cooperation—or any other coordinated action undertaken by several rational actors. There are many ways that the goals of multiple agents can interact, but when they are combined or distributed so that the agents work as a goal-directed system we have examples of *cooperative behavior.* Mechanisms for achieving and controlling this task are the topics of this book. First, though, some examples.

Cooperating With Shared Goals

First, consider several example scenarios in which two agents share a goal:

1. *Two children want a piece of pie. They fight over it and one gets it.
2. Two children want a piece of pie. They talk about it and cut it in half.
3. Two children want a piece of pie. They talk about it and one of them decides to skip the pie today if he or she can have some tomorrow.
4. *Two children want a piece of pie. They fight over it and it breaks in half. Each grabs a half and eats it.
5. *Two children want a piece of pie. Their mother brings each of them a piece.

In cases 1 and 3, one agent achieves the shared goal and the other agent fails. In cases 2 and 4, both agents achieve their mutual goal (at least partially), and in case 5 both agents fully realize their common goal. Cases 1, 4, and 5 are judged to be not cooperative while cases 2 and 3 are judged to be cooperative.

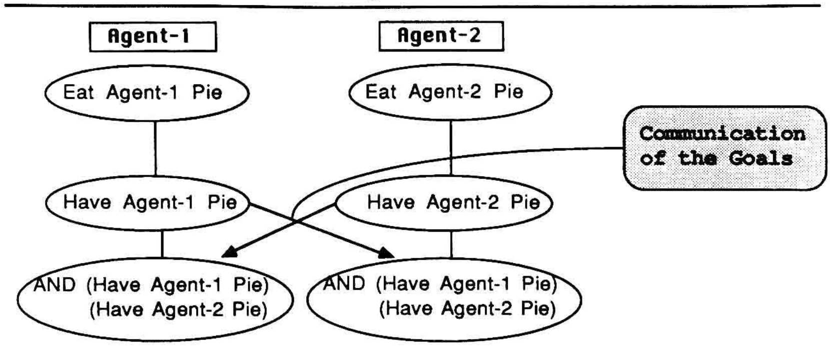
Examples 1–5 suggest that it is not necessary that agents achieve their goals to be considered cooperative, and it is not necessarily the case that agents have cooperated just because they achieve mutual goals. Case 5 reinforces the latter point by showing that agents can share goals and realize them completely, but still not be considered to have cooperated. The distinguishing feature of the two cooperative cases (2 and 3) is that the outcome is the result of coordinated behavior of the agents. In both cases, the original goal is renegotiated in such a way that the new goal formulation can be achieved. In case 2, for example, each agent decides to have less pie so that both can have some.

While the outcome in case 4 is the same as in case 2, it is not achieved in the same way and is judged to be noncooperative because the goal reformulation is not arrived at by mutual problem solving but is the result of a change in the state of the world. Similarly, in case 5, the mutual goal is arrived at by the action of another agent and is not accompanied by mutual problem solving. In case 5, in fact, the agents need not have been in communication at all for the scenario to have developed. Thus, it seems that the nature of problem solving, more than the outcome, is critical to cooperative behavior.

Figure 1 presents an analysis of the situation depicted in examples 2 and 3 in terms of the goal configurations of the agents. The goals of the two agents are shown in different columns. The superordinate goal of both agents has the form (Eat *agent* Pie). In both cases, we assume that the agents have generated a subgoal of the form (Have *agent* Pie) in order to satisfy a precondition for eating the pie. (This analysis is consistent with any number of planning or goal decomposition mechanisms in the problem-solving literature.)

At the second level of decomposition, where both agents have generated the subgoal to have pie, we cannot say that they have a mutual goal. In fact, if pie is a limited resource, then the agents have a goal conflict. All of the situations depicted in examples 1–5 are consistent with the goal analysis in Figure 1 up to the second level of decomposition. A precondition for the cooperative acts in examples 2 and 3 is that the agents formulate

Figure 1.



a truly mutual goal, a goal with the same form and content. In order to do this, each agent must incorporate the goal of the other. This is indicated at the third level of decomposition by a conjunctive goal which includes both agents' goals to have pie. Also indicated is the fact that in order to formulate a goal with another agent as the subject, communication about this goal must take place. In examples 2 and 3 this is achieved by conversation about the mutual goal.

The point of working toward the conjunctive goal in Figure 1 is that planning processes will be qualitatively different under the conjunctive goal than under its parent goal. For example, under the parent goal (Have *agent* Pie), grabbing the pie and running off is a satisfactory plan, but it is not satisfactory under the conjunctive goal. Also, because the conjunctive goal contains its parent, satisfying the conjunctive goal means satisfying the parent goal. Thus the conjunctive goal is really a reformulation of the parent goal which allows planning toward satisfaction of both the parent goal and another agent's goal.

Figures 2 and 3 show further goal decompositions of situations 2 and 3

Figure 2.

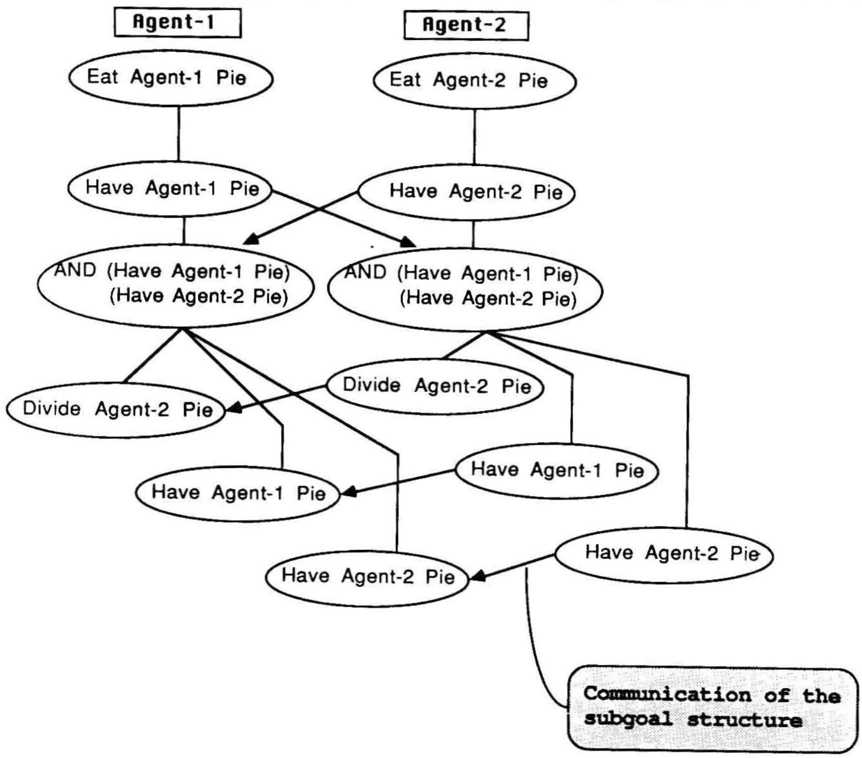
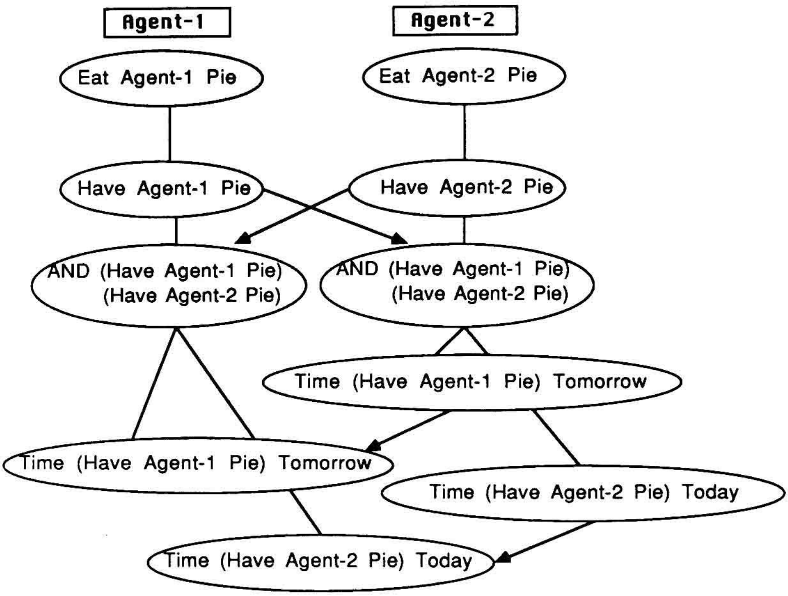


Figure 3.



respectively. In Figure 2, agent-2 decides that the pie object can be redefined such that having and eating half of a piece of pie will be sufficient to satisfy the goal. Thus a plan to divide and share the pie resolves the goal conflict. Note that this plan must be communicated to agent-1, allowing the other agent to reformulate the conjunctive goal in the same way so that when the plan is carried out, both agents will agree that it satisfies their respective goals. Figure 3 shows a similar reformulation of the conjunctive goal, but this time the plan is to spread the limited resource out in time.

The act of communication about goals is critical to cooperative behavior. This is true for two reasons. First, it is not possible to represent another agent's goals unless they are communicated in some way (we will take up a subtle variation of this point, inferring the goals of another, later). Second, once mutual goals are established, the planners must agree about what will satisfy their goals. One agent may generate a plan which will be rejected by the other, for example, if the agent in example 3 does not want to wait until tomorrow to eat pie. Cooperation requires that an agent know that another agent will be satisfied with a particular plan for achieving conjunctive goals.

Irrelevance of Goal Outcomes

To further make the point about the importance of communication and agreement on plans for the achievement of mutual goals, consider the following scenarios in which the outcomes differ:

6. Two children want a piece of pie. One suggests that they divide it in half. The other says that he or she wants a whole piece of pie.
7. Two children want a piece of pie. They talk about how to get one. Their mother brings each of them a piece.
8. Two children want a piece of pie. They pool their money and buy one. A dog snatches it and runs away.

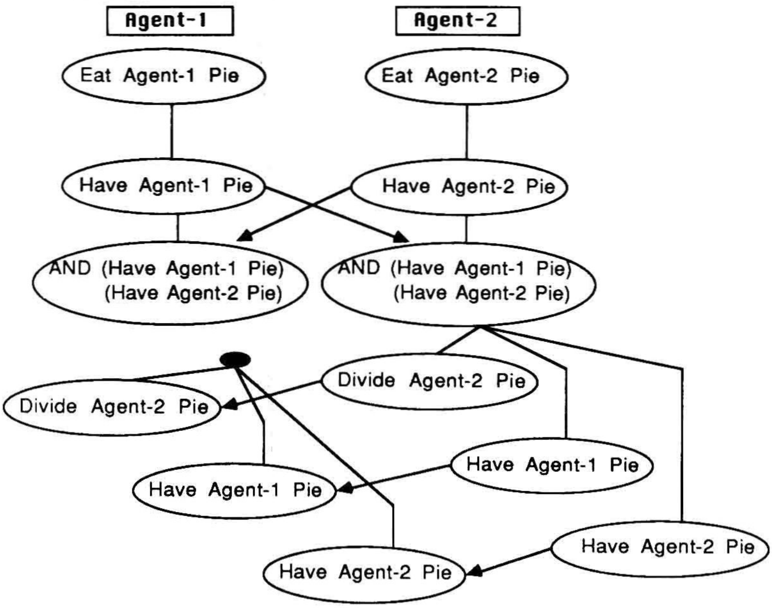
Cases 6, 7, and 8 all involve mutual goals and depict actions taken in common toward those goals. All of these examples are considered instances of cooperation.

Case 6 is the case discussed earlier in which one agent rejects the plan of another agent to achieve a mutual goal. Figure 4 presents a goal analysis of this example. In this case, communication of the plan from agent-2 to agent-1 does not result in adoption of the plan by agent-1 as it did in example 2 (compare Figure 2). Note that this does not preclude further planning toward the mutual goal and thus we consider it still an instance of cooperation.

There is a definitional ambiguity here that results from the nature of the example and that is potentially confusing. Some might say that agent-1's action in example 6 is *noncooperative* because he or she did not go along with the plan. This assessment is based on a judgment that agent-1 could be satisfied with the pie division plan but that he or she is just being difficult. In the analyses presented here, it is assumed that plans are accepted and rejected on sound principles. Thus the rejection is part of a process of communication about plans for achieving common goals and is part of cooperative problem solving.

In case 7, the goal is achieved in the same way that it was in case 5, by the intervention of an agent external to the cooperating group. Interestingly, we say that the agents were cooperating in case 7 and not in case 5 because, in case 7, they engaged in problem-solving activity toward a mutual goal. A similar point is made in case 8, where the agents were engaged in problem solving toward a mutual goal but an external agent intervened to thwart the goal. Thus, just as acceptance or rejection of a plan by other agents is irrelevant to the cooperative nature of problem solving, the outcome of the planning process and the actions that are taken are also irrelevant.

Figure 4.



Situations 1–8 begin to suggest that shared goals and communication about plans for achievement of shared goals are central features of cooperative acts. They stress the fact that the outcome of goal-directed behavior is irrelevant to its cooperative nature since we have noncooperative examples where goals are achieved and cooperative examples where goals are not achieved. The examples above also suggest strongly that cooperation has something to do with the nature of the problem-solving process as much as it has to do with the existence of mutual goals.

Cooperating Without Sharing Goals

Consider next some examples where goals are not shared:

9. *Two children notice a piece of pie on the table but they are so stuffed that they don't give it a second look.
10. *One child wants a piece of pie. She takes a piece and offers some to her friend who refuses.
11. One child wants a piece of pie. She can't reach it. She asks her taller friend who gets it and gives it to her.
12. One child wants a piece of pie. She can't reach it. Her taller friend notices and gets it for her.