The agreement process: an empirical investigation of human-human computer-mediated collaborative dialogues

Barbara Di Eugenio*
University of Illinois at Chicago

Pamela W. Jordan†
University of Pittsburgh

Richmond H. Thomason‡
University of Michigan

Johanna D. Moore§
University of Edinburgh

In this paper, we investigate the empirical correlates of the agreement process. Informally, the agreement process is the dialogue process by which collaborators achieve joint commitment on a joint action. We propose a specific instantiation of the agreement process, derived from our theoretical model, that integrates the IRMA framework for rational problem solving (Bratman, Israel, and Pollack 1988) with Clark's work (1992; 1996) on language as a collaborative activity; and from the characteristics of our task, a simple design problem (furnishing a two room apartment) in which knowledge is equally distributed among agents, and needs to be shared. The main contribution of our paper is an empirical study of some of the components of the agreement process. We first discuss why we believe the findings from our corpus of computer-mediated dialogues are applicable to human-human collaborative dialogues in general. We then present our theoretical model, and apply it to make predictions about the components of the agreement process. We focus on how information is exchanged in order to arrive at a proposal, and on what constitutes a proposal and its acceptance / rejection. Our corpus study makes use of features of both the dialogue and the domain reasoning situation, and led us to discover that the notion of commitment is more useful to model the agreement process than that of acceptance / rejection, as it more closely relates to the unfolding of negotiation.

1. Introduction

The last few years have seen greatly increased interest in collaboration and negotiation (Grosz 1996), no doubt partly because of the new opportunities for collaboration between both human and software agents offered by the Internet and the World Wide Web; see (Etzioni and Weld 1994; Maes 1994) inter alia. The problem of collaboration is being approached in a variety of ways: researchers are studying its philosophical, linguistic and psychological foundations (Bratman 1992; Clark 1992; Clark 1996), and are developing formal, computational models (Cohen and Levesque 1990; Grosz and Kraus 1996), including models of negotiation (Rosenschein and Zlotkin 1994).

* Electrical Engineering and Computer Science Department, Chicago, IL, USA, *dieugeni@eecs.uic.edu*
† Intelligent Systems Program, Pittsburgh, PA, 15260, USA, *jordan@isp.pitt.edu*
‡ AI Laboratory, Ann Arbor, MI 48109, USA, *rich@thomason.org*
§ Division of Informatics, Edinburgh EH8 9LW, UK, *jmoore@cogsci.ed.ac.uk*

As communication is deemed necessary for successful collaboration by at least some researchers (Cohen and Levesque 1991; Grosz 1996; Grosz and Kraus 1996), and as language is one of the main means humans use to communicate, collaborative dialogues have received much attention in the recent computational literature; see (Grosz and Sidner 1990; Lochbaum, Grosz, and Sidner 1990; Lochbaum 1994; Walker 1996; Rich and Sidner 1997; Chu-Carroll and Carberry 1998) inter alia. Interestingly, most of this work concerns collaborative problem solving dialogues. The fact that collaboration in language has been studied mostly in problem solving scenarios is not surprising, as collaboration often spontaneously arises when people *discover they have a problem best solved together* (Grosz 1996, p. 69). More importantly, problem solving scenarios provide a direct window into the mutual influence that rational problem solving behavior and communication exert on each other.

One of the most widely accepted models of rational problem solving behavior within the CL community is IRMA (Intelligent, Resource-Bounded Machine Architecture) (Bratman, Israel, and Pollack 1988; Pollack 1992)—see for example (Walker 1995; Webber 1999). Of course, the view of language as action presupposes viewing speaker and hearer as rational agents, and goes back to at least (Austin 1962; Searle 1965). However, the model developed by Bratman and his colleagues is especially appealing because it brings to the fore the issue of *resource-boundedness*, i.e., the fact that *[agents] are unable to perform arbitrarily large computations in constant time* (Bratman, Israel, and Pollack 1988, p. 349). IRMA accounts for both means-end reasoning and the need to weigh alternative options for action, and for the successful interaction of these two processes. What is missing in IRMA is an explicit link to collaboration, particularly in dialogue. Although perception is taken into account in IRMA, this architecture does not directly explain how negotiation unfolds in dialogue, how conversants come to agree on a solution, how they interpret and produce language, and the discourse strategies they use. Clark and his associates' work (Clark and Wilkes-Gibbs 1986; Clark and Schaefer 1987; Clark 1992; Clark 1996) provides a model of collaboration in dialogue that is an ideal candidate to bridge the gap, as it explains how the mutual belief needed for an agreement can be reached. We believe we should be able to model collaborative problem solving dialogues more effectively by integrating these two frameworks.

Informally, the *agreement process* is the dialogue process by which collaborators achieve joint commitment on a joint action. We model this process by integrating IRMA with the basic present/accept mechanism used to establish the mutual beliefs that constitute agreement. In this paper, we will propose a specific instantiation of the agreement process, attuned to the characteristics of our dialogues. Note that we are not advocating a conceptual departure from the collaborative cycles that have been proposed in the computational literature (Sidner 1992; Walker 1993; Sidner 1994; Chu-Carroll and Carberry 1998). Rather, we believe that these cycles, including ours, are all instantiations of the same abstract agreement process. The specifics of each cycle, including the one we present in this paper, can be seen as manifestations of the recursiveness of the agreement process, coupled with characteristics of resource-bounded practical reasoning and its manifestations in language that different researchers want to explore. We believe that the framework we propose can potentially explain every instantiation of the agreement process, as long as the framework is informed by the appropriate features of the corresponding task, such as the distribution of knowledge between agents. We will attempt to show this as far as our task is concerned, and we will speculate on how the framework could be applied to other collaborative cycles from the literature.

Given this is a general framework for analyzing collaboration in dialogue, the main contribution of our paper is an empirical study of some of its components, as specified in the particular instantiation of the agreement process we explore. The goal of our empirical study is to systematically identify features of utterances that allow the conversant (and ultimately a computer version of a conversant) to recognize the function each utterance performs within the agreement process. To gain insights into the agreement process, our empirical corpus study focused on how information is exchanged in order to arrive at a proposal, on what constitutes a proposal, and on its acceptance / rejection. (Or, more generally, on the acceptance / rejection of a *refashioned*, i.e. modified, proposal (Clark and Wilkes-Gibbs 1986)). Clearly, a corpus study would not really be necessary if our dialogues included only explicit proposals, acceptances and rejections / counterproposals. However, as the theory of speech acts has pointed out (Austin 1962; Searle 1965; Searle 1975; Brown 1980), surface form is not a clear indicator of the speaker's intention. We exploit dialogue history and the effect of the *task*, i.e., of the domain reasoning situation, on context, to reach the appropriate interpretation.

The theoretical contributions of this study are as follows. First, we distinguish *proposals* proper from *partner decidable options*; these two constructs correspond to an instantiation of the deliberation process in IRMA derived from the specific characteristics of our task (a design problem in which knowledge is equally distributed, and needs to be shared). *Proposals* correspond to situations where an agent can deliberate to the point of making a commitment, and *partner decidable options* to situations in which the agent's knowledge does not permit her to do so, at least if she is cooperative. Moreover, in light of our corpus study we conclude that the notion of *commitment* (Cohen and Levesque 1990; Bratman 1992) is more useful than that of *acceptance / rejection* in order to model the agreement process. By tracing how the commitment of the two partners changes with respect to a certain proposal, we can account for how negotiation unfolds over several turns, and we can overcome the problem of recognizing implicit and / or "passive" acceptances. In this way, we are able to identify more reliably whether a certain proposal is jointly committed to at a given point in the dialogue.

To our knowledge, previous empirical work on the agreement process has focused on just one component of the process. For example, Walker (1996) studies acceptances and rejections, but does not try to characterize what is accepted or rejected, as we do here. Other empirical work addresses a more abstract level of analysis, namely, it tries to characterize strategies conversants adopt to reconcile their views after a disagreement (Chu-Carroll and Carberry 1998). Our work goes one step further by trying to correlate different dialogue patterns with the theoretical constructs, such as partner decidable options and proposals, that motivate them.

The context of our empirical analysis is the COCONUT project.[1] COCONUT's long-range goal is to create a unified architecture for collaborative discourse, accommodating both interpretation and generation. Our computational approach (Thomason and Moore 1995; Thomason and Hobbs 1997) uses a form of weighted abduction as the reasoning mechanism for both interpretation and generation. The regularities that emerge from the empirical analysis of the data are intended to be incorporated in the computational model in order to constrain computation: they will limit the set of axioms the abductive reasoner has at its disposal at any given time, and adjust the weights on those axioms. Given the goal of this paper, we will not discuss our computational model; we refer the reader to (Thomason and Moore 1995; Thomason and Hobbs 1997) for further details.

A caveat before proceeding. As we collected *computer-mediated* human-human dialogues for a simple design task, it is legitimate to ask whether our findings can be applied to other kinds of human dialogues. We claim that they can, because, as we will show, the basic agreement process is not affected by the particular modality of the dialogue. In

1 See http://www.isp.pitt.edu/~intgen.

Section 2 we discuss the features of our task and setting, their correlations to other types of computer-mediated communication, and our reasons for claiming that the agreement process is not affected. In Section 3 we describe the theoretical framework we are assuming, instantiate it on the basis of the features of our task, and specify the aspects that we validate with our corpus analysis. In Section 4 we describe our coding scheme, and in Section 5 we use the coded data to validate our theoretical claims via the correlations we found in our corpus. Finally, in Section 6 we speculate on how to apply our model to other types of collaborative dialogues, and we conclude.

2. The COCONUT Corpus

2.1 Overview of the Task

We collected 24 computer-mediated design dialogues in which two people collaborate on a simple design task, buying furniture for the living and dining rooms of a house.[2] The task is related to those described in (Walker 1993; Whittaker, Geelhoed, and Robinson 1993) but differs in the communication setting and the emphasis and complexity of the task[3] as will be described below.

For the COCONUT task, each person is given a separate budget and inventory of furniture that lists the quantities, colors, and prices for each item in that inventory.[4] Neither participant knows what is in the other's inventory or the money that the other has. The participants have the same types of knowledge but different instantiations of it. By sharing information about their different instantiations during their conversation, the participants can combine their budgets and can select furniture from each other's inventories. The problem is collaborative in that all decisions have to be consensual; funds are shared and purchasing decisions are joint. The participants are equals in that there is no master-slave or expert-client relationship. Both participants have been briefed on the task goals, incentives and tools and have had no prior contact.

The participants' main goal is to negotiate the purchases; the items of highest priority are a sofa for the living room and a table and four chairs for the dining room. The participants also have specific secondary goals which further complicate the problem solving task. Participants are instructed to try to meet as many of these goals as possible,[5] and are motivated to do so by associating points with satisfied goals. The secondary goals are: 1) Match colors within a room, 2) Buy as much furniture as you can, 3) Spend all your money.

There are two other cases in which participants might need to negotiate: 1) when the goals are not all achievable they must negotiate which ones to pursue, 2) when one participant wants to explore alternatives and the other is not as motivated to get a good score and is willing to settle for the first reasonably good solution (i.e. their notions of satisficing solutions (Simon 1955) are different).

2.2 An Overview of the Task Setting

The participants are in separate rooms and can communicate via the computer interface only. They are asked to maintain private graphical representations of their discussions and incremental agreements. The participants share dialogue windows but the inventories,

2 9 of the 24 dialogues were analyzed by 2 annotators. We also collected 12 trial dialogues that are not included in the corpus.
3 Walker's similar task is performed by two artificial agents whereas our task and that in Whittaker et. al. is performed by two humans. Whittaker's dialogues are spoken whereas ours are written.
4 In Walker's task this information is committed to memory but in our task the participants have this information in written form.
5 In Whittaker's task the incentives and goals are simpler.

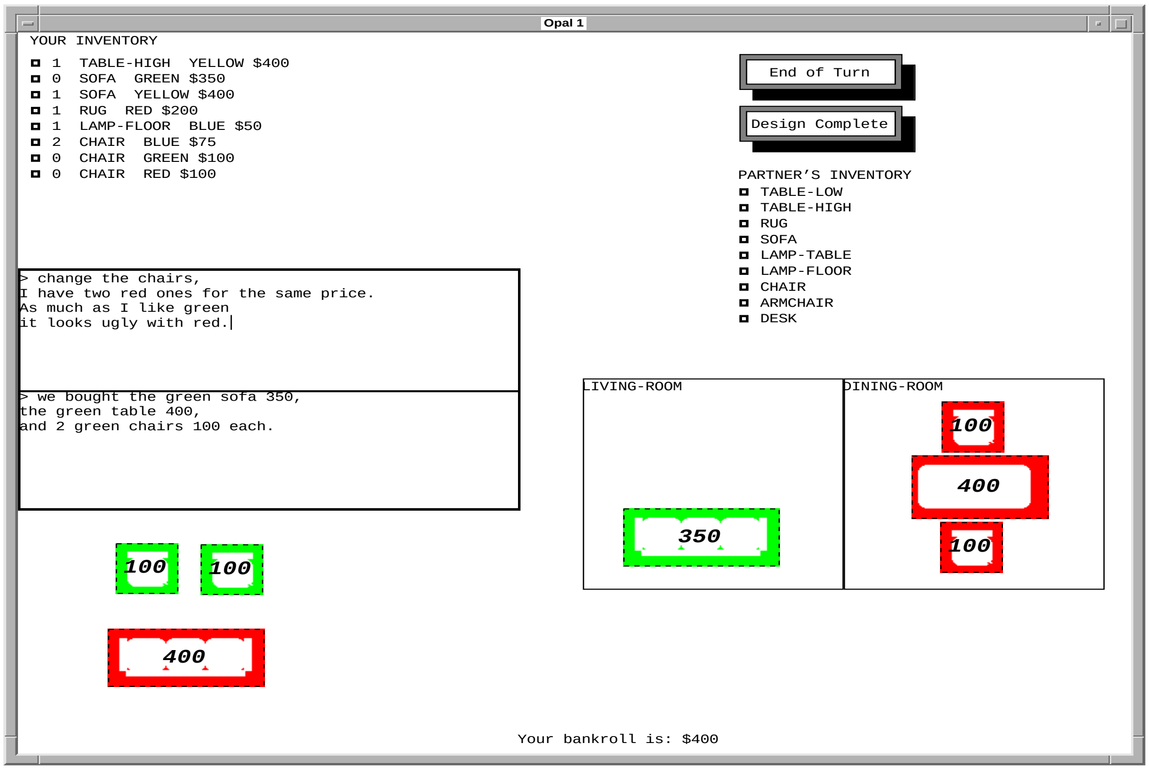

Figure 1
A View of the COCONUT Interface

budgets and updated floor plans are private and appear only on the owner's color display. Figure 1 shows the interface as it looks in the middle of a design session.

The buttons in the upper right corner of Figure 1, "End of Turn" and "Design Complete", enforce turn-taking and initiate the incremental recording of the conversation and the graphics updates. No interruption of the partner's turn is allowed. Also note that only the participants' current turns are available, i.e., the turn being currently held in the top dialogue box and the partner's previous turn in the bottom one.

During an incremental recording, the most recently transmitted message is recorded as well as the state of the sender's graphics display. The graphics display record is a description of the furniture icons in the two rooms as well as those that have been created but not assigned to any room. The interface has the additional feature that each furniture icon is initially displayed with a dashed outline around it. The participants are given the option of turning off the dashed outline to note that they believe agreement has been reached on using the item in the solution. In Figure 1 all furniture icons have dashed outlines: the sofa and chairs in the lower left corner have been mentioned, but not assigned to any room, whereas the rest of the furniture has been assigned to a specific room but not yet committed to. Given the previous turn (displayed in the bottom dialogue window), the participants would have turned off the dashed outlines on the furniture icons if the participant hadn't backtracked. However, in practice the participants are not consistent in using the dashed indicator to reflect incremental agreements. They either did not use the option or delayed until the last turn of the dialogue. However, the participants do consistently incrementally update the floor plan by placing the furniture icons in meaningful locations. Whenever possible we have used this private information in our corpus analysis as partial evidence of what the speaker's utterance meant and what the hearer understood. However, the primary purpose of the graphics display is as a memory aid for the participants and is only intended secondarily to help clarify possible sources of misunderstanding during analysis.

Note that since a participant does not know what furniture his partner has available, there is a menu (see the mid-right section of the display in Figure 1) that allows a participant to define furniture icons that represent what he understands his partner to have, as his partner shares this information with him. There is nothing to prevent the participant from creating an icon for a piece of furniture the partner does not actually have since the menu is general. An icon for a non-existent item could result from either a misunderstanding of his partner's item description or an error in selecting feature values for the item. At minimum the participant must know the type of the furniture item (e.g. chair, table). If the participant does not know or is uncertain about any of the other feature values of the furniture item, he can leave that feature unspecified (i.e. color and purchase price).

The participants first did a trial problem to familiarize themselves with the task and the communications setting. During this time they could ask for guidance on using the interface and clarification of the goals and the point system. We do not include the dialogues from the trial problem in the corpus. The participants then solved 1-3 scenarios where the inventories and budgets vary. The problem scenarios ranged from ones where items are inexpensive and the budget is relatively large to ones where the items are expensive and the budget relatively small. When the primary goals are harder to achieve, it can lead to either backtracking to find a better solution (see Figure 9),[6] or goal changes, or both.[7]

2.3 Possible Effects of the Setting and Task on the Dialogues

Our corpus is a collection of computer-mediated dialogues. Computer-mediated communication is a genre which is attracting more and more interest, and whose specific features are being actively studied, see for example (JCMC ; CMCM). However, since in this paper we are making general claims about how people achieve agreement, we need to question whether our findings from the COCONUT corpus will generalize to other tasks and settings. Our task is different from a real one in at least two ways: first, it is simpler; second, our participants were college students satisfying a psychology requirement, and thus, with no real motivation to engage in the kind of behavior we sought to analyze. Moreover, some constraints we imposed on the interaction, in particular strict turn taking, do not reflect face-to-face conversation.

2.3.1 Characterizing the Task Effects. Most studies on collaboration in computational linguistics have concentrated on dialogues that involve planning or scheduling tasks (among many others, see for example (Lochbaum, Grosz, and Sidner 1990; Ramshaw 1991; Sidner 1994; Walker 1996; Chu-Carroll and Carberry 1998)). It is not unprecedented to select a simple task taken out of a larger context in order to control the situation and potentially allow for a more objective analysis. In the case of the COCONUT task, we too opted for a simple decontextualized task but we chose to analyze a design

6 All the dialogue excerpts included in this paper appear as they have been recorded, i.e., typos have not been corrected. However, the dialogues are presented broken into utterances, according to an algorithm based on the one proposed by Passonneau (1994) Further details on our turn-breaking algorithm can be found in (Di Eugenio, Jordan, and Pylkkänen 1998).

7 In this overview of the dialogues, our goal is to give the readers a general impression of the corpus. None of the characterizations in this section have been empirically validated unless otherwise indicated. See Section 5 for the empirically validated hypotheses.

task instead. Note that there is nothing about design tasks in general that should result in a contrived or artificial dialogue, and the participants in our experiments were free to say whatever they wished. Also, we expect the agreement process to be applicable across many types of tasks, given the domain independent nature of IRMA and Clark's acceptance process (both will be described further in Section 3).

We view the design task as part of a larger product realization problem that also encompasses both planning and scheduling as subtasks; see, for example, (Lyons 1995). For product realization, one typically needs to plan and schedule the design subtasks, and the resources and processes needed to manufacture the product. However, the design task itself primarily involves the negotiation of product features and the constraints between dependent design subtasks. The design task is most advantageously represented as a constraint satisfaction problem whereas this is not usually the case for planning and scheduling tasks.[8]

To better understand what we mean by a design task, consider a group of electrical engineers who have the task of designing a circuit board.[9] They have the required functionality defined at a high-level but have to decide on which off-the-shelf integrated circuits (ICs) to use to help achieve this functionality. They have a variety of ICs to choose from, where the choices offer a different but overlapping set of functions and have different costs and different impacts on the overall design of the new circuit board. Their goal is to make the board as cheaply as possible but to consider future enhancements and all the products that this particular board might be part of. Their choices may prove beneficial for some products and detrimental for others. Because of this, the team of electrical engineers need to communicate with one another to negotiate the options at the goals level. Since it might not be possible to meet all of the goals with a single design, they may have to give up on trying to meet some of the constraints. All of the designers involved may know of some ICs that are available but since there is a vast IC market, it is reasonable to assume that there will be some possibilities that are not mutually known among the team.

In general, there are many areas that participants might negotiate in a problem solving situation. Among the possibilities are goals, actions, and parameter values. These possibilities arise because of resource-boundedness, which affects a participant's ability to generate solutions and deliberate about them, and because of the distribution of knowledge. These possible areas of negotiation should apply to all problem solving tasks.

When information is nearly evenly distributed, as with information exchanges, where the types of knowledge are divided, and as with design tasks, where instantiations of knowledge differ, we expect to see more instances of negotiation since all the participants have something to contribute towards finding a solution (Walker 1995). Dialogue initiative is a potential indicator of who is contributing to the dialogue. (Walker 1995) shows that dialogue initiative is more evenly distributed in information exchange than in instructional dialogues. The same should hold for design tasks as for information exchanges.

Furthermore, we expect both the type of task and the information distribution to influence which area one typically sees being negotiated. Planning tasks in which knowl-

8 While it has been hypothesized that planning problems (Joslin 1996) and scheduling problems (Qu 1997; Walker et al. 1997) can be translated into constraint satisfaction problems, we think that it is still useful to distinguish planning, scheduling and designing and that it may be more helpful to represent these tasks with specialized languages. However, it seems intuitively clear that the existing planning languages offer few advantages for most design problems, due to their complex interdependencies and the sparse space of domain action types.

9 For other design scenarios, see electromechanical (Lyons 1995) and architectural (Lottaz 1996; Lottaz and Smith 1997) design task descriptions.

edge of possible actions and desires are either shared or evenly distributed would seem to focus on negotiations about actions or which desires to address. This happens in some of the family interaction dialogues collected by Condon and Čech (1996). In one instance the participants negotiate their activities while planning their weekend. On the other hand, planning tasks in which knowledge of the problem lies with one participant and actions with another, as with advisory dialogues, would seem to focus on negotiations about understanding the problem and its solution. In an advisory dialogue, the problem itself is not typically negotiable (e.g. if something is broken and you need advice on fixing it, the problem is typically static) but the solution might be. With scheduling tasks, the knowledge of parameter values may be evenly distributed, so that the negotiation is typically focused on parameter value assignments. See the dialogue in Figure 2 from to the VERBMOBIL corpus (http://www.dfki.de/dfki.html). With design tasks, when the goals are mutually known but not necessarily all achievable, the focus may include negotiating what the goals should be (see Figure 3, utterances [12] and [15]). If the values for parameters are evenly distributed then negotiation will also emphasize assigning values to parameters.

For the COCONUT task, parameter values are equally distributed and the goals are not always all achievable. Because of this we expect to see more negotiation in these two areas. Therefore the COCONUT dialogues can offer additional insights into the agreement process that we might not gain by looking just at advisory dialogues or scheduling and planning dialogues.

```
B: I think the first day that is really good for me
   is the eighteenth that is a Tuesday
A: okay want to have lunch
B: that sounds pretty good
   are you available just before noon
A: we can meet at noon
B: sounds good
   on campus or off
A: your choice
```

Figure 2
Scheduling - negotiating parameter values

The size of our design task is smaller than what is typical of "real world" design problems,[10] but the size should not affect the nature of the agreement process relative to the goals or the parameter values. We simply expect to see fewer instances of the agreement process when solving one instance of a COCONUT design problem than when solving a real design problem.

2.3.2 Characterizing the Setting Effects. Since communication settings are often viewed as derivatives of face-to-face communication (Clark 1996), we will consider how our setting differs from face-to-face communication and whether we expect these differences to have an impact on the general agreement process. Future work will show whether our expectations are correct.

The turns tend to be longer than in face-to-face communication, because we did not allow the participants to interrupt each other, and because feedback cannot be as fine-

10 We only have 6 basic constraints and 21 parameters, whereas real design tasks are much larger. For example, the construction domain example (Lottaz and Smith 1997) notes that it had 42 constraints and 54 parameters.

```
Ju: [3] i have a variety of high tables...
    [4] green, red and yellow for 400, 300, and 200 respectively.
    [5] so let's begin in the dining room.
    [6] do you have any high tables

Jo: [7] Yes i do,
    [8] A high table blue,green,red,at 400,200,400 respectively.
    [9] My cash flow is 450.

Ju: [10] hmmm...
    [11] well
    [12] i we can try to color coordinate with chairs.
    [13] i have 3 green (50) 5 yellow (150) 2 red (50)
    [14] but colors aren't that many points...
    [15] maybe we should just try to accumulate furnature.
    [16] in that case: you about your green table adn some cheap chairs?

Jo: [17] well
    [18] i have 6 chairs
    [19] that are green for 100
    [20] or 2 chair for 50?
    [21] so how what do you think?
```

Figure 3
Designing - negotiating goals

grained.[11] By comparing the COCONUT dialogue in Figure 4 with the more interactive dialogue from (Whittaker, Geelhoed, and Robinson 1993) in Figure 5, one sees that the turns are longer and less interactive.

The initial turns can also be much longer than later ones when the participants have a strategy of giving each other a complete dump of all that they know is not mutually known, as in the excerpt in Figure 6. Although the participants can sometimes solve the problem in fewer turns by sharing everything that isn't mutually known at the beginning,[12] this does not mean that the participants are able to completely avoid having to negotiate (see [5]-[18] in Figure 6). A plausible reason why negotiation is still needed is the difficulty of generating and deliberating about all the satisficing solutions (due to memory and time limitations). This is a possible reason why one might find it beneficial to collaborate and why an artificial agent might be a good design collaborator (see (Lottaz 1996; Lottaz and Smith 1997)).

Without doing a formal study of the setting effects, we can only project a range of expected effects by examining the results of formal studies of similar settings. In our case, we will consider studies of the videoteleconferencing setting. The ability to interrupt is seriously disrupted in this setting and the participants tend to develop a formal method to relinquish the floor to the next speaker to compensate for this difficulty.[13] The noted effect of fewer interruptions is that the turns are longer and mutual understanding is harder to achieve (Whittaker 1995). In COCONUT, we prevent interruptions and provide a

11 (O'Conaill, Whittaker, and Wilbur 1993) notes the relationship between turn length, feedback and interactiveness.
12 7 of the 24 COCONUT dialogues used this strategy, and 2 of these 7 are part of the 9 that we annotated. The design corpus from (Whittaker, Geelhoed, and Robinson 1993) also has some instances of this type of strategy.
13 The difficulty is attributed partly to half-duplex voice transmission and partly to the loss of directionality of voice (Whittaker 1995).

```
S-1:  [1]: I have $550 to spend.
      [2]: How much do you have?

G-1:  [3]: I have 450 dollars

S-2:  [4]: So we have a total of $900 to spend.
      [5]: The cheapest sofa I have costs $300
      [6]: and it is yellow.
      [7]: How much are your sofas?

G-2:  [8]: My sofa's are more expensive
      [9]: so buy your $300 yellow sofa.
     [10]: Also.... Your $550 + my $450 is $1000

S-3: [11]: You are right,
     [12]: we have $1000 to spend.
     [13]: I will go ahead and buy the $300 yellow sofa.
     [14]: That leaves me with $250 to spend.

G-3: [15]: I have table-high for $200 green.
     [16]: Do you have any cheaper?
```

Figure 4
An interruption-free design dialogue

```
A-1: [1] I have-hello?

B-1: [2] Hello.

A-2: [3] one big piece of 40 that fits in the left window-in the left
room exactly.  It is-the bigger size is 12 x 10.

B-2: [4] 12 x 10.  It is a sort of triangle.

A-3: [5] Yes, a sort of triangle again.

B-3: [6] And er ok, could you describe it?

A-4: [7] And I have one-let me describe the pieces that I have first.

B-4: [8] Have you got a big pieces-piece for the other, for the right
rectangle.
```

Figure 5
An interruptible design dialogue

formal method for relinquishing the floor, so we expect the COCONUT dialogues to have these features. However, there is good reason to believe that any additional difficulties in achieving mutual understanding were overcome by our subjects: all the COCONUT participants believed they had reached an agreement and their final floor plans were identical in all but one case. To achieve mutual understanding the participants must have compensated by being more explicit. Having more explicitness in the dialogue can only be advantageous for our studies of the agreement process, since the process is more clearly visible.

The COCONUT dialogues appear to have more overlapping discourse segments than one might expect for spoken dialogues. This may be due to the longer, less interactive

```
J-1: [1]: have $450.
     [2]: 1 table high 400/1 table high green 200 1 table high red 400/
          1sofa blue 400/1 sofa green 550/1 sofa yellow 350
          1 rug yellow 150/1 lamp floor blue 250/6 chair green 100/
          2 chair red 50

D-1: [3]: i have $550.
     [4]: 1 table high green 400, 1 table high yellow 200/
          1 table hi red 300/1 sofa blue300/ 1 sofa yellow 400/
          1 sofa red 550/ 1 rug blue 250/1 lamp floor yellow 150/
          3 chair green 50/ 5 chair yellow 150/2 chair red 50.
     [5]: I suggest that we buy my blue sofa 300,
          your 1 table high green 200, your 2  chairs red 50,
          my 2 chairs red 50
     [6]: and you can decide the rest.
     [7]: What do you think

J-2: [8]: your 3chair green my high table green 200and  my 1 chair green 100.
          your sofa blue 300 rug blue 250.
     [9]: we get 700 point.
     [10]: 200 for sofa in livingroom plus rug 10
     [11]: 20 points for match.
     [12]: 50 points for match in dining room
     [13]: plus 20 for spending all.
     [14]: red chairs plus red table costs 600
     [15]: we only 650 points without rug and bluematch in living room.
     [16]: add it up
     [17]: and tell me what you think

D-2: [18]: Your perfectly right
     [19]: you are so much better than I am at this stuff.
```

Figure 6
An "initial dump" strategy design dialogue

turns. For example, in Figure 7 utterance [11] presents an optional item for the living room (an inference supported by the graphics displays) and then goes on to discuss items for the dining room in [12]. Here, the current turn holder starts the discussion of the new parameter before getting feedback from his collaborator, presumably on the assumption that the partner should accept the solution presented in utterance [11]. In this type of situation, one turn involves several separate agreement processes. The agreement process for the optional item in [11] is still open when the agreement process starts for the dining room items in [12]. In [19] the agreement process for the lamp and table are addressed so that the agreement process for the lamp and dining room items cross each other with respect to the previous and current turns. In a spoken dialogue with no imposed turn-taking mechanism, one would expect immediate feedback on the lamp before M goes on to address other furniture item decisions. However, the interdependencies of the task can cause overlaps as well. In Figure 8 (from the corpus described in (Whittaker, Geelhoed, and Robinson 1993)), a decision about an item worth 40 points which is first described in [1-7] and an item worth 20 points that is first described in [9] are left pending until utterance [24] for the 40 point item and utterance [77] for the 20 point item. Since this dialogue setting is interruptible speech, it is questionable to what degree, if any, the setting is responsible for the overlapping discourse segments.

However, whether the agreement processes overlap or not, the basic process is still the same. The overlapping discourse structure should not negatively affect our results,

```
M-1: [7]-[9]
     [10] I do not have a sofa for a better price
     [11] but, i do have a lamp-floor, blue (250).
     [12] i have a green table (200) and four chairs for (75) a piece.
     [13]-[15]

D-2: [16]-[18]
     [19] the lamp and table sound good,
     [20] but the chairs seem expensive.
```

Figure 7
Overlapping Agreement Process in Interruption-Free Dialogue

```
1-7. [description of item worth 40 points for the left room followed
     by a request for permission to describe others]

8. B: Have you got a big pieces-piece for the other, for the right
rectangle.

9. A: For the right rectangle, 6 x 12.  I have the same-that's the
size of the room, 6 x 12.  I have one piece that is 20 points worth,
and that is 6 x 6, the same as you had before in the previous thing.

10-22 [description of item worth 20 points]

23.A:
yes, I have-let me describe what I have-all the pieces I have.  I have two
pieces of one of 20 and one of 25, that both fit in the right-hand er room;
and I have this other piece of 40 and a piece of 35, that I think fit
together in room 1, the left-hand room.

24-31. [decision to use items worth 40 and 35]

32-59.  [describe items worth 40 and 35 points to orient in room]

60-65  [describe fill-in items for the room]

xx.A:
Now I have two pieces for this room: one is value 20, and one is
value 25.

66-76  [describe item worth 25]

xx.B:
Wow!  So I have fit both your pieces, 20 and 25, and another piece of mine
which is worth 20.

xx.A:
Hm.

xx.B:
And, if I am not wrong, I've filled completely the rectangle.

77.A:
Ok.
```

Figure 8
Overlapping Agreement Process in Interruptible Dialogue

which concern the structure of the agreement process rather than the discourse structure itself.

```
S: [51] wait.
   [52] I think we should change our ugly red dining room around.
   [53] I have a green table that only costs 200.
   [54] I'll get the table
   [55] and you buy the green chairs to go with it.
   [56] even if i have to by two of my 100 dollar green chairs
   [57] we would save $100.
   [58] what do you think?

R: [59] 1, I only have two green chairs,
   [60] do you have two more,
   [61] and 2, what do you plan on doing with this 100, we saved.
   [62] I say we stick with what we got and  get outta' here.

S: [63] yes i have 2 green chairs.
   [64] With the extra 100 i planned on
   [65] getting rid of the blue sofa,
   [66] getting a yellow one i have for 350
   [67] and get a matching yellow rug with it for only 150.
   [68] don't worry
   [69] my screen still has the original stuff

R: [70] Alright,
   [71] Now I have 4 $50, green chairs in the DR, two from your
        screen two from mine = $200, 1 green $200, table in the DR.
   [72] In the LR I have 1 yellow $150 rug and one yellow $350 sofa.
   [73] = $900 and 660 pts.
   [74] Right?
```

Figure 9
An example of backtracking

Finally, as our dialogues were elicited during experiments (as mentioned, the participants were college students satisfying a psychology course requirement), there might be a risk that the data is simplified with respect to natural collaborative dialogues. We already extensively argued that our task and setting do not affect the basic agreement process. Furthermore, we feel that the constraints on the task, such as the budget constraint, and the point system that we associated with achieved goals, were sufficient to ensure that our participants engaged in full fledged negotiation (some suggestive evidence is provided by Table 4 in Section 5). Not only did participants in our experiments negotiate solutions, they often negotiated which goals to achieve, and they sometimes even backtracked on a solution already agreed upon, as shown by the dialogue excerpt in Figure 9. The two participants S and R have already committed to a complete dining room set. However, in [52] S starts an extensive backtracking that will result in changing the sofa ([65]-[66]) they had also already agreed on.

3. Modeling Collaborative Problem Solving Dialogues

In this section, we provide the theoretical background for our version of the agreement process, and we highlight the predictions that we wish to investigate. To situate the discussion, let us start with a claim that we will substantiate later, that our participants' collaborative behavior can be modeled according to a *Balance-Propose-Dispose* agreement process, schematized as follows:

(1a) balance information and partially deliberate

(1b) propose options

(1c) dispose of proposal

This characterization of the agreement process for our dialogues derives from combining IRMA (Bratman, Israel, and Pollack 1988), the basic present/accept mechanism used to establish the mutual belief needed for an agreement to be achieved (Clark 1996), and an ability to reason about the possible contributions of a collaborating partner.

```
Ju-4: [32]: well
      [33]: how about we use 2 of your chairs and 2 of my red.
      [34]: we will have a christmas room...
      [35]: i have a blue sofa for 300.
      [36]: it's my cheapest one.

Jo-4: [37]: I have 1 sofa for 350
      [38]: that is yellow
      [39]: which is my cheapest,
      [40]: yours sounds good.

Ju-5: [41]: ok
      [42]: i logged in 2 of your chairs and 2 of mine..
      [43]: both red.
      [44]: I'll order that blue sofa.
```

Figure 10
Example of a partner decidable option ([35])

To make the discussion more concrete, we will refer to the examples in Figures 10 and 11. As [35] is the first mention of a sofa in the conversation, we contend that it opens a new agreement process, concerning the choice of sofas. Because the knowledge preconditions (Moore 1985) for deliberating to the point of making a commitment (we will call this *full deliberation*) have not been met, we claim that [35] cannot count as a proposal for a solution that includes Ju's sofa. Whereas Ju in [35] offers her sofa for consideration, i.e., offers it to Jo as an option he can use in problem solving (what we will call a *partner decidable option*), it becomes effectively proposed only after the exchange of information in [37]–[39]. However, note how [15] in Figure 11, which is semantically equivalent to [35] in Figure 10, would count as a proposal because it could be fully deliberated.[14] We claim that we can explain the difference between [35] and [15] in the two excerpts by appealing to IRMA and considering them as belonging to different phases of an agreement process that emphasizes being able to deliberate to the point of making a commitment.

We will start by discussing the two theories that are the basis for our model, namely, IRMA and Clark's view of language as collaboration. We will then discuss how our model applies to our dialogues, and the predictions we make on the basis of the model, predictions that we will later verify with our empirical study.

We will conclude this section by showing that many collaborative cycles described in the literature (Sidner 1992; Walker 1993; Sidner 1994; Chu-Carroll and Carberry 1998)

14 The omitted utterances in Figure 11 have nothing to do with the task—the two participants joke about a hypothetical party that destroyed all the furniture.

```
D-1: [1]-[3]
     [4] I have a red sofa 400 and a table high 400
     [5] with 550 to spend
     [6] let's start here.

M-1: [7]-[10]
     [11] I have 400.
     [12] that is it.
     [13] everything is expensive.
     [14] lets just get the basics.
     [15] i have a sofa, green for 350.

D-2: [16]-[18]
     [19] the sofa sounds good,
     [20] i have 2green chairs
     [21] we could put with the diningroom table 100 each.
     [22] and that would spend all of our money.
```

Figure 11
Example of a fully deliberated proposal ([15])

can all be seen as instantiations of processes that integrate an architecture for rational agency such as IRMA and a model of language as collaboration such as Clark's. The specifics of our and other researchers' agreement processes can be seen as manifestations of the recursive nature of the process, coupled with the characteristics of resource-bounded practical reasoning and its manifestations in language that different researchers wish to explore.

3.1 Theoretical Underpinnings

IRMA will be the framework from which we derive predictions about the agreement process and its realization in dialogue. If one thinks that language is action and should be resource-bounded, then IRMA should work for both language and domain actions. However, IRMA is not sufficient, as there is a gap between the model of resource-bounded rational problem solving that it provides and how such a model affects communication. Although perception is taken into account in IRMA, this does not directly explain how conversants interpret and produce language, the discourse strategies they use, and the unfolding of negotiation in a dialogue. Clark and his collaborators' work provides a model of collaboration in dialogue that is an ideal candidate to bridge the gap. This is why we believe we should be able to model collaborative problem solving dialogues more effectively by integrating the two models.

3.1.1 IRMA: Means-end Reasoning and Resource Bounds. The Intelligent, Resource Bounded Machine Architecture commonly known as IRMA[15] was proposed in (Bratman, Israel, and Pollack 1988; Pollack 1992) as an architecture for resource-bounded practical agents. This architecture addresses the issue of how a single resource bounded agent can perform means-end analysis, weigh competing alternatives and act on its intentions. Figure 12 schematizes the IRMA architecture.

IRMA is based on Bratman's fundamental idea (1990) that the agents' planning commitments (*intentions structured into plans* in Figure 12) constrain subsequent reasoning

15 As far as we know, the acronym first appeared in (Pollack 1992), and denotes a slightly simplified architecture with respect to the one in (Bratman, Israel, and Pollack 1988). However, Pollack appears to consider the two architectures as equivalent, and refers to both as IRMA.

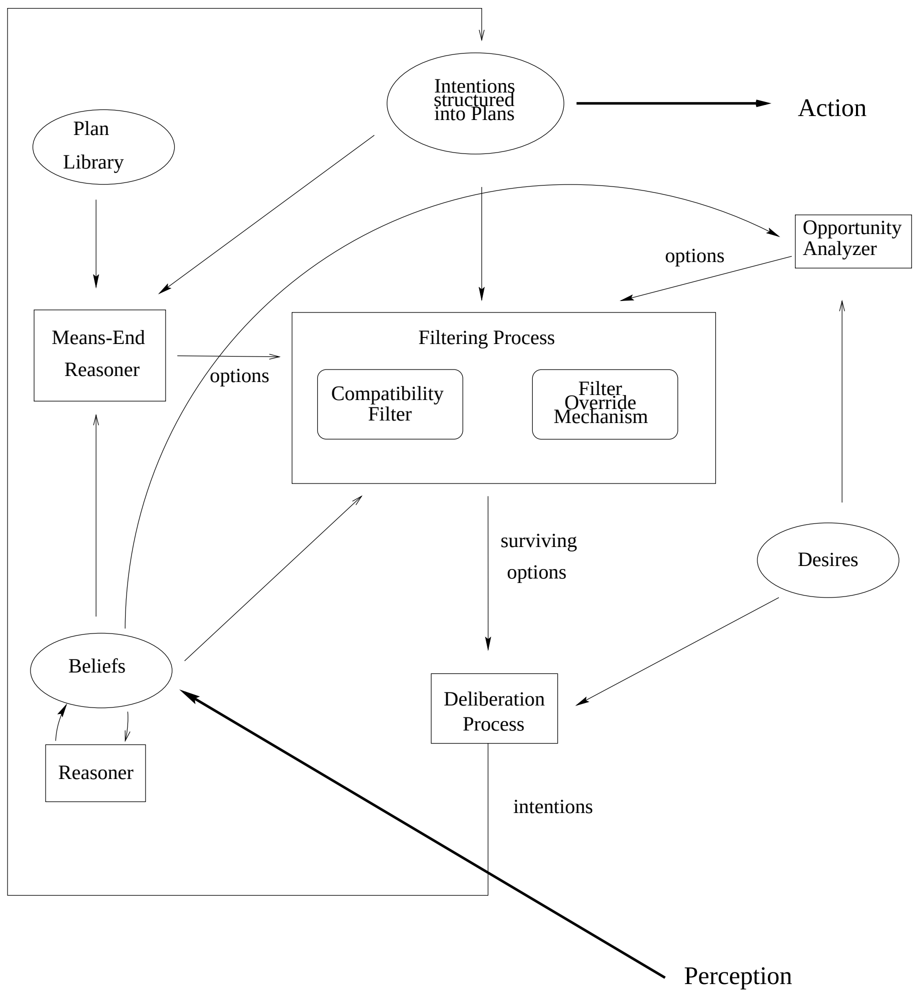

Figure 12
The IRMA architecture

in two ways. As input to the means-end reasoner, they guide reasoning: for example, the means-end reasoner will fill in the details of partial plans by drawing on its plan library. As input to the filtering process, they limit the scope of *deliberation* to the options that are compatible with them. Naturally, a previous commitment to a plan may be subject to reconsideration or abandonment in light of changes in belief. However, an agent can't constantly reconsider its plans, so that plans are relatively stable.

Options are produced both by the means-end reasoner while filling in partial plans, and by the *opportunity analyzer*; the latter is the component of the architecture that responds to perceived changes in the environment, and proposes new options to an agent. Options are subject to the compatibility filter, which checks whether they are consistent with previous plans. When options don't survive the compatibility filter, they have the potential of triggering a *filter override*. This mechanism encodes the conditions under which some portion of the agent's existing plan should be suspended and weighed against some other option. Finally, options that survive the filtering process are passed to the *deliberation process*, which weighs one option against the other and produces intentions to be incorporated in the agents' plans.

3.1.2 Clark's model. Clark advocates the view that *speaking and listening are not autonomous activities, but parts of collective activities* (Clark 1992, p. xvi). Two issues that Clark considers necessary to investigate in order to support this view are: what constitutes the *common ground*, i.e., the information shared by both participants, and how *collaboration* in language works. Common ground is naturally dynamic; new beliefs are added to it through the process of *grounding*.

In his most recent book, Clark (1996) articulates a theory of *joint action in conversation*. Of particular interest to us is his contention that joint actions in conversation occur at many different levels (he refers to a *ladder of joint actions*). He argues that the following four levels are necessary, and others may be possible. Given two speakers A and B, at level 1 (the bottom level), A executes behaviors and B attends to them; at level 2, A presents a signal and B identifies it; at level 3, A signals something to B, and B recognizes what A means; and at level 4, A proposes a joint project and B considers taking up A's proposal (*disposition* of a proposal).[16]

Grounding occurs at levels 1 through 3, in order to support the joint project at level 4. Grounding takes place by means of *contributions*, which Clark defines as a *signal successfully understood* (1996, p. 227). Contributions in turn are composed of two phases, *present* and *accept*, as follows (Clark 1996, p. 227):

> *Presentation phase*. A presents a signal s for B to understand. She[17] assumes that, if B gives evidence $\boldsymbol{e}$ or stronger, she can believe that B understands what she means by it.
> *Acceptance phase*. B accepts A's signal s by giving evidence $\boldsymbol{e'}$ that he believes he understands what A means by it. He assumes that, once A registers $\boldsymbol{e'}$, she too will believe he understands.

Both *presentation* and *acceptance* phases can be complex and each can have a hierarchical structure, i.e., they may contain embedded contributions. There are various types of evidence that B can employ, which correspond to the four levels of joint actions required by communication. B's evidence may be provided at level 4 by an appropriate *disposition* of A's proposed joint project, as when A asks a question and B answers it.

16 Note that a disposition of a proposal is not the same as agreeing to what is being proposed.
17 Suppose A is female and B is male.

We will not make much explicit use of Clark's notion of grounding in this paper, as in our dialogues the vast majority of the times B's evidence is in fact provided at level 4. However, we owe the idea of *present/accept* phases, and of *disposition*, to his work.

3.2 The Agreement Process in COCONUT

In general, the specific realization of the agreement process in our dialogues is, as we pointed out earlier, derived from IRMA along with the basic present/accept mechanism used to establish the mutual belief needed for agreement. To agree, both parties must mutually believe that both are committed to the decision. Presenting an option in a context where the speaker has fully deliberated shows the speaker's commitment to using that option in the solution, and accepting the utterance shows the hearer's disposition of the utterance. If the hearer correctly understands the utterance then he will have recognized the speaker's commitment. As a result of the disposition of the utterance, if the hearer agrees about using the option in the solution, he can either explicitly express his own commitment or move on to a new part of the problem (as predicted by Clark's observation regarding strengths of evidence). The two agents will then have reached joint commitment with respect to that option. One added element of our instantiation of the agreement process is the balancing of information about negotiable elements. We assume that both parties are aware of one another and reason at an abstract level about what they expect the other partner to be able to contribute to the collaborative effort. In this way they can use the context of the expected state of the other's problem solving effort to help recognize when a commitment should be possible.

We will now present our proposed integration of IRMA with some of the ideas from Clark's work. To do so, we will first discuss our interpretation of IRMA as applied to the COCONUT task, focusing first on a single agent, and then relating the single agent process to the collaborative agreement process.

We will only attempt to explain how IRMA applies to the design task and dialogues to the degree needed to support our instantiation of the agreement process. Many of the details of how IRMA might apply to language and domain actions remain to be worked out. Our goal is to explain the basis for the predictions we wish to empirically investigate. As more empirical investigations are conducted, we expect to get a clearer picture of whether such a mapping is reasonable and if so what the mapping should be.

We believe ours is the first attempt to relate IRMA to language in such detail. The only other researcher we know of who explicitly invokes IRMA as underlying collaborative conversation is Walker (1993; 1995). However, in Walker's work the agreement process is not directly motivated by IRMA and communicative decisions are made peripherally to IRMA. Discourse strategy decisions depend on what is salient in memory, and while saliency and memory content are effected by IRMA's domain reasoning, IRMA is not used directly to reason about, filter or deliberate about communicative options.

3.2.1 An Interpretation of IRMA for the COCONUT Task. We will attempt to describe what we believe is the most plausible interpretation of IRMA for the COCONUT task. In this description we will take the viewpoint of a single IRMA agent (the agent—a female) interacting with another IRMA agent (the partner—a male). We will also refer to dialogue excerpts in order to give readers an intuitive feel for the mapping we are advocating, even if we are not making any direct predictions about communication yet.

First we will assume that the agent's intentions can be structured into plans for language and constraint satisfaction problems (CSPs) for the design task. Some relevant intentions for a CSP are to set particular constraints and to make particular assignments for constraint equation parameters. Also, we will assume that the means-end reasoner might more generally include the ability to solve constraint equations. There could be many solutions for a partially specified set of constraint equations and each solution would be an option that gets passed to the filtering process. We will assume that for COCONUT each agent starts with a default CSP that is based on the problem description and priority goals (see Section 2.1).

The Opportunity Analyzer reacts to perceived changes in the beliefs and creates options based on the general desires of the agent with regard to the problem. This allows the agent to consider options that change the constraint equations themselves (e.g. whether or not to require color matches in a room, see [12]-[15] in Figure 3).

The filtering process considers options from both sources and passes on those that do not appear to conflict with any current intentions. Or, if there are noticeable conflicts, it may decide to override the filtering if the option appears to be worth pursuing. If at some point the means-end reasoner is unable to identify any options, this might lead to a filter override that will eventually allow changes to the domain constraint equations. Filter overrides would also allow the agent to consider alternative solutions for a parameter even though that assignment has already been committed to (either jointly or singly). For example, in the excerpt in Figure 9 at the end of Section 2, S and R have already committed to a dining set. However, with utterances [51]–[58] S reopens the parameter decisions for the table and chairs.

The options that pass the filter then must be evaluated by the deliberation process. During this process it might be the case that the agent is unable to make any commitments to a particular constraint change or parameter value assignment. This could happen if the agent does not have enough information to deliberate well. One strategy for overcoming this obstacle is to utilize a partner if one is available. We will say more about this in Section 3.2.2.

So far we have focused on how the agent goes about solving problems but we also need to consider what might happen when the partner communicates something to the agent. We expect that the partner will be very similar to the agent, so that the agent will have certain expectations about the partner that are based on IRMA.

When something is communicated to the agent, it is a perception that the agent must reason about to understand and act on. Alternative interpretations would be subject to filtering and deliberation. If a best option, i.e., a best candidate for the interpretation of an utterance cannot be determined then the agent might form an intention that will eventually communicate lack of understanding.[18] If the agent understands (i.e., she can select an option that explains the utterance) then the agent might add the contents as a new belief.

While attempting to understand the partner's utterances, the agent's beliefs will be changing. These changing beliefs may eventually lead to new task options being produced by the means-end reasoner or the opportunity analyzer.

3.2.2 Relating the Single Agent Processes to the Collaborative Agreement Process. Now that we have given our interpretation of the single agent processes we can show how we arrived at our instantiation of the collaborative agreement process. Although IRMA per se has not been augmented to account for collaboration, Bratman (1992) has extended his theory of intention to *shared cooperative activity (SCA)*, i.e., *collaborative activity*.[19] Bratman shows that the three features that identify SCA are mutual responsiveness, commitment to the joint activity, and commitment to mutual support. Some of the observations we will present here can be seen as going in the

18 This rarely occurs in the COCONUT dialogues.
19 Grosz (1996) shows that Bratman's criteria for shared cooperative activity are equivalent to hers for collaborative activity.

direction of making IRMA compatible with SCA, in particular as regards commitment to the joint activity.

Other areas of IRMA that need further work are opportunity analysis, filtering and deliberation: what happens during these processes is contentious, since they have not yet been studied to the same extent as has means-end reasoning.[20] We will mainly focus on the deliberation process because this is where commitments are made and because our instantiation of the agreement process as *Balance-Propose-Dispose* emphasizes deliberation.

For the multi-agent design case where the information needed to complete the task is equally distributed, we need to be able to explain what happens when an agent can't commit to an option that solves the domain problem. Because the information needed to solve the problem is equally distributed, both agents may be unable to meet the preferred goals on their own. Here, an awareness of what the partner might be able to contribute is useful. An agent who knows that the task information is equally distributed can reason that the partner might know of something that would help her find or identify good options.

One possibility for involving the partner is to ask him for the missing information. Another possibility is to ask him to provide a solution to the problem (Biermann et al. 1993). However, since the information needed to solve the problem is equally distributed, this might well mirror the original impasse. The agent might anticipate this problem by providing the partner with additional information that could better enable the partner to deliberate and find a good solution. In specific cases, the reasoning the agent has already done in partially solving the problem may suggest a focus on a particular goal, or cause her to present all the best contenders among the options, as [35]-[36] in Figure 10, etc. It depends on the context and the plans and strategies that the agent has available for overcoming this particular obstacle.

So we claim that when neither agent is able to deliberate to the point that they can make a commitment to a change in the problem state, they will simply do partial deliberations and balance the information distribution. This, then, is the first phase in the agreement process, step (1a). The *Balance* phase continues until at least one of the agents is in a position to fully deliberate and make a commitment to a problem state change. [35-36] in Figure 10 belong to this phase. Balancing is the phase to which most grounding activity belongs, as participants are building their common ground.

When an agent is able to find a good solution she is willing to commit to, she must get her partner's commitment in order to reach an agreement. She might do so by forming an intention to get a joint commitment to intend that change. The full deliberation and commitment of a single agent constitute the propose phase of the agreement process, step (1b).

One way for an agent to show her commitment is to explicitly propose the change, as with utterance [9] of Figure 4. If the change is to assign a value to a constraint parameter and the value isn't already mutually known then another way might be to communicate the existence of the value as with utterance [15] of Figure 11. Whether the latter is feasible depends on the context and the agent's plan library and his other beliefs as to how mutual commitment can be achieved (Thomason and Moore 1995). At the language action level, various options would be produced and filtered and deliberated about before a particular communicative intention is committed to. These constitute discourse strategy decisions, as empirically studied in (Walker 1993).

20 There are research efforts to further explore some of these processes; for example, (Ephrati, Pollack, and Ur 1995) investigates filtering.

If the partner recognizes the agent's commitment then he is expected to dispose of the proposed change and deliberate about it. Deliberation might require triggering a filtering override if the proposed change conflicts with previous commitments. If the partner didn't believe the agent was committed to the option then he should be less likely to override the conflicts.

The partner's expectations with respect to the agent can help him determine whether he thinks the agent has committed to the option she addressed in her utterance. It is important whether the partner believes the agent was in a position to deliberate about options, since (according to IRMA) one cannot felicitously make commitments without deliberation.

The partner's ability to fully deliberate about the action the agent has committed to and the partner's commitment constitute the disposition phase, step (1c). If the partner chooses not to commit to a change, the collaborative process must be restarted at either the balance or propose phase (see [15] in Figure 11[21]). On the other hand, if the partner chooses to commit to the proposed change he can either indicate his commitment or allow the agent to infer it. Again the choice will depend on the context and the partner's abilities. The process is then completed for this part of the problem, although the disposition phase at its completion can involve making explicit a commitment that the partner may have just inferred (e.g., [44] in Figure 10). At the conclusion of the disposition phase, the agent and the partner have achieved joint commitment towards that specific option.

The collaborative agent will have to arrive at a joint commitment. However, IRMA's notion of commitment does not take the partner into account. To repair this difficulty, we appeal to the notion of commitment which is *conditional* on the hearer's agreement. An agent is conditionally committed to an option if she is committed to it (in the sense of IRMA) *if* the partner's agreement is secured. A collaborative agent is unconditionally committed if such agreement has been expressed or inferred.

So far, we have illustrated the *Balance-Propose-Dispose* instantiation of the agreement process for our type of collaborative task and information distribution. For different types, the balance phase may be optional. Only the propose and dispose phases would be required in such cases. Participants can interrupt a *Balance-Propose-Dispose* agreement process and put it on hold in order to open a new one, and can return to another interrupted process that deals with a different part of the collaborative problem solving effort. In this sense, our model of the agreement process is recursive; we will show some empirical correlations of this in Section 5.1.2. This assumption of recursion at the domain level is motivated by Clark's view of recursiveness of contributions at the grounding level, and by the widely observed hierarchical nature of discourse.

Before turning to how to further confirm our interpretation of deliberation, we will first discuss an alternative interpretation that seems less plausible. According to this interpretation, the agent always deliberates even if she doesn't know what the partner might have available, and she always commits to changing the problem state in some way and trying to convince the partner to also commit.

First, this does not explain why an agent presents several alternative options in a single utterance, as for example [13] in Figure 3. It is unlikely here that the agent is committing to conflicting alternatives. Instead it seems more plausible that the agent has, as part of the deliberation process, a way of representing open options, and of forming meta-intentions to resolve them.

Second, if we insist that the agent always deliberates to a commitment to a problem

21 Note that we looked at this utterance before from the viewpoint of the agent but are now looking at it from the viewpoint of the partner. Every utterance has two viewpoints.

state change, we must conclude that the agent's commitment to this change is weak. She must be easily willing to override that commitment, because she must be prepared for a rejection of her proposal. However, one of IRMA's theoretical points is that the agent should form strong commitments, not tentative ones that she is willing to easily give up. This interpretation seems to undermine IRMA's rationale for making a commitment: commitments bound the reasoning process.

But the appropriate strength of an IRMA agent's devotion to its commitments has always been open to debate. It is generally agreed that individual intentions may need to be revised in light of new circumstances. As we have seen, collaboration requires a phase in which commitments are only conditional. It may also provide new reasons for withdrawing commitments.

There is also room for differences in the extent to which an agent in a collaborative task will take her partner into account in the course of the deliberation. In the example in Figure 13, the participants have apparently suspended the need to agree on commitments.

```
G-1: [1]: I have $550
     [2]: what about you

D-1: [3]: i have $450.

G-2: [4]: I got a $300 red table and 4 $50 red chairs
     [5]: I have $50 left

D-2: [6]: i bought a $350 yello sofa and $150 yello rug.
     [7]: We have no money left now
     [8]: ok?

G-3: [9]:  I only have two red chairs in my inventory
     [10]: Do you think you can get a blue sofa and two red chairs.
     [11]: I can get a blue rug with the money I have

D-3: [12]: blue sofa $600 i can get,
     [13]: i also can get two red chairs for total of $100
```

Figure 13
An ill-formed agreement process

3.3 Validating the Agreement Process

If our interpretation of the deliberative process is correct, we expect to see a correlation between the phases of the agreement process and the set of utterances that address a particular constraint equation parameter in the CSP.

3.3.1 Theoretical Predictions. Many things can happen during each of the phases of the agreement process. During the balance phase there could be simple beliefs that are shared that have no direct association with a CSP parameter (e.g. how much money the COCONUT agents each have). Also included in the balance phase are any discussions of options that have not been fully deliberated about, which we call *partner decidable options*. For an option to be partner decidable, the agent must believe that all the partner's knowledge preconditions for deciding whether to make a commitment to that option are satisfied, while the agent's own preconditions are unsatisfied. For example, [35] in Figure 10 (*I have a blue sofa for 300*) counts as a *partner decidable option*. If the presented option were group or agent decidable then it would be a proposal. Note: we will call *decision preconditions* those knowledge preconditions that must be satisfied in order for an agent to decide whether to make a commitment to a specific option.

The propose phase can include discussion of an option that has been fully deliberated and discussion of commitment to a particular option. For example, [15] in Figure 11 counts as a *proposal*. Note that there is an inferrable connection to a task action any time a possible parameter value is presented, but only certain specific circumstances, determined by both task and dialogue context, allow one to infer a commitment that is not strictly part of the meaning of the sentence. In Sections 4 and 5 we will discuss some of the circumstances under which such inferences are warranted.

The propose phase may also include evidence of deliberation such as evaluations and options that the proposal has been compared against. We will call the options that aren't committed to, though they are explored in the course of deliberation, *unendorsed* options, e.g. [37]-[38] in Figure 10.

Finally the dispose phase can include evidence of full deliberation and a show of commitment to a proposal by the partner and potentially a follow-up explicit commitment by the agent who proposed the option.

The transitions between phases can be implicit because awareness of the partner and expectations can allow some components of the phases to be inferred. If after a full deliberation, the best option corresponds to an option presented by her partner, the agent can reason that the partner may have had some good reasons to tell her about it. That is, the partner may have done some deliberation but have been unable to make a commitment to the option. The agent could then reason that with some balancing the partner would arrive at the same option as she did and would be willing to commit to it. Also, the partner now expects the agent to propose because she should be in a position to fully deliberate. This should make it easy to infer that the partner will commit to the option. The agent can simply give evidence of deliberation without explicitly proposing, since she can infer the partner will be committed. This happens in Figure 10 (take Jo to be the agent and Ju the partner). Jo re-enters the balance phase with [37-38]. Later in the turn this will be recognized as an *unendorsed option*. The next utterance, [39], is part of the propose phase since it gives evidence of deliberation and makes it more likely that [37-38] is an unendorsed option. This means that [37-38] is a merging of the balance and propose phases. Next we expect a proposal but instead [40] gives evidence of a full deliberation and thus is part of the dispose phase. From this turn and the expectations, Ju should be able to infer that there was a proposal and that Jo believes Ju has committed to this proposal. Unless Ju objects a joint commitment has been established.

The partner (Ju in the above example) can explicitly commit now or simply go on to the next part of the problem. If the partner commits at this point it may be evidence that confirms he was not previously committed to the assignment.

If the partner had intended to commit before he had an opportunity to fully deliberate, then we expect that the partner should be explicit about his commitment to the option, as in [4]-[5] in Figure 13. Otherwise the agent will reason that the partner must not have fully deliberated yet since he probably did not have enough information to do a good job. If the commitment were implicit and missed by the agent and the agent intended to put the agreement process on hold, then the agents could become uncoordinated. Since an agent can infer acceptance of a proposal unless there is evidence otherwise (Walker 1996), if the partner has committed to a change and the agent doesn't recognize that commitment then she might fail to block the partner's default inference of her own commitment. The partner would think that the process was closed while the agent thought it was just on hold.

One last prediction we can make is that if the agent has followed the strategy of considering the partner's knowledge before deliberating to the point of committing, more of these options will be mutually committed to as a joint action. If the agent has not taken the partner's possible contribution into account then mutual commitment will be less likely.

3.3.2 Empirical Testing. So far, we have discussed the agreement process and justified our *Balance-Propose-Dispose* instantiation theoretically. We have substantiated our observations with examples taken from our dialogues, showing that our *Balance-Propose-Dispose* instantiation is plausible and explanatory. In Section 3.3.1, we discussed a number of general predictions that arise from our model. We now turn to more detailed predictions that support our model, and for which we can provide empirical evidence. This is the contribution of our empirical study, and will be the focus of the rest of the paper. The corpus correlations that we present in Section 5 are the first step towards the full mapping we seek to uncover: they concern how the different components of the agreement process correlate with simple notions of context, such as those that can be recognized in a single utterance (a simplified notion of illocutionary force, reference relations, certain types of utterance subject matter) and the current state of problem solving. Although these simpler features are very often determined on the basis of at least part of the discourse history, we leave a detailed empirical study of the features that characterize this larger context for future work.

As a starting point in providing evidence for our *Balance-Propose-Dispose* instantiation of the agreement process, we need to recognize the boundaries of each agreement process. In this way, we can assign each utterance to the relevant agreement processes. This is necessary for empirical testing. Moreover, the end of the process reflects decision points, which we would like to be able to identify automatically.

Once we identify the different agreement processes, we can investigate the distinction between partner decidable options, unendorsed options and proposals, and so indirectly investigate the distinction between the *Balance* phase, to which partner decidable options belongs, and the *Propose* phase, to which most unendorsed options and all proposals belong. Features of the context will provide evidence for this distinction. We assume that for both participants, context is partly determined by the domain reasoning situation, in addition to the preceding dialogue. For instance, if the suitable courses of action are highly limited, this will make an utterance more likely to be treated as a proposal, whereas if the suitable courses of action are not yet limited, this will make an utterance more likely to be treated as a partner decidable option. This correlation is indeed supported by our corpus analysis, as we will show in Section 5.1.3.

Our most interesting result has to do with the *dispose* phase. In fact, originally we had expected to model this phase in terms of agreement proper, namely, in terms of *acceptance* and *rejection*, like many other researchers (Sidner 1994; Walker 1996; Chu-Carroll and Carberry 1998). However, the inability to reliably annotate our corpus for *acceptance* and *rejection* (see Section 4.3) forced us to look for other correlates of the dispose phase. We have found that tracing how the agents' commitment towards the presented options unfolds and changes helps trace negotiation more effectively than accept/reject, and is empirically testable. As we will show in Section 5, commitment in collaborative problem solving dialogues evolves through negotiation: normally, commitment to a certain action starts as tentative or conditional (in fact, the speaker may even present the option as a good one, but show he is unable to commit yet, as with partner decidable options), and may become absolute or unconditional, according to the outcome of subsequent negotiation.

There is another source of evidence for the *dispose* phase. A proposal represents a state of the dialogue in which an explicit *disposition of a proposal* is expected. This is because the agent is expected to be in a position to decide whether to commit to the option. Instead, it would be surprising for a *partner decidable option* to be followed directly by a *disposition* since some decision preconditions are missing. For that option to become part of the solution, further balancing of information is necessary. As we will see, the dialogue does unfold differently in response to a *partner decidable option* and a *proposal*.

We see the major contribution of our empirical study as follows. First, as far as we know, it is the first study that attempts to find correlations of all phases of the purported agreement process in corpus data. In this respect, our study is more systematic than previous work by computational linguists. Sidner, for instance, does not empirically support her claim that sequences of proposal / acceptance and proposal / rejection are the most typical discourse correlates of negotiation. Walker and Chu-Carroll & Carberry, on the other hand, provide empirical investigations of utterances that accept or reject, but not of what is accepted or rejected by them.

3.4 Related Work: Modeling Collaborative Problem Solving Dialogues

To summarize so far, we have presented our view on how IRMA can be instantiated for our task, and how it can be extended to account for collaboration in dialogue by taking Clark's model into account. On the basis of this, we have shown that the agreement process can be seen as the result of combining IRMA, the basic present/accept mechanism used to establish the mutual belief needed for an agreement to be established, and an ability to reason about the possible contributions of a collaborating partner. We also discussed some general predictions we can derive from our model. Before detailing how we are going to verify some of our general predictions, we describe related computational work on collaborative problem solving dialogues. We will show that, like our specific *Balance-Propose-Dispose*, most of the work that follows can be seen as specific instantiations of the general agreement process we have delineated. Any cycle that spells out the components of the agreement process, even if indirectly, can be seen as a compilation of some of these different factors as applied to a specific type of dialogue, in terms of e.g. different distributions of knowledge as we discussed in Section 2. Such compilations may be useful from the point of view of empirical analysis and / or implementation.

Many researchers have explored what we call the agreement process in collaborative dialogues. Whereas some researchers have taken grounding into account (Novick and Ward 1993; Traum 1994; Heeman and Hirst 1995), most of them, like us, have focused on level 4 in Clark's ladder of joint actions, i.e. *proposals* and their disposition in terms of *acceptance* or *rejection* (Ramshaw 1991; Lambert and Carberry 1992; Sidner 1994; Walker 1993; Walker 1996; Chu-Carroll and Carberry 1998).

Some of the work just mentioned models negotiation by means of discourse planners that represent actions at different levels.[22] For example, Lambert and Carberry (1991; 1992) postulate a problem solving level that mediates between a discourse level, which concerns only communicative actions, and a domain level (the discourse and domain levels were first proposed by Litman (1985)). The problem solving level models the process by which two agents build a plan so that one of them can accomplish a certain goal. Ramshaw (1991) also appeals to discourse and domain levels, but adds to them an exploration level. The exploration level partly concerns problem solving, as in Lambert and Carberry's model, but highlights the exploration of alternative plans and actions. In terms of IRMA, we might interpret domain reasoning as referring to the means-end reasoner. It is not exactly clear where Lambert and Carberry's problem solving level fits with respect to IRMA: perhaps it partly concerns deliberation and partly the reasoner that updates the agent's beliefs, as one of their problem-solving operators models providing values for the

22 These are not the same as Clark's levels.

parameters in the agent's domain plan. Ramshaw's exploration level more directly maps to IRMA, as it appears to concern deliberation, but also the opportunity analyzer and possibly the filtering process.

Chu-Carroll and Carberry (1998) build on Lambert and Carberry's work to provide a model of cooperative response generation in collaborative problem solving dialogues. Chu-Carroll and Carberry propose a recursive *Propose-Evaluate-Modify* cycle as their general framework:

> We view collaborative planning as agent A *proposing* a set of actions and beliefs to be added to the shared plan being developed, agent B *evaluating* the proposal based on his private beliefs to determine whether or not to accept the proposal, and, if not, agent B proposing a set of *modifications* to the original proposal. Notice that this model is a recursive one in that the modification process itself contains a full collaboration Cycle—agent B's proposed modifications will again be evaluated by A, and if conflicts arise, A may propose modifications to the previously proposed modifications.

Within this framework, Chu-Carroll and Carberry focus on modeling *information-sharing* subdialogues and *collaborative negotiation* subdialogues that may arise while realizing the abstract evaluate and modify steps. Information-sharing subdialogues are initiated by an agent when she has to evaluate the proposal made by the other agent, but she realizes that she does not have sufficient information to do so. Collaborative negotiation subdialogues are initiated when an agent detects a conflict between the agents' beliefs with respect to a proposal, i.e., when the agent evaluating the proposal holds beliefs that would cause her to reject it.

We view Chu-Carroll and Carberry's collaborative cycle as a direct instantiation of the agreement process, tuned to the specifics of their corpus and of their interests. For example, Chu-Carroll and Carberry's corpus is composed mainly of advice- and information-seeking dialogues, in which different kinds of knowledge reside with each partner (a college counselor and a student, a travel agent and a customer, see Section 2.3.1). We assume they don't need a separate *Balance* phase as we do, because the different information distribution de-emphasizes balancing except at the top-most level of abstraction in the dialogue (we will speculate more about this in Section 6). On the other hand, they make explicit the *Evaluate* and *Modify* phases. This is because they focus on how an agent's beliefs affect her behavior in dialogue; they are also interested in explicitly modeling modifications of the original proposal. We instead follow Clark in considering *refashioning* (Clark and Wilkes-Gibbs 1986) as part of the *dispose* phase. Moreover, for us *evaluation* may be part of both the *propose* phase (via *unendorsed options*) and the *dispose* phase.

Heeman and Hirst (1995) model collaboration on referring expressions by means of a *Present-Judge-Refashion* cycle, that, contrary to Chu-Carroll and Carberry's *Propose-Eval-Modify* cycle, is directly inspired by Clark's presentation/acceptance phases. Heeman and Hirst make use of two levels (or tiers), the *planning* tier and the *collaborative* tier. The planning tier accounts for how utterances are both interpreted and generated; the collaborative tier accounts for the collaborative behavior of agents by providing a link between the mental state of the agent and the planning process. Although Heeman and Hirst don't directly cast their model in terms of IRMA, some of their observations on how the collaborative tier interacts with the planning tier resemble our observations on how the single agent processes in IRMA are related to the collaborative activity. We feel that in our application of IRMA to collaboration we have gone one step further than Heeman and Hirst, as our resulting model is more detailed, and directly anchored to a well accepted architecture for rational agency.

Researchers who attempt implementing Clark's ideas more directly than Heeman and Hirst include Novick and collaborators (Novick and Ward 1993; Novick et al. 1998), and Traum (1994).

(Novick and Ward 1993) presents a model of multiparty dialogue in air traffic control that takes into account not only interlocutors, but also overhearers (Schober and Clark 1989). (Novick et al. 1998) examines the meaning and the advantages and disadvantage of using Clark's model of co-presence and his levels of acceptance as models for cooperative, interactive systems.

Traum's model of conversation acts (Traum 1994) sees acts at different levels corresponding to Clark's ladder, for example, grounding and core speech acts, the latter corresponding to the usual notion of illocutionary acts. Interestingly, he sees each act as putting the hearer under some discourse obligation, and he suggests that such obligations should be integrated into IRMA (Traum and Allen 1994)).

The researcher who most directly relates her work to IRMA is Walker (1993; 1995; 1996). She explores informational redundancy and resource bounds in dialogues by means of empirical analysis of collaborative dialogues, and computer simulations of discourse strategies. The architecture of her computer testbed is a modified version of IRMA that makes explicit agents' resource bounds in terms of their memory limitations. On the basis of this architecture, she models the basic dialogue structure the agents are engaged in as a recursive process, in which one agent proposes options to the other agent; the second agent may ask for clarifications on the proposal, accept it, possibly implicitly, or reject it. As far as we know, Walker is the only researcher who draws an explicit link between IRMA and language; however, her mapping isn't as detailed as ours, even if she explores *acceptances* and *rejections* in human-human dialogues (Walker 1996). She notes that those are an important means by which conversants remain coordinated on what is in the *common ground* (Clark and Marshall 1981; Thomason 1990), and she explores some empirical correlations that help deal with difficult cases.

Chu-Carroll & Carberry's and Walker's accounts are based on empirical analysis. However, in both studies, only one coder coded the data. We believe that a single coder analysis of this sort might be problematic, as it appears that accepts / rejects are difficult to identify consistently, see Section 4.3.

The *SharedPlans* approach develops a multi-agent planning framework and uses this to study the effects of collaboration on discourse and discourse structure. While *SharedPlans*(Grosz and Sidner 1990; Lochbaum 1994; Lochbaum 1995; Grosz and Kraus 1996) does not emphasize an explicit agreement process, Grosz and Sidner (1990) do point out the necessity of a present/accept pairing as part of achieving the mutual beliefs underlying a *SharedPlan*.

Because the *SharedPlans* approach emphasizes the participants intentions and agrees with Bratman's definition of an SCA, it honors the spirit of IRMA and provides us with some of the details needed to implement language actions in IRMA. In fact, Lochbaum points out in (1994; 1995) that IRMA options and potential intentions in *SharedPlans* are equivalent.

Although our current approach does not make use of a multi-agent planning formalism, our account is broadly compatible with *SharedPlans*. On both approaches, the agreement process can be regarded as a discourse mechanism for coordinating multi-agent intentions that need to be agreed upon, and that sometimes need to be negotiated.[23] And

23 Note however that the *SharedPlans* approach does not include an explicit notion of an agreement process. The basic computational component used in that framework is that of *plan augmentation* (Lochbaum 1994), that does not differentiate among different phases within the process.

for both approaches, intention recognition is central in determining the relation between discourse units and the changes that they effect on an evolving collaboration. There are, however, some differences, due to the genres on which we have concentrated and to different theoretical emphases.

Most of the *SharedPlans* research concentrates on group tasks that decompose into separate but coordinated individual plans; (Lochbaum, Grosz, and Sidner 1990), for instance, deals with a joint cooking project in which the interlocutors share responsibility for a meal they are planning. Our discourse task did not ask the participants to plan how the furniture would be purchased, to decide which participant would buy which furniture item; neither is the solution affected by the temporal order in which furniture items are bought. The exchanges that we collected therefore concentrated on information exchange and on negotiation of broad goals. There is little or no domain reasoning that could properly be called planning in our task.

Within the framework of *SharedPlans*, Sidner (1992; 1994) follows a different approach. She focuses on an artificial language to model sequences proposal / acceptance and proposal / rejection, that she claims are *the most typical characterization of negotiation in discourse* (Sidner 1992). In recent work, Rich and Sidner (1997) build a collaborative interface by integrating her negotiation language with the SharedPlans approach.

4. Coding Scheme

In the previous sections we provided a number of informal observations with regard to the macro and micro structure of our dialogues. We turn now to our corpus study. This allows us to flesh out some of these observations by uncovering correlations between illocutionary form, highly limited problem solving alternatives, and the predictions just discussed.

Two coders coded 9 of the 24 dialogues we collected, for a total of 482 coded utterances. The coders were expert linguists and computational linguists. Although our conclusions are based on an analysis of just 37% of our corpus, we believe they are warranted because:

- First, as in any data driven enterprise, data used for development (training set) should not be included in the coded data (test set). We mainly used three dialogues to develop our coding scheme,[24] this leaves 21 dialogues total for coding. [25]
- We had twelve pairs of participants. As we mentioned, they solved between 1 and 3 scenarios: since each pair's session was limited to two hours, the number of scenarios solved varied according to how difficult they found it to come up with a solution and to the complexity of their conversations. It is important to note that everybody solved the same scenarios, and in the same order. Thus, all 12 pairs solved the first scenario, 9 solved the second as well, and only three had enough time to solve the third. As manual coding is highly labor intensive, we realized we couldn't code the whole corpus. To code a representative subset of dialogues, we tried to code data from as many pairs as possible, to make our results independent from possible individual idiosyncrasies. As everybody solved the first scenario, we started from the 12 dialogues concerning the first

24 In some very early trials we also used three other dialogues that belong to the coded corpus.
25 The 12 trial dialogues were also disqualified.

scenario, of which three had been used as a training set. For reasons that are too long to explain, the nine coded dialogues ended up including 8 of the 9 first scenario dialogues and one of the second scenario dialogues. Note that although the 9 coded dialogues represent 43% of the 21 dialogues left after development, they actually account for 51% of the utterances. In fact, on average the dialogues for the first scenario comprise 51 utterances, the ones for the second scenario 42, and for the third only 24. As the first scenario is less constrained than the other two, it presumably leaves more space for negotiation. Intuitively, one would think that looking at longer dialogues is better from our point of view, as they potentially contain more instances of the agreement process, or at least more negotiation.

- At this point in time, there is no consensus on how much coded data is necessary to obtain meaningful statistics; and given the time consuming nature of coding for discourse features, it is common to report results based on an analysis of a subset of a corpus. For example, Core and Allen (1997) report on an experiment in which they coded 604 utterances out of the 114 dialogues of the TRAINS 91-93 corpus (see (Gross, Allen, and Traum 1993; Heeman and Allen 1995)). We will also informally note that, after having coded 5-6 dialogues, Kappa values appeared to stabilize.

We coded for two aspects of the conversations we collected: the dialogue features proper, and the domain reasoning situation. As the reader will notice, the former is far more complex than the latter. Whereas many different linguistic aspects of the conversation are potentially relevant for discovering the correlations we are interested in, the potentially relevant aspects of the domain reasoning situation are sparse. Although dialogue and problem solving features encapsulate different aspects of the context, we will see that one important aspect of our task, namely, partiality of information, affects both, via the definitions of *general/specific* actions and of *solution size* respectively.

The reader may wonder why we do not code directly for the theoretical categories we previously introduced, such as partner decidable options and proposals or the phases of the agreement process. First, we expect it would be extremely difficult for coders to reliably code such categories: they are fairly complex, as they depend on various facets of context, and thus very difficult to define. Instead, the tags we use in our study have simpler definitions, simple enough so that it is not too far fetched to envision a computer system that perhaps through training could reliably recognize at least some of them. Second, we must contend with implicitness in discourse which can make it more difficult to reliably code. In fact, the coder may recognize that a certain inference has been drawn, such as that a furniture item has been proposed, without being able to unambiguously pinpoint the utterances to which the inference should be related (for examples of inferences see Section 3.3.1).

4.1 Coding for Dialogue Features

We designed this part of our coding scheme to conform with the standards developed within the Discourse Resource Initiative (DRI).[26] The DRI is a cooperative response to the recently increased interest in developing tagging resources appropriate for discourse modeling (Passonneau 1994; Nakatani et al. 1995; Moser, Moore, and Glendening 1996; Carletta et al. 1997). DRI has produced a draft annotation scheme called DAMSL (DAMSL 1997).

26 See http://www.georgetown.edu/luperfoy/Discourse-Treebank/dri-home.html.

Two dimensions we code for are taken from DAMSL: *Forward-Looking Functions*, that characterize the effect that utterance U_i has on the subsequent dialogue, and that roughly correspond to the classical notion of an *illocutionary act* (Austin 1962; Searle 1965; Searle 1975); *Backward-Looking Functions*, that indicate whether U_i is unsolicited, or provides a response of some sort to a previous U_j or segment. The main modifications we introduced with respect to DAMSL are operationalizations of the tests used to decide whether a tag applies. Moreover, we code for two new dimensions, *Gist* tags, that capture the gist of the utterance in terms of features relevant to problem solving (properties of furniture items, money and points); and *Reference* tags that encode a simple notion of reference relations.

It is probably clear to the reader why we thought that these aspects would be relevant to our problem: the agreement process is described in terms of illocutionary force, i.e., Forward-Looking Functions (primarily the *balance* and *propose* phases), and of providing a response to a previous proposal, i.e., Backward-Looking Functions (the *dispose* phase). As we will show, the *Gist* and *Reference* tags highlight different and simpler aspects of utterances that are relevant to uncovering their functions as well.

Note that each utterance U_i is tagged by limiting lookahead to the current turn. Reasons for this choice, further details about our scheme, and about the specific differences from DAMSL can be found in (Di Eugenio, Jordan, and Pylkkänen 1998). Further details on our work as an assessment of DRI can be found in (Di Eugenio et al. 1998).

4.1.1 Forward-Looking Functions. Forward-Looking Functions capture the effect that U_i has on the subsequent dialogue. As each U_i may achieve many different effects simultaneously, it can be coded along four different dimensions: *Statement, Influence-on-Hearer, Influence-on-Speaker, Other-Forward-Function.* We briefly discuss Statement and Other-Forward-Function, then we concentrate on Influence-on-Hearer and Influence-on-Speaker. These two tags will be the most relevant to the discussion in Section 5.

The primary purpose of *Statements* is "to make claims about the world".[27] To operationalize the notion of "making claims about the world", we provide the following test. U_i is a *Statement* if it is a declarative sentence that is

- past; or
- non past, and containing a stative verb; or
- non past, and containing a non-stative verb in which the implied action:
  - does *not* require agreement in the domain;
  - or is supplying agreement.

For example, *We could start in the living room* is not tagged as a statement if it is meant as a suggestion, i.e. if it requires agreement; it is tagged as a statement if it is meant as *according to the rules of the game, we're allowed to start in the living room.* The latter case does not require agreement, as that fact is known to both participants.

Other-forward-function acts do not form a natural category, but are grouped together because of their relative rarity. They include conventional conversational acts such as greetings, explicit performatives, and exclamations.

27 *Statements* are categorized into *Assert*, used when S is trying to change H's beliefs, and *Reassert*, used if the claim has already been made in the dialogue. We will not comment on this distinction in this paper.

Influence-on-Hearer and Influence-on-Speaker. The primary purpose of a U_i tagged along the *Influence-on-Hearer* dimension is to influence H's future action, whereas a U_i tagged along the *Influence-on-Speaker* dimension potentially commits S to some future course of action.

Given the nature of problem solving in our domain, the vast majority of actions by our agents are joint, even when the surface form of the utterances is not, as in *I will order that blue sofa*. For the moment, we follow DAMSL in considering *joint* actions as decomposable into independent *Influence-on-Hearer/Speaker* dimensions, even if this may create problems (Tuomela 1995). Thus, in practice in our corpus a U_i tagged along *Influence-on-Hearer* will almost always be tagged along *Influence-on Speaker* as well.

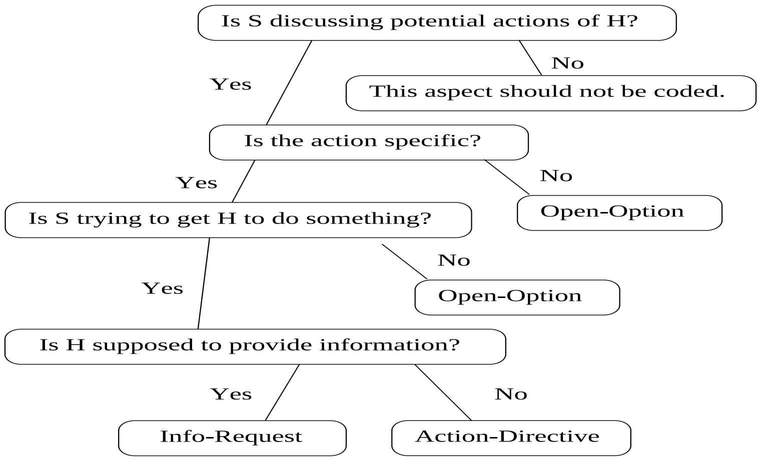

Figure 14
Decision Tree for Influence-on-Listener

Figure 14 shows the decision tree that coders traverse for *Influence-on-Hearer* tags. A distinction is drawn between S merely laying out options for H's future actions (*Open-Option*), and S putting H under obligation to act (Traum and Allen 1994). The lack of obligation may derive from S not providing H with enough information to act (*is the action specific?*), or from S not endorsing the act (*is S trying to get H to do something?*). *Info-Request* includes all actions that request information, in both explicit (*How much is your blue table?*) and implicit (*Tell me what color your table is*) forms. All other *Directives*, including imperatives such as *Let's use my red sofa* and questions such as *What about using my red sofa?* are *Action-Directives*.

As concerns *Influence-on-Speaker*, the only distinction is whether the commitment is conditional on H's agreement (*Offer*) or not (*Commit*). Thus, in both cases, S is committed to the action in the IRMA sense, but with an *Offer*, S makes it clear she wants to achieve joint commitment—recall the discussion in Section 3. As we will see, *Offer* and *Commit* will figure prominently in how we trace reaching joint commitment.

Assigning an *Influence-on-Listener* and/or *Influence-on-Speaker* tag depends on whether a *potential action* underlies a certain utterance. We provide a definition for actions in our domain, and heuristics that correlate types of actions with Forward Functions.[28]

There are two types of potential actions in COCONUT; they correspond to *meta-actions* or to *domain actions*. *Meta-actions* underlie utterances that explicitly address

28 Our definition of actions does not apply to *Info-Requests*, as the latter are easy to recognize.

the experimental procedure or the problem solving process, including strategies to follow in the current problem solving scenario, such as *Let's start from the living room.*[29] There are two types of *domain actions*: *put furniture item X in room Y* and *remove furniture item X from room Y*.

As we mentioned above, one reason why S may not put H under obligation to act is that S does not provide H with enough information to do so. We try to approximate this notion by characterizing when an action description is incomplete, namely, we distinguish between *specific* (potential) actions and *general* (potential) actions. As we will see in Section 4.2 on problem solving features, the distinction between general and specific actions captures just one aspect of partiality of information; another aspect will be captured by distinguishing between indeterminate and determinate solution size. In general, a specific action has all necessary parameters specified, a general one does not. We will now give more details about domain actions; meta-action won't be discussed further. For a domain action, necessary parameters are *type of furniture item*, *price* and *room*; whether *color* is necessary depends on context. Note that the parameter *room* should not be taken as set by default: whereas *room* is indeed instantiated by default to *living room* for *sofa* and to *dining room* for *table* and *chairs*, no such 1:1 correspondence exists for the other furniture items.

General actions arise in two ways: either because not all necessary parameters are set, as in *I have a blue sofa* uttered in a null context (the price is missing), or because the action is an abstraction of different choices that S may list in a single U_i, as in *I have a red sofa for 150 or a blue one for 200.*

In general, coders are encouraged to see declarative utterances regarding furniture items such as the ones just mentioned as referring to actions in the domain, i.e., to answer the question at the root of the tree in Figure 14 positively. This coding convention was adopted to make coders' decisions easier to take. Further, the manual provides examples of specific cancelling contexts, such as negative statements like *I don' t have a blue sofa*.

If the coder recognizes that there is a potential action underlying the current utterance U_i, s/he has to take the next decision in Figure 14, namely, whether the action is specific. If not, the action is tagged as *Open-Option*, and it is not tagged along the *Influence-on-Speaker* dimension.

If the action is specific, a third decision has to be taken: *is S trying to get H to do something?*, namely, is S endorsing the option for action presented to H. It is hard to devise a comprehensive test for this, but some clear special cases can be isolated. For instance, S may refer to one action that the participants could undertake, but in the same turn S may make it clear that that action is not to be performed. This happens in excerpt (10) in Section 3, repeated at the end of this section in Figure 15 alongside some of it tags. A specific action (*get Jo's yellow sofa for 350*) underlies [38], which thus would qualify as an *Action-Directive*, just like [35]. However, because of [40], it is clear that Jo is not offering his yellow sofa as part of the solution, thus, Jo is not endorsing using his own sofa. [38] is therefore tagged as an *Open-Option*.

Whereas examples like the one we just discussed make it clear that the speaker is not endorsing the option for action, there are other cases in which there is no evidence to show whether the speaker endorses it or not. This happens, for example, when an option for action arises in the answer to a question, and the respondent does not express any opinion with respect to the option in question (as [31] in Figure 19, Section 5). Such cases are tagged as *Open-Option* as well.

29 Meta-actions are fairly consistently labeled with the *Strategize-Action* tag along the *Information-Level* dimension, which we do not have space to discuss here.

Intuitively, the cases tagged as *Open-Option* because of lack of explicit endorsement should not be coded along the *Influence-on-Speaker* dimension. After all, if the speaker is not endorsing that option, she is not potentially committing to the corresponding action. However, since certain contexts, such as answers, are in a sense "neutral" with respect to lack of endorsement, and because of lack of clarity in the coding scheme, some of these *Open-Option*'s have indeed been coded along the *Influence-on-Speaker* dimension, always as *Offer*.

To summarize, if the coder recognizes that U_i refers to a joint action, then the following heuristics apply:

- If U_i refers to a *general* action it is tagged as *Open-Option*, and is not tagged along *Influence-on-Speaker*.
- If U_i refers to a *specific action* but there is a lack of explicit endorsement, it is tagged as *Open-Option*. Whether it is tagged along *Influence-on-Speaker* depends on context.
- In all other cases, U_i is tagged as *Action-Directive*, as long as U_i does not qualify as *Info-Request*. U_i is also tagged along *Influence-on-Speaker*.

4.1.2 Backward Looking Functions. Backward-Looking Functions capture (part of) the relationship between an utterance or group of utterances $\{U_i\}$ and the previous discourse: namely, $\{U_i\}$ may be unsolicited, or respond to a previous U_j or segment.[30] Tags in this dimension categorize as follows. **Answer** is used when U_i answers a question. **Understanding** tags are used to tag *Acknowledgements* and the like. We won't discuss the Understanding tags, as they occur very rarely in our corpus (5 occurrences out of 482 utterances). **Agreement** tags are used when U_i expresses S's attitude towards a belief or option for action embodied in its antecedent, and include:

- *Accept (Accept-part)*: U_i accepts (part of) the content of its antecedent, e.g. [19] in Figure 7: *The lamp and chairs sound good.*
- *Reject (Reject-part)*: U_i rejects (part of) the content of its antecedent, e.g. [20] in Figure 7: *but the chairs seem expensive.*
- U_i performs a *Hold* if it does not express an attitude towards its antecedent, but leaves the decision open pending further discussion. For example the segment from [18] to [20] in Figure 3 qualifies as a hold, because it doesn't directly address Ju's option for action in [16].

Any U_i's coded with one of these tags is also annotated with an explicit link to its antecedent.

4.1.3 Gist. These tags capture (part of) the meaning of the utterance by encoding what is relevant to problem solving in terms of money, points or furniture items. Possible dimensions are:

- *budget* related tags: U_i discusses the initial budget (*budgetAmount* tag) or budgetary consequences (*budgetRemains* tag: *That will leave us with $500*);
- *point* related tags: U_i discusses the amount of points associated with a solution (*This would give us a score of 670*).

30 Space constraints prevent discussion of segments, both as antecedents and as responses.

- *furniture* related tags:
  - *haveItem*: S states that she has a particular item (*I have a blue sofa for $300*);
  - *elaborateItem*: S elaborates the description of an already introduced item (*my red chairs are $100 each*);
  - *getItem*: S discusses selecting a particular item (*shall we buy the two red chairs*);
  - *otherItem*: U_i concerns item(s) of furniture, but none of the other *x-Item* tags applies, as in *so I say let's start with the sofa*.

- *evaluation* related tags: S evaluates a specific furniture item (*the chairs seem expensive*), or a solution (*I like this plan you have suggested*); sometimes the evaluation is expressed by comparing the relevant features of two items, as in [8] in Figure 4 (*My sofa's are more expensive*). Note that *evaluation* related tags apply only when the utterance has an *explicit* evaluative connotation. Just mentioning the consequences of a certain choice, even if seen as positive as in *This combo will give us 200 [points] for the sofa and 200 for the table*, or as negative as in *We can't buy that sofa, we'd go over the budget*, is no grounds for using an *evaluation* tag.

Each *Gist* tag closely reflects the surface form of the utterance. Coders are instructed to infer one of these tags if they don't explicitly appear in surface form if the corresponding verb (*have, get*) can be either substituted or inserted in the utterance (if e.g. the utterance is elliptical). So, for example, *buy the chairs* will always be tagged as *getItem*; however, *I have a sofa for $300* will never be tagged as *getItem*, always as *haveItem*, independently from its Forward Function. This also explains why certain utterances, such as *so let's begin in the dining room* are tagged with a *nil* gist tag: none of our predefined gists applies. Finally, note that more than one gist tag can apply to the same utterance.

4.1.4 Reference Relations. These embody a simple notion of reference relations, i.e., they capture how furniture items discussed in one utterance are related to those previously discussed. *SameItem* is used when U_i is related to its antecedent via the same item or set of items. *Subset* is used if U_i discusses a subset of the items in its antecedent. The tag *MutuallyExclusive* is used when U_1 mentions a set of items $\mathcal{S}_1$, U_2 provides an alternative $\mathcal{S}_2$ to that same set of items, and $\mathcal{S}_1$ and $\mathcal{S}_2$ are mutually exclusive. As we will see, *MutuallyExclusive* characterizes some proposals that are not committed to. A *Reference* tag explicitly points to its antecedent.

4.2 Coding for Problem Solving Features

Recall that our assumption regarding how domain reasoning affects negotiation concerns whether the branch factor of the problem solving situation is large (see Section 3.3.2). We now show how we code the data for this aspect in terms of the solution size (or number of options in terms of IRMA) for a set of constraint equations.

We view the problem space as a set of constraint equation parameters $\{parm_i\}$ that must have a single value or a set of values of a certain cardinality assigned to them for a solution to exist. The main parameters of interest for our corpus are the objects of type t in the goal to put an object in a room (e.g. $parm_{sofa}$, $parm_{table}$ or $parm_{chairs}$). For a solution to exist to the set of constraint equations, each $parm_i$ in the set of equations must have a solution, and the cardinality of the assigned value set must match the cardinality designated for it. For example, if $parm_{chair}$ has a designated cardinality of 3, and a value

Forward Functions				Backward Functions	
Statement	Listener	Speaker	Other	Answer	Agreement
.83	.72	.72	.93	.79	.54

Table 1
Kappa values for Forward and Backward Functions

set of cardinality 4 (i.e. there are 4 instances of chairs that are known), but no more than 2 instances of chairs can ever be assigned without violating the budget constraint, then $parm_{chair}$ is unsolvable. No solution will be found in cases where all combinations of value assignments to some parameter of interest violate some constraint. Just one unsolvable parameter is all that is needed to render the problem unsolvable.

Based on this view, we code each utterance that relates implicitly or explicitly to the problem solving parameters with which parameters are addressed and the solution size for the set of constraint equations that are related to those parameters. Note that coding for problem solving parameters is independent from coding for Reference relations. By coding which constrained parameters are addressed, we can identify the different agreement processes taking place in the dialogue given our assumption that, in general, each parameter requires a separate agreement process.

As regards solution size for a set of constraint equations, we characterize it as determinate if there is one or more solutions for a set of constrained parameters. It is important to note that the set of possible values for each $parm_i$ is not known at the outset since this information must be exchanged during the interaction. If there is no solution to the problem or the value set for some $parm_i$ is open we characterize it as indeterminate.[31] A value set is open, e.g. if S supplies appropriate values for $parm_i$ but does not know what H has available for it. A value set for a certain $parm_i$ can be reopened, and thus, solution size may revert from determinate to indeterminate, e.g. if S asks what else H has available for a closed $parm_i$. The value *indeterminate* for solution size in situations in which the value set for some $parm_i$ is open captures an aspect of partiality of information different from the notion of *general* action discussed earlier. A general action, i.e., an incomplete action description, is local to the utterance U_i that contains the action description, independent of the previous context. Instead, an indeterminate solution size due to the value set for a parameter $parm_i$ being open reflects that not enough information has been exchanged in the dialogue preceding U_i regarding that parameter.

4.3 Reliability of the Coding Scheme

Tables 1 and 2 report values for the Kappa coefficient of agreement (Krippendorff 1980; Carletta 1996), which factors out chance agreement between coders. Recall that we have 9 dialogues exhaustively doubly coded, for a total of 482 utterances. Note that we don't report intercoder reliability measures for coding for solution size and parameters because the coding is straightforward. We ran a pilot study in which we computed Kappa on two dialogues doubly coded for these features; as we obtained values over .8 for both solution size and parameters, only one coder coded the remaining 7 dialogues.

The columns in the tables read as follows: is utterance U_i tagged for tag X, and if yes, do coders agree on the specific subtag? For example, the possible set of values for *Influence-on-Listener* are: NIL (U_i is not tagged along this dimension), Action-Directive,

31 We grouped together no solution and open value set because we were initially only interested in the CSP branching factor. In future work, we may decide to code separately situations in which there is no solution to the problem, as they are likely to correlate with cases in which the participants change their goals.

Gist	Eval Tags	Reference
.86	.82	.84

Table 2
Kappa values for Gist and Reference

Open-Option, and Info-Request. The last two columns probe Backward Functions: was U_i tagged as an answer? was U_i tagged as *accept*ing, *reject*ing, or *hold*ing the same antecedent? Computing Kappa for the backward tags takes into account whether the coders linked U_i to the same antecedent: thus, a situation in which both coders code U_i as *Accept*, but disagree on what antecedent U_i accepts, counts as a disagreement. We report one Kappa value for Gist in general, and one specific to the *evaluation* subtags. We checked the latter because their definitions call on the coder's judgement more than those of the other *Gist* tags; we were pleased to see that we reached excellent agreement in this case as well.

Kappa's possible values are constrained to the interval $[0, 1]$; K=0 means that agreement is not different from chance, and K=1 means perfect agreement. To assess the import of the values $0 < K < 1$ beyond K's statistical significance (all of our K values are significant at p=0.000005, except for *Other-forward-function* at p=0.0005), the discourse processing community uses Krippendorf's scale (1980).[32] Krippendorf's scale discounts any variable with K < .67, and allows tentative conclusions when .67 < K < .8, and definite conclusions when K≥.8. Thus, Table 1 suggests that Forward Functions and Answers can be recognized far more reliably than Agreement functions. This will have consequences for the way we model the *Dispose* phase in Section 5.

There may be various reasons why Agreement tags are less reliable than the others. First, they are much rarer, and this may negatively affect K. As (Grove et al. 1981) pointed out, the low frequency of a tag may change the upper bound for K, that corresponds to perfect agreement, from 1 to a value sometimes much lower than 1. Second, we did not put as much effort into revising the original DAMSL manual for Backward Functions as we did for Forward Functions, because our pilot coding experiments did not highlight any problems with Agreement tags. Also Core and Allen (1997), who used the DAMSL manual without modifications to tag 604 utterances in spoken task-oriented dialogues, reported that agreement tags are unreliable. We refer the reader to (Di Eugenio et al. 1998) for a longer discussion of this issue.

4.4 A Coded Example

We conclude this section with a coded example to help the reader get an informal grasp of the meanings of the tags more relevant to the discussion in Section 5. Figure 15 lists the most important tags for Example (10) from Section 3. We include [41-43] in the dialogue for completeness, but as they refer to a different agreement process, concerning $parm_{chairs}$, we don't include their analysis in Figure 15.

As the reader can see, all utterances have been coded as *statements*, and they all concern $parm_{sofa}$. The other tags deserve more discussion. First of all, according to our conventions, there are action descriptions underlying [35], [37], [38], [40], and [44]. Of these, [35], [38] and [40] refer to *specific* actions (all parameters are known), whereas [37] is *general*, because the color of the sofa is not made explicit till [38]. This explains why [37] is labeled as Open-Option. Also [38] is labeled as Open-Option, but for different

32 More forgiving scales exist, e.g. the one in (Rietveld and van Hout 1993), but have not yet been assessed by the discourse processing community.

```
Ju-4: [35]: i have a blue sofa for 300.
      [36]: it's my cheapest one.

Jo-4: [37]: I have 1 sofa for 350
      [38]: that is yellow
      [39]: which is my cheapest,
      [40]: yours sounds good.

Ju-5: [41]: Ok
      [42]: I logged in 2 of your chairs and 2 of mine.
      [43]: both red.
      [44]: I will order that blue sofa.
```

	Dialogue Features				Problem Solving Features	
	Statement	Listener	Speaker	Agreement	Parameters	Solution size
[35]	yes	Action-Dir	Offer	nil	sofa	indet
[36]	yes	nil	nil	nil	sofa	indet
[37]	yes	Open-Option	nil	nil	sofa	indet
[38]	yes	Open-Option	Offer	nil	sofa	det
[39]	yes	nil	nil	nil	sofa	det
[40]	yes	Action-Dir	Commit	Accept	sofa	det
[44]	yes	Action-Dir	Commit	nil	sofa	det

Figure 15
A dialogue excerpt and its tags

reasons: because of [40], it is clear that Jo is not trying to get Ju to use his own sofa. As regards *Influence-on-Speaker*, [37] is not labeled in this dimension because it is a general, and not specific, action. [35] and [38] are labeled as *Offer*, because the action described is specific, and the hearer's agreement is necessary for a commitment to that potential course of action.[33] [40] and [44] are labeled as *Commit* because the speaker's commitment is not contingent on the hearer's agreement.

As regards solution size for sofas, it stays *indeterminate* till [38], when it changes to *determinate*—at this point, Jo and Ju have exchanged enough information to take a decision about sofas. Note that *determinate* does not mean only one solution is possible, but that the solution size is finite: in fact, at this point two solutions are still possible, Ju's sofa or Jo's sofa.

This example highlights some combinations of tags that we will use in section 5 to show how the agreement process is carried out. Specifically, we will show that *partner decidable options* correspond to utterances tagged with an *Influence-on-Hearer* tag and indeterminate solution size, as with [35] and [37]. *Proposals* will correspond to utterances tagged as Action-Dir+Offer with determinate solution size (none in this example). *Unendorsed Options* will correspond to utterances tagged as *Open-Option* with determinate solution size, such as [38].

Negotiation and agreement will be modeled through changes in commitment. In [40], Jo's expresses his agreement to buying Ju's sofa (presented in [35]) via his unconditional

33 [38] is an example of an *Open-Option* due to lack of endorsement that should probably not have been tagged along the Influence-on-Speaker dimension, see discussion in Section 4.1.1.

commitment. And in [44] Ju expresses her own commitment to buying her blue sofa. Note that [35], as it is indeterminate, will count as a partner decidable option, and thus, will not count as expressing commitment in the IRMA sense, even if it is labeled as an *Offer*. This inconsistency between theory and coding scheme is due to the need to keep coding simple, and to keep coding dialogue features and problem solving independent.

5. Corpus Correlations of the Agreement Process

In Section 3, we discussed how the IRMA model and the characteristics of our design task, in particular the initial distribution of knowledge, allow us to make some predictions with respect to our dialogues. In particular, we claimed: that we should see both *partner decidable options* and *proposals*; that a *proposal* is more likely to refer to an action on which the two partners are going to agree than a *partner decidable option*; that for a *partner decidable option* to become part of the final solution further balancing of information is necessary, whereas a *proposal* represents a state of the dialogue in which a disposition in the form of an explicit evaluation, agreement or rejection is necessary, or at least strongly expected. Moreover, we also hoped we would find correlates of the start and end of each agreement process and in general, that we would be able to trace the negotiation process as it evolves.

We expected to use the tags we coded for to verify these claims. More specifically, we expected to analyze partner decidable options in terms of *Open-Options* (those due to general actions, not to lack of endorsement), proposals in terms of *Action-Directives*, and to trace the results of the negotiation process by means of gist tags of type *evaluation* and of the agreement tags. Given our Kappa results, it is clear that our plan is going to fall short as far as the agreement tags are concerned. As we discussed in Section 4.3, the coding experiments conducted by Core and Allen (1997) yielded the same result, thus it is plausible to conclude that agreement tags are difficult to reliably code for, even if this may be partly due to the coding manual needing further work and/or to their relative rarity. This result thus calls into question the empirical foundations of studies, such as Walker's (1996) and Chu-Carroll and Carberry's (1998), that are based on a single coder's annotation of acceptances and rejections. Since both Walker and Chu-Carroll & Carberry report low frequencies for accept/reject in their respective corpora, it is possible that, if they had doubly coded their corpora, their intercoder reliability scores would also be low.

The issue is how to trace the agreement process, given that what a priori would appear to be the most explicit evidence for it is not available. It was only after we had to face this problem that we realized that explicit agreement is just the most obvious way of tracking the agreement process. Other ways are available: specifically, given that our subjects are negotiating joint potential actions, they indicate their attitudes towards such actions by expressing their commitment towards them. The notion of conditional / unconditional commitment that the tags *Offer* / *Commit* capture, when applied to collaborative problem solving dialogues, can be recast as tracking changes in agents' commitment: commitment to a certain action X starts as conditional and may become unconditional, according to the outcome of negotiation. Moreover, given that an *Offer* expresses a commitment contingent on the hearer's agreement, whereas a *Commit* expresses an unconditional commitment, this distinction partly captures the speaker's view of the state of the negotiation. The pattern of change in commitment on the part of the

	Accept	Reject	Hold
Commit	27	2	0
Others	21	16	9

Table 3
Commits and Agreements

two speakers can be abstracted as follows:

(2a) S_1: Offer (regarding action X)

(2b) S_2: Commit (regarding action X)

(2c) S_1: Commit (regarding action X)

Such a commitment pattern is the most explicit way speakers may carry out their agreement process. Although it rarely occurs in as complete a form as in (2), the pattern in (2) can be considered as the safest way to ensure mutual belief that a commitment to action X has been reached by both partners. Not surprisingly, the step which is most often missing is (2c). The reason it is left out is not redundancy ((2c) is not redundant because it shows that S_1's conditional commitment has become unconditional), but rather, that S_2 can easily infer (2c), given (2a). Note that sometimes (2a) may be null in terms of commitment, i.e., an explicit offer with regard to X may be missing. This happens when S_2 commits to a *partner decidable option* presented by S_1: at the stage in which S_1 presented it, she was unable to commit to it yet, thus we would expect (2c) to be explicit. Another case in which we expect (2c) to be explicit is when S_2 does not express any attitude regarding action X, i.e., when (2b) is null from a commitment point of view. In these cases, if S_1 infers an implicit acceptance on the part of S_2, S_1 can use (2c) as a way of making explicit such an inference, and of providing S_2 with another opportunity of expressing his view regarding action X.

In general, the advantages of commitment over agreement to model negotiation are as follows. First, commitment captures the speaker's evolving attitude towards an option for action better than agreement: the speaker can show she is deliberating, but unable to commit yet, as when she provides a partner decidable option; she can conditionally commit, as when she proposes an option for action; and she can unconditionally commit to an option for action initially provided by the speaker herself or more often by the partner. Instead, accept and reject per se do not capture how the speaker's attitude towards an option for action evolves. Second, the notion of commitment also partially embodies the agent's view of the partner's attitude towards that option for action, whereas accept/reject only embodies one speaker's point of view: acceptance is inherently an attitude that one speaker can express only in response to an option for action presented by the other speaker. This is reflected in our coding conventions, that speakers cannot accept their own proposals (although sometimes they reject them).

Luckily, our intercoder reliability score for *Influence-on-Speaker* is good enough to make these tags usable. Moreover, note that although our coding for *Agreement* is not as reliable, we found a correlation (Table 3) between *Commits* and *Accepts* in one coder's tagged data (χ^2 = 17.7, p<0.001).

5.1 Tracking the Agreement Process

5.1.1 Some General Trends. Table 4 is meant to give the reader an informal and very high level impression of the distribution of forward functions and gist tags in our dialogues. In particular, it highlights that many utterances in our dialogues are not concerned with building the solution in the most direct way. Rather, they are concerned

Stat. only	Open-Option	Action-Dir + Offer	Action-Dir + Commit	Info-Request	Action-Dir only
183	33	63	48	59	2

Budget	Points	HaveItem	ElaborateItem	GetItem	Evaluate	otherItem	nil
69	28	74	39	93	50	12	33

Table 4
Distributions of Forward Functions and Gist Tags

with balancing non-negotiable information, with explicitly evaluating solutions and with exploring the consequences of certain choices in terms of points gained, money spent or goals achieved. This is shown by the high number of utterances that are tagged as Statements *only* or as Info-Requests, and by the high number of occurrences of gist tags pertaining to money, points or evaluation.[34]

Some categories from Table 4 need clarification. First, the column *Stat(ement) only* refers to those utterances whose only Forward-Looking Function tag is *Statement* (e.g., [36] and [39] in Figure 15); as Forward-Looking Functions are not mutually exclusive, some utterances tagged along the *Influence-on-Hearer/Speaker* dimensions are tagged as *Statements* as well. Second, given the nature of our task, almost all actions are joint. This, coupled with the heuristics we presented in Section 4.1, means that almost all actions tagged as *Action-Dir* co-occur with either *Offer* or *Commit*. The two *Action-Dir only* in Table 4 tag two utterances that ask the partner to check their "point" computations. A U_i tagged as *Open-Option* due to a general action description in U_i is not tagged along *Influence-on-Speaker*; a U_i tagged as *Open-Option* due to the speaker not endorsing the presented action may or may not be tagged along *Influence-on-Speaker*, depending on context. In the latter case, if U_i is tagged along *Influence-on-Speaker*, it is always tagged as *Offer*, never as *Commit* ([38] in Figure 15). However, because the coding of the second type of *Open-Option* as *Offer* doesn't appear to be consistent (see Section 4.1.1), we don't consider *Open-Option* subdivided into *Open-Option only* and *Open-Option+Offer*.

Although we have computed cross distributions between gist tags and forward functions, we will not tabulate them here because such a table would require a long explanation. Rather, we will now turn to discussing the recognition of the end of individual agreement processes, how negotiation occurs in the presence of partner decidable options and of proposals, and how commitment unfolds in negotiation.

5.1.2 End of an Agreement Process. We are interested in tracking the end of each process because we want to be able to assign each utterance to one agreement process for empirical testing. Moreover, the end of the process corresponds to decision points, which we would like to be able to identify automatically. Given our definitions of parameter value set and of determinate / indeterminate solution size, one way of recognizing the end of a process is to track when subjects turn to a different parameter to solve, and solution size reverts from determinate to indeterminate: this potentially means that the subjects have reached an agreement on the previous parameter, and are moving to a different part of the problem space. For example, [27] in the excerpt in Figure 16 marks the end of the process about table and chairs. This is because [27] still concerns table and chairs, and solution size is coded as determinate; in [28] there's a change to a new

34 All numbers in this section are based on one coder's tagged data, as allowed by our good reliability results.

parameter (sofa) and solution size reverts to indeterminate (J and K haven't discussed sofas yet). However, this correlation is going to give wrong predictions in cases such as [32]. The same pattern as for [27]-[28] occurs in [31]-[32]: [31] is coded as determinate because J and K have shared all available information about their sofas, and [32] as indeterminate because J and K have never discussed lamps. However, it is clear that no solution regarding sofas has been achieved, thus [31] should not qualify as the end of the process about sofas. As it turns out, the agreement process about lamps is embedded in the one about sofas, and the two processes end together in [42].

```
J-3: [21]: Here is my suggestion.
     [22]: We should buy your green table for 200, my three green chairs
           for 50 a piece and your one green chair for 100.
     [23]: That would be a total of 450,
     [24]: correct?

K-4: [25]: yes, that would be 450,
     [26]: leaving us with 550 to design our living room,
     [27]: that works for me.             [parm=table,chairs;  sol size=det]
     [28]: I have three sofa              [parm=sofa; sol size=indet]
     [29]: a blue one for 400, a green one for 550,
           and a yellow one for 350,      [parm=sofa; sol size=indet]
     [30]: how about you?

J-4: [31]: I have a yellow one for 400, a red for 550, and
           a blue one for 300.                [parm=sofa; sol size=det]
     [32]: What kind of lamp do you have?     [parm=lamp; sol size=indet]

K-5: [33]: i have a blue floor lamp for 250,  [parm=lamp; sol size=det]
     [34]: but we should first decide         [parm=sofa; sol size=det]
           on the sofa,
     [35]: so we know how much we have left to spend.

J-6: [36]: Well,
     [37]: if we buy my blue sofa for 300 and
           your blue lamp for 250,            [parm=sofa,lamp; sol size=det]
     [38]: that would total 550,
     [39]: using all our money
     [40]: and mataching colors.
     [41]: Do you agree?

K-6: [42]: That sounds good to me.
```

Figure 16
Embedded agreement processes

Thus, we looked at other features of pairs of utterances where a transition from determinate to indeterminate solution size occurs, and we arrived at the following rule. An agreement process ends at U_i if U_i has determinate solution size, U_{i+1} indeterminate, and:

- U_i regards $parm_x$ and U_{i+1} $parm_y$;
- U_{i+1} is not tagged with a reference relation, i.e., $parm_y$ has not been discussed yet;
- U_i is a *Commit* regarding $parm_x$, or, if not, there is a preceding U_{i-j} tagged as a *Commit* regarding $parm_x$, and all the utterances between U_{i-j} and U_i are

tagged only as Statements that concern what is left of the budget or the accumulated points.

This rule correctly identifies 12 ends of agreement processes, and correctly rejects 16 cases in which there is a transition from determinate to indeterminate solution size, but in which the current agreement process is kept open, such as [32] in Figure 16. The rule fails 3 times: in two cases, it fails to identify the end of a process, in the remaining one, it predicts the end of a process that is still open. In all cases, it is the third constraint on the rule that fails.

5.1.3 Partner Decidable Options, Proposals and the Agreement Process. In Section 3.3.1, we pointed out that given the nature of our task, and the fact that information is initially private but needs to be shared to reach a solution, we should see proposals occur only in contexts where enough information has been exchanged. As we noted in Section 4, partiality of information is captured in two different circumstances in our coding scheme: via the notions of general action and of indeterminate solution size. As a first hypothesis, we would expect *partner decidable options* to correspond to *Influence-on-Hearer/Influence-on-Speaker* tag pairs co-occurring with indeterminate solution size, and *proposals* to correspond to *Influence-on-Hearer/Influence-on-Speaker* tag pairs co-occurring with determinate solution size. To start exploring this hypothesis, first of all we tabulate all *Influence-on-Hearer/Influence-on-Speaker* pairs with respect to solution size (recall from Section 5.1.1 that we don't distinguish *Open-Options* that don't co-occur with an *Influence-on-Speaker* tag from those co-occurring with *Offer*).

In Table 5 we distinguish between *meta-* and *domain* actions. As discussed in Section 4.1.1, *meta-actions*, such as *Let's start from the living room*, explicitly direct problem solving. In the following, we will only discuss negotiation at the domain level, i.e., at the level of choosing values for the parameters in the associated constraint satisfaction problem. We have identified some preliminary patterns that link meta-actions to the rest of the dialogue. For example, as Table 5 shows, meta-actions are very often indeterminate Action-Dir+Offer; moreover, they are implicitly accepted, i.e., the dialogue proceeds according to the strategy advocated by the meta-action, without any negotiation in this regard. A full account of our dialogues that explains meta-actions as well is left for future work.

If we just consider domain actions, Table 5 confirms there is a correlation between *Influence-on-Hearer / Influence-on Speaker* tag pairs, and solution size ($\chi^2 = 17.58$, $p < 0.001$).

	Indeterminate		Determinate	
	Domain	Meta	Domain	Meta
Open-Option	15	1	17	0
Action-Dir + Offer	14	9	37	3
Action-Dir + Commit	3	0	45	0
Total	32	10	99	3

Table 5
Forward Functions and solution size

If we now abstract away from the specific tags, and correlate bundles of tags with conceptual stages in the agreement process, we can postulate the following conceptualizations.

- *Partner Decidable Options.* We claim they correspond to either Open-Option or Action-Dir+Offer, both with indeterminate solution size. Namely, partner decidable options are references to actions that occur in a context in which not enough information has been shared for the agent's full deliberation to take place. Note that Open-Option and Action-Dir+Offer with indeterminate solution size differ in terms of partiality of information. In both cases, the fact that U_i is marked with *indeterminate* solution size for $parm_j$ indicates that not enough information regarding $parm_j$ has been exchanged in the dialogue preceding U_i for a decision to be made. However, the label Open-Option for U_i indicates another source of partiality of information, a *general*, i.e., incomplete action description in U_i.

The alert reader may have noticed an inconsistency in considering utterances tagged as Action-Dir+Offer with indeterminate solution size as partner decidable options. Namely, when we defined partner decidable options in Section 3, we mentioned that they occur when the agent is unable to commit to the presented option. As a consequence, no utterance tagged as *Offer* should qualify as a partner decidable option, given that, as we mentioned in Section 4.1.1, *Offer* does encode commitment in the IRMA sense, even if conditional on the partner's agreement. In retrospect, one could say that these utterances should not have been tagged as *Offer*, and perhaps not even as *Action-Dir*, but simply as *Open-Option*. However, this would have required yet another heuristics to be added to the coding manual, and this heuristic would have required the coder to take into account the state of the problem solving in addition to the state of the dialogue. We prefer to keep definitions for tags as simple as possible, and when a theoretically incorrect label is assigned, to disregard it in favor of other features we coded for. The only case in which this happens is the one under discussion, in which we disregard the incorrect commitment label in favor of solution size. Note that it is necessary that U_i is labeled with an *Influence-on-Hearer* or *Influence-on-Speaker* tag to count as a partner decidable option: indeterminate solution size is not sufficient per se, as U_i must refer to a domain action.

- *Unendorsed Options* are Open-Options with determinate solution size. As we discussed in Section 4.1, an Open-Option will co-occur with determinate solution size when the action described is specific, but the speaker appears not to endorse the presented option (as in [38] in Figure 15). Whereas presenting *unendorsed options* could seem unnecessary in terms of the IRMA architecture, it makes sense because it can satisfy at least two potential goals at the same time. First, an unendorsed option provides evidence that the agent did deliberate. Second, it balances information in anticipation of the interdependencies of the CSP parameters: it may be advantageous to know which options were close contenders if later on the agent has to do backtracking.

- *Proposals* correspond to utterances tagged as Action-Dir+Offer with determinate solution size.

- *Unconditional Commitments* correspond to utterances tagged as Action-Dir+Commit. In this case we don't distinguish whether the associated solution size is determinate or indeterminate, because *Commit* should occur only when the solution size is determinate: by definition, a certain parameter is solved only when the solution size for that specific parameter is closed, so that is the only occasion in which subjects can commit to a solution. In fact, the 3 *Commits* that occur with indeterminate solution size (cf. Table 5) are

> ill-formed: they correspond to utterances [4], [6] and [12] in Figure 13, Section 3. The tagging reflects the fact that the two subjects start buying items in [4] and [6] without being in a position to fully deliberate. The two turns G-2 and D-2 appear to be intended to both balance and commit to a proposal, without considering the necessary knowledge preconditions for deliberation (note that the preconditions for deliberation call for a balancing of the information distribution). In G-3, after G realizes he doesn't have as many red chairs as he thought he had, backtracking occurs and a more "standard" process begins: in fact, G does not have enough information to solve the subproblem on his own and has to ask for information, to which D answers with [12]-[13]. The dialogue could of course have ended with G-3, if G had said "OK, we're done", but given G realizes his mistake, the dialogue continues for 12 more turns. Although [12] belongs to the part of the dialogue that more closely conforms to the agreement process, it is ill-formed all the same, because the two subjects have not discussed their options with respect to sofas.[35]

Table 6 recasts the utterances regarding domain actions from Table 5 in terms of the categories we just discussed.

Partner decidable options	29
Unendorsed options	17
Proposals	37
Unconditional Commitments	48
Total	131

Table 6
Dialogue functions

We now have to provide evidence for the validity of our categorization. We do so by analyzing the antecedents of *Commit*s and by analyzing how the dialogue develops in certain cases. We expect to see that most of the antecedents of Commit will be Proposals. We also expect to see some partner decidable options and possibly some unendorsed options as antecedents of Commit: however, the dialogue between these types of antecedents and the corresponding Commit should evolve differently than in the case of proposals. Finally, we also expect to see some unconditional commitments as antecedents of other commitments, because of the commitment pattern we presented in (2). In this case, we expect the two unconditional commitments to be uttered by two different speakers.

Table 7 lists the antecedents of Commit in our data, in terms of the notions of *partner decidable options*, *proposals* etc. that we just discussed. It also highlights whether the antecedent to which the agent commits was presented by the agent himself or by the partner. Not surprisingly, 82% of the times the agent commits to something presented by the partner.

As antecedents of *Commits* are not tagged, we reconstructed them by exploiting the *parameter* tagging or via the antecedent of the Accept tag, if the utterance is tagged as both Commit and Accept.[36] Note that Table 7 lists 60 antecedents for *Commit* although

35 If the reader is puzzled by the fact that [12] is coded as a *Commit*, s/he should remember that utterances tagged with *Commit* describe an action by the speaker that is not conditional on the hearer's agreement: clearly *getting the blue sofa* is not conditional on G's agreement, given G suggested it. Moreover, *Commit* is used for all unconditional commitments, independent from their strengths.

36 Recognizing that a *Commit* has an antecedent calls into question the fact that it is considered only as a *Forward-Looking*, but not as a *Backward-Looking*, function. This issue has been brought up within the Discourse Resource Initiative as well.

	Self	Partner	N/A	Total
Partner Decidable Options	1	9		10
Unendorsed Options	0	2		2
Proposals	1	24		25
Unconditional Commitments	2	11		13
Other	1	3		4
No clear antecedent			6	6
Total	5	49	6	60

Table 7
Antecedents of Commit

	# Committed	# Uncommitted	Percent Committed
Partner Decidable Options	10	19	31%
Unendorsed Options	2	15	12%
Proposals	20 (25)	17	54%
Commits	12 (13)	N/A	N/A

Table 8
Percent of utterances performing dialogue functions of interest that are committed to

the number of unconditional commitments in Table 6 is only 48. This is because certain *Commits* have more than one antecedent. The category *Other* includes miscellaneous items, such as three *Info-Requests* that have gist tag *getItem* (it is possible that they should have rather been tagged as *Open-Option* or as *Action-Directive*). In the 6 *no clear antecedent* cases it is unclear to what exactly the speaker is committing. Both *Other* and *no clear antecedent* cases need further analysis.

Table 7 confirms we are on the right track. Proposals are the most frequent category that appears as an antecedent of *Commit*. We are of course not surprised that very few unendorsed options appear as antecedent of *Commit*. On the other hand, we did not expect so many partner decidable options and *Commit* to appear as antecedents of *Commit*, but after examining these examples more closely, we can actually see confirmations of the commitment pattern in (2), as we will discuss below.

Table 8 shows the percent of utterances tagged with that particular category that are committed to (such a percentage doesn't makes sense in the case of unconditional commitments). There is a significant difference among partner decidable options, unendorsed options and proposals in Table 8 in terms of whether they are committed to or not, $\chi^2 = 9.01$, $p < 0.025$.

In Table 8, proposals and commits are reported with two numbers. The lower number refers to distinct proposals and commits that appear as antecedents of *Commit*. We discuss here briefly the issue of redundant commits, namely, of utterances by the same speaker labeled *Commit* and which commit to exactly the same antecedent (note that our definition of *same* for antecedents is purely syntactic, i.e., refers to a specific utterance label, not to the content of the utterance). As far as proposals are concerned, they appear 25 times as antecedents of *Commit*, but 5 occurrences are antecedents of a redundant commit, i.e., each of these 5 already appears among the other 20 occurrences as the antecedent of another *Commit*. In these cases, S_2 expresses his/her commitment, but in the same turn, in fact in the next utterance, s/he repeats it, as in Figure 17.

[34] and [35] are both tagged as *Action-Dir+Commit*, and both have [32] as antecedent. Clearly, in this case [35] repeats the commitment in [34]. This repeated function may be due to a shortcoming of our definitions of the categories we are using in this section, that don't take into account the gist tags: what distinguishes [34] from [35]

```
G-6: [30]: That leaves us with 300 dollars total
     [31]: I have a green chair
     [32]: I'll buy for 100 dollars.    [Proposal]
     [33]: Okay?

S-7: [34]: That sounds good.            [Commit 32]
     [35]: Go ahead and buy it.         [Commit 32]
```

Figure 17
Redundant Commits

is their gist tags, *eval* and *getItem* respectively. A similar pattern occurs in the case in which two *Commit*s have as antecedent exactly the same *Commit*, which results in only 12, not 13, distinct *Commit*s functioning as antecedents of another *Commit*.

For the moment we will only discuss the 20 proposal antecedents and the 12 commit antecedents that are distinct one from the other. We leave a full account of redundant *commits* for future work.

5.1.4 Proposals, Partner Decidable Options, and Dialogue Patterns. We now look at how the dialogue unfolds according to the type of antecedent—we will only consider cases in which the commitment is to the partner's antecedent. We expect the dialogue to unfold differently according to whether the antecedent of *commit* is a partner decidable option or a proposal. For a partner decidable option to become part of the final solution, further balancing of information is necessary. On the other hand, a proposal represents a state of the dialogue in which an explicit agreement or rejection is called for, namely, either an expression of commitment to that proposal, or if not, evidence of deliberation and proposal of an alternative.

Options. Let's first examine the 9 partner decidable options, and the two unendorsed options that appear as antecedents of a *Commit* towards the partner. By definition, one would expect partner decidable options to be followed by negotiation; remember that they are characterized by an indeterminate solution size. As a minimum, we expect that the other utterances that occur between the U_i characterized as a partner decidable option, and the corresponding *Commit* to reduce the solution size for $parm_i$ to determinate.[37] This happens in all cases. However, we also have stronger expectations. We expect that S_2 will collaborate by at least balancing information, whether solicited or unsolicited by S_1; we also expect S_2 to provide evidence that s/he has performed deliberation as support for the *Commit*. In 6 out of 9 cases,[38] balance of information occurs, and it is explicitly initiated by S_1 in 4 of these cases by means of an *Info-Request*. Further, in 4 of these 6 cases, S_2 provides evidence that s/he has deliberated. These exchanges are exemplified in Figure 15: we can now recast its coding in the terms we have been using here, as shown in Figure 19 in Section 5.1.5. [35] constitutes a partner decidable option for sofas, as [35] is the first mention of a sofa. In [37] and [38] Jo informs Ju of a possible alternative which he negatively evaluates in [39], thereby showing he is not endorsing the alternative. Finally, in [40], Jo commits to [35] with an explicit positive evaluation of [35]. Thus, Jo both balances information, and provides evidence he has indeed performed deliberation.

37 In the case of unendorsed options, this is trivially true, as an unendorsed option has solution size determinate.
38 The 3 failed cases don't show any serious flaws for our stronger expectation.

The deliberation pattern in the other 2 cases in which there is balancing of information is instead more complex, as they stem from S_1 backtracking on a previous commitment, i.e., providing a new solution—namely, S_1 has reconsidered previous decisions, and proposes a *filter override* to S_2 (see Figure 9). The indeterminate solution size derives from the fact that the parameters in question are reopened (utterances [52], [55] and [65]). As S_2 initially objects towards this change of plans, it is S_1 that actually provides evidence he has performed deliberation by showing the new options provide more points. S_2 then commits to the new solution. These two cases can still be classified as *partner decidable options*, even if S_1 has performed deliberation: S_1 may not be willing to commit in the IRMA sense at this stage, as he recognizes S_2 still needs to perform deliberation.

As regards the two unendorsed options, one of them occurs in the complex context just discussed of a *filter override*. The other one is an answer to a question, i.e., to an *Info-Request* on the part of S_1 about the existence of a certain furniture item: clearly the pair *Info-Request* and its answer display balancing information, however no further deliberation occurs in this case, partly because some deliberation has actually already happened. Namely, the *Info-Request/Answer* pair occurs at the end of negotiating the dining room set, and concerns the existence of 2 red chairs, needed to complete the set and match colors.

Proposals. We will now consider how dialogue unfolds in the case of proposals. In this case, not only will we look at the 20 committed proposals, but also at the 17 that do not appear as antecedents of *Commit*—cf. Table 8. Examining how proposals that do not appear to be committed to are dealt with in the dialogue provides some further evidence that our characterization of proposals is on the right track. Further evidence that partner decidable options and proposals are indeed different will come from analyzing how the dialogue develops in the cases in which a partner decidable option is not committed to. This is left for future work.

First, let's examine the 19 distinct proposals that are committed to by the partner (the remaining proposal is committed to by the agent, as discussed below with respect to [33] and [42] in Figure 19). We expect that in this case, contrary to the options case, there is no balancing of information or deliberation going on between the proposal and the corresponding commit, as both balancing of information and deliberation must have already occurred when the proposal is uttered. In fact, this is the case: in 11 cases, the proposal (possibly paired with an utterance that elicits agreement, such as *do you want to do that?*) is immediately followed by the partner's commit, or the only utterances between the proposal and the commit concern how much money is left, or how many points have been gained. In 5 other cases, other items unrelated to the ones being proposed are brought up between the proposal and the commit, with 2 of these cases beginning a new agreement process. The remaining three cases are more complex, because they occur in the *filter override* case discussed above.

Finally, the analysis of the 17 proposals that do not appear as antecedents of *Commit* reveals that we can account for the vast majority of them in terms of our model. In fact, if our characterization of a proposal is correct, we expect that each U_i we have labeled as a proposal should be responded to in some fashion. In a collaborative setting such as ours, a partner cannot just ignore a proposal as if it had not occurred, i.e., he must give evidence of *disposition*. In this case, moving to another part of the problem is not evidence of disposition. On the other hand, if those 17 proposals were mostly ignored, our definition of proposal would probably need to be revised. It turns out that of these 17 proposals, 10 are committed to, although not in the most direct way; 5 are committed to without explicit evidence of disposition (or at least, without evidence that our coding scheme manages to capture); and only two are ignored.

- 10 proposals are indeed committed to, although not in the most direct way. This is the reason why they appear in the *Uncommitted* column in Table 8.

  — Five of them are indirectly committed to in the dialogue. Three of these 5 proposals are elaborated or simply repeated by another proposal that immediately follows, as in *[1] My blue sofa costs 300. [2] I'll buy that*: both [1] and [2] are coded as *Action-Dir+Offer* with a determinate solution size, and therefore qualify as two proposals, even if they are equivalent one to the other; because of our coding for antecedents of Commit, only the second proposal is marked as committed to. The other two of these five proposals that are indirectly committed to appear in one turn and the last utterance in that turn (a third proposal) summarizes them: only this last proposal is then counted as the antecedent of the corresponding *Commit*. Clearly, these five cases would be more perspicuously dealt with if we had taken into account the notion of redundancy (Walker 1993). Alternatively, our analysis of proposals may need to include the gist tag of the utterance (*elaborateItem* for [1] and *getItem* for [2]), in the same way that gist tags may need to be included to distinguish between apparently redundant *Commits*, as discussed below. This is left for future work.

  — Five other proposals are committed to, but not via the dialogue: i.e, the items proposed by these proposals are actually included in the final solution (we verify this by means of the graphics), but without an explicit commitment being expressed in the dialogue. Three of these cases occur in the two dialogues in which the participants follow the "initial dump" strategy (see Section 2.3 along with Figure 6), that makes it more difficult to label *Offer* and *Commit*. This is because S_1 makes an offer that includes several items at once, and S_2 replies with another offer that includes some, but not all, of the items proposed by S_1. Namely, S_2's reply is part *Offer*, part *Commit*, but as it is not possible to mark it as both, it is only marked as *Offer*. Note that the agreement tags in the coding scheme would allow partial acceptance / rejection to be marked, via *Accept-/Reject-part*.

- 5 other proposals are not committed to in any way, but they are responded to. Four among these 5 are linked to a subsequent proposal with a *MutuallyExclusive* reference tag, indicating that the partner is offering an alternative solution (which does not necessarily represent a rejection of the initial proposal yet). Recall our claim that a proposal represents a state of the dialogue in which either an expression of commitment to that proposal, or if not, evidence of deliberation and proposal of an alternative are called for. This subsequent alternative proposal is later committed to.

  The last proposal among these 5 is explicitly rejected (the partner who utters the proposal mistakenly thinks there still is some money left).

- Thus, in the end, only two proposals appear to have no consequence on the rest of the dialogue, neither in terms of commitment, nor in terms of a more general response. Not surprisingly, one of them occurs in the dialogue from which Figure 13 in Section 3 was extracted. That dialogue was an example of two non collaborative partners. The other case should actually have been linked to a subsequent proposal with a *Mutually exclusive* tag.

5.1.5 Unfolding of Commitment in Negotiation. We conclude this section by providing some evidence for the unfolding of commitment in negotiation that we presented above in (2). If we recast the commitment pattern in (2) in terms of the categories we have been discussing in this section, we obtain the following:

(3a) S_1: Partner Decidable Option (i.e., S_1 unable to commit yet) /
Proposal (i.e., Offer)

(3b) S_2: Commit (antecedent (3a))

(3c) S_1: Commit (antecedent (3b))

We will now provide evidence for the existence of this commitment pattern by examining commitments to an antecedent also tagged as unconditional commitment, i.e., we will examine in which contexts (3c) is made explicit.

Commit to an antecedent Commit uttered by the partner. The 10 distinct cases in which S_1 commits to a commitment from S_2 are candidates for verifying our explicit commitment pattern.[39] In 6 of these 10 cases we see exactly this pattern. The seventh case among the 10 adds a further step to one of the 6 full patterns: this optional fourth step, (4d), is in turn also a *Commit* with another *Commit* as antecedent. It is in fact a repetition of an unconditional commitment ((4b)) that has been already expressed—note this is a redundant commit, but of a different nature than those that we discussed at the end of Section 5.1.3. Earlier, we identified redundant *Commit*s just in syntactic terms, namely, when the antecedents of the two *Commit*s have exactly the same label. The antecedent of (4d), instead, has a different label than the antecedent of (4b); nonetheless, (4d) is redundant with respect to (4b):

(4a) S_1: Partner Decidable Option (i.e., S_1 unable to commit yet) /
Proposal (i.e., Offer)

(4b) S_2: Commit (antecedent (4a))

(4c) S_1: Commit (antecedent (4b))

(4d) S_2: Commit (antecedent (4c))

We had mentioned above that we expected commitment to unfold in just the explicit way represented by pattern (3), i.e., step (3c) (the same as (4c)) to be explicit more often when step (3a) (the same as (4a)) is a partner decidable option than when it is a proposal. Although the numbers are too small to draw any real conclusion, the trends go in the direction we were expecting: of the 6 full commitment patterns we found (with *full* we mean as in (3)), 4 occur when step (3a) is a partner decidable option, i.e., almost 50% of the times in which a partner decidable option is committed to by the partner; one

39 These 10 cases are obtained as follows. There are 13 *Commit*s committed to in Table 8, 12 of which are distinct—cf. the discussion at the end of Section 5.1.3. The redundant *Commit* commits to a *Commit* uttered by the partner. Going back to Table 7, this means that the 12 distinct *Commits* include the 2 *Commit* whose antecedent was uttered by the agent herself. Hence, eliminating these 2 *Commit*s, we are left with 10 distinct *Commit*s whose antecedent is a *Commit* uttered by the partner.

of these 4 processes is the one that includes the extra step (4d) (i.e., case 7 discussed above). On the other hand, only two of these full processes occur when step (3a) is a proposal, i.e. 10% of the times a proposal is explicitly committed to.

In the other 3 cases what is missing is actually step (3a), i.e., there was no explicit partner decidable option or proposal for that specific parameter. Since step (3a) is missing, S_1 has not expressed any attitude towards that specific option for action yet, not even shown that he was unable to commit to it. For example, in Figure 18, [27] expresses a commitment to two different antecedents: [19] and [20]. As far as [19] is concerned, P shows she is now committed to her original partner decidable option in [18], i.e., this is one of the 6 full patterns we discussed above. As far as [20] is concerned, P had not expressed any attitude towards J's green chairs, as P didn't even know such chairs existed. The question of course may be why [20] is labeled a *commit* and not as an *offer*: clearly, the coder took into account that P's *Info-Request* in [17] shows she is willing to entertain a solution that includes J's green chairs. However, [17] doesn't qualify as a referring to an action, thus it can't even be labeled as a partner decidable option.

```
P-1: [13]: Our total budget is $1000 (450+550).
     [14]: I have a red table $400, and 2 red chairs $50 each,
     [15]: can we make a match for cheaper than $800?
     [16]: Do you have any red chairs?
     [17]: Or green chairs?                      [parm=chairs; Info-Request]
     [18]: I have a green table $200,
           & 2 green chairs $100 each.           [parm=table,chairs; Partner Dec. option]

J-2: [19]: how about your green table and chairs  [parm=table,chairs; Commit 18]
     [20]: and i have 2 green chairs for $50 each.  [Answer 17; Commit, no antec.]
     [21]: that is $500.
     [22]: that is the cheapest we can do.
     [23]: that still leaves us with $500 for the living room.
     [24]: i have 3 sofas.
     [25]: 1 yellow - $400, 1 red - $550, and 1 blue - $300.

P-2: [26]: good,
     [27]: the dining room is done!!             [parm=table,chairs; Commit 19,20]
```

Figure 18
A commit without an explicit antecedent ([20])

Committing to an antecedent uttered by the agent. Let's consider now the 5 cases in which a speaker S_1 commits to an antecedent that S_1 had uttered herself (see Table 7). The two cases in which S_1 recommits to an action she had already unconditionally committed to are clear examples of redundancies, and the case labeled *Other* needs further analysis; more interesting are the two cases in which S_1 commits to one's own partner decidable option or proposal. These cases pattern as in (5), where the *Null* step on the part of S_2 means that S_2 does not express any attitude towards (5a) (not that S_2 is silent!).

(5a) S_1: Partner Decidable Option (i.e., S_1 unable to commit yet) /
Proposal (i.e., Offer)

(5b) S_2: Null

(5c) S_1: Commit (antecedent (5a))

Walker (1996) points out that if participants in a conversation follow the collaborative principle, they *must provide evidence of detected discrepancy in belief as soon as possible.*

```
Ju-3: [22]: i think the 50 ones [chairs] are better
      [23]: how about 2 of yours             [parm=chairs; Partner dec. option]
      [24]: and i have 2 also.               [parm=chairs; Partner dec. option]
      [25]: mine are green and red.
      [26]: which?
      [27]: and what colors are yours        [parm=chairs; Info-Request]
      [28]: so i can log them in.
      [29]: i just put in your table.
      [30]: living room next.....

Jo-3: [31]: The 2 chairs that i have are red.  [parm=chairs; Answer 27]

Ju-4: [32]: well
      [33]: how about we use 2 of your chairs  [parm=chairs; Proposal]
            and 2 of my red.
      [34]: we will have a christmas room...
      [35]: i have a blue sofa for 300.      [parm=sofa; Partner dec. option]
      [36]: it's my cheapest one.

Jo-4: [37]: I have 1 sofa for 350
      [38]: that is yellow                   [parm=sofa; Unendorsed option]
      [39]: which is my cheapest,
      [40]: yours sounds good.               [parm=sofa; Commit 35]

Ju-5: [41]: ok
      [42]: i logged in 2 of your chairs     [parm=chairs; Commit 33]
            and 2 of mine..
      [43]: both red.
      [44]: I'll order that blue sofa.       [parm=sofa; Commit 40]
```

Figure 19
Two instantiations of the commitment pattern ([33]-[42] and [35]-[40]-[44])

Walker defines the first opportunity that one conversant has to express their opinion with respect to a certain proposal as the *attitude locus*. In our framework, the attitude locus is the first turn a subject has after a certain option for action has been presented: given our policy of strict turn-taking, our subjects often have to fulfill several *obligations* from the previous turn (Traum and Allen 1994). Sometimes subjects don't fulfill all their obligations, in particular, they may not express any attitude towards their partner's option for action. When this happens, the partner S_1 who presented the option for action may express that his/her explicit commitment to the original proposal, initially unattainable or conditional, is now unconditional. Basically S_1 makes it explicit that s/he has interpreted S_2's lack of an explicit rejection as an implicit commitment, and offers S_2 an opportunity to voice his disagreement, if S_1's inference is wrong. This happens in the excerpt in Figure 19, that includes the excerpt from Figure 15 we have been discussing at length. Jo in [31] only answers the question in [27], but does not show he has performed any deliberation with respect to the partner decidable option in [23]-[24]. Ju goes on to making a specific proposal in [33], and again, Jo does not express any opinion with respect to it in his turn. At this point, Ju makes it clear in [42] that her conditional commitment in [33] has become unconditional.

6. Conclusions and Future Directions

In this paper, we have explored corpus correlations of the agreement process by examining how utterances related to a single task purpose function in the negotiation of a solution. While other researchers have studied components of the process (specifically acceptances and rejections) and strategies to reconcile disagreements, we have attempted to look at the process as a whole to see how agreements on solutions are arrived at and how the context of the problem solving situation can help guide the collaboration.

6.1 Generality of the Agreement Process

The agreement process is motivated by general models of problem solving and collaboration along with information about the problem at an abstract level. We generally assume for all problem types that in a collaborative setting both parties are aware of one another and reason at an abstract level about what the other's role is in the collaborative effort. From this assumption and the models of problem solving and collaboration we can form expectations for how the dialogue should unfold.

The particular problem type we have empirically explored here is a design problem in which instantiations of knowledge are evenly distributed. We expect that different information distributions will alter how the agreement process is realized. In future work we hope to explore the agreement process in more generality, beginning with cases in which there are uneven information distributions.

For example, in an advisory setting such as those studied in (Walker 1996; Chu-Carroll and Carberry 1998) the participants should generally expect knowledge types rather than instantiations of knowledge to be distributed. In particular, the expert would have more knowledge about actions and parameter values while the client has more knowledge about the goals. In this case we would assume that the client (male) reasons that the expert (female) may have some options to suggest whenever he is unable to solve the problem. That is, the client realizes he does not have enough information to deliberate to the point of making a commitment. The client has to decide what to tell the expert so that the expert can fully deliberate and propose an option that she is willing to commit to. If the client just gives a single goal to the expert then the expert isn't likely to be able to come up with good options. Even in simple cases, clients seeking help in making a decision will have multiple goals, and some of these goals may conflict. In advising the client, the expert needs to arrive at a sense of these goals and of their relative priorities. The expert may also need to know many problem specifics, as well as any relevant commitments that have already been made by the client. The expert's information needs call for a (sometimes elaborate) interview in which the balancing will negotiate preferences, constraints, and background information. Without this information gathering phase, the expert cannot hope to come up with a good option or to explain it to the client.

Modalities of the conversation will also affect the discourse. In an interruptible dialogue situation, we would expect the balancing to be more efficient than in a non-interruptible dialogue since the expert can interrupt when she has enough information to do a full deliberation.

Once the expert has a solution she can commit to, then she has to give the client enough information so that the client can deliberate to the point of making a commitment. In this case it would be information about the actions and parameters that contribute to the solution. The client can't always just be told the actions and parameters, he must understand how the actions interact and contribute to the solution. Otherwise he cannot properly deliberate about the option the expert will propose.

So, in general we would expect to see a more extended balancing phase at the highest level of the collaboration than with design dialogues because of the differences in the knowledge distributions (see for example the information-sharing subdialogues studied by Chu-Carroll and Carberry (1998)). However, the more expansive the client-expert common ground at the outset of the dialogue, the closer the instantiation of the agreement process should become to a design situation.

This example of the expected effect of the different knowledge distributions for an advisory task is speculative and is meant to just give an idea of how we think the general process might extend to other types of dialogues. There will be different challenges to address for each type of problem. For example, how balancing itself unfolds in the advisory case might be strongly affected by an interruptible vs. non-interruptible setting. Perhaps the balancing by the expert and the client will be more interleaved than in the design case regardless of interruptibility. The main requirement for making use of the agreement process as a predictive mechanism is identifying when the agents might be in a position to fully deliberate. Obviously, whether an agent has been told anything at all about the parameter values the partner has and how many options are available is a good simple indicator for the design case, but there may not be such simple indicators available for other types of problems. If no good indicators can be found for when the agents are in a position to deliberate to the point of committing, then it will be difficult to infer a single agent commitment in the IRMA sense. Perhaps we will find that in the absence of good indicators, single agent commitment is much more explicit.

6.2 Collaboration Patterns

Our empirical study has shown that tracking commitment at the collaborative level and at the single agent level, including situations in which the agent shows she is unable to commit to an option for action yet, can provide a better sense of how an agreement is reached than attempting to pinpoint which utterances accept and which reject a proposal. We have identified two related problems. One is that these particular functions can be implicit. The other is that an accept or reject only considers one agent's attitude toward an action and doesn't give us a clear sense of joint agreement. We need to track how commitment evolves from an inability to commit (*partner decidable option*), to conditional commitment (*proposal*), to unconditional commitment, in order to model how agreement is reached.

In addition to the general finding about the advantage of commitment vs. acceptance for recognizing an agreement, we saw support for the idea of combining IRMA and Clark's acceptance process in accounting for discourse concerning a collaborative design task. In particular, we expected IRMA options to be realized in this collaborative task as: partner decidable options, unendorsed options and proposals. We confirmed these categories first by projecting how we expected them to behave and their expected context and then empirically checking for correlations in the tagged corpus. In particular, we found trends that indicate that what we defined as a proposal is more likely to refer to an action that the partners will agree on than what we defined as a partner decidable option. Also we found that what we identified as proposals are generally responded to with an utterance that shows whether the collaborator is willing to commit to it, whereas for a partner decidable option to become part of the final solution further balancing of information and deliberation are necessary.

From the agreement process we also expected certain patterns of negotiation based on what could and could not be inferred. We had two basic situations to consider: one in which a partner decidable option becomes the agreed to solution and one in which a proposal becomes the agreed to solution. We confirmed the following expected patterns for these two situations: it is more likely that later in the agreement process, S_1 explicitly expresses that her commitment is unconditional if she earlier presented a partner decidable option (showing S_1 was unable to commit to it at that time), than if she earlier presented a proposal (showing she was conditionally committed at that time). Another situation in which S_1 makes her unconditional commitment explicit later in the agreement process is if S_2 "passively" accepted S_1's option.

As regards future work proper, we mentioned that we need to still explore how the dialogue develops when a partner decidable option does not become part of the solution; we expect that this will shed further light on the difference between partner decidable options and proposals. We also mentioned in Section 5 that we need to investigate cases of redundant *Commit*, and whether we should take into account the gist of an utterance to distinguish them. Finally, we would like to investigate *contingent proposals* in which a decision about an action allows one to infer the agreement status for hierarchically related actions for which agreement is still pending. The notion of *meta-actions* briefly discussed in Section 5.1.3 is related to contingent proposals.

Acknowledgments
This material is based on work supported by the National Science Foundation under Grant No. IRI–9314961. The work was conducted while all the authors were affiliated with the University of Pittsburgh. We wish to acknowledge project members Megan Moser and Jerry Hobbs; a special mention to Liina Pyllkänen for her contributions to developing our coding schema, and to the coding effort proper. We also wish to thank Marilyn Walker for stimulating discussions and for making some of the spoken furniture design dialogues available to us, Steve Whittaker for suggesting studies on dialogue modalities different from face to face, and the reviewers and editors of the special issue for their constructive suggestions.

References

Austin, John L. 1962. *How to Do Things With Words.* Oxford University Press, Oxford.

Biermann, A., C. Guinn, D. Hipp, and R. Smith. 1993. Efficient collaborative discourse: A theory and its implementation. In *ARPA Human Language Technology Workshop.*

Bratman, Michael. 1990. What is Intention? In P. Cohen, J. Morgan, and M. Pollack, editors, *Intentions in Communication.* MIT Press.

Bratman, Michael. 1992. Shared cooperative activity. *Philosophical Review*, 101:327–341.

Bratman, Michael, David Israel, and Martha Pollack. 1988. Plan and resource-bounded practical reasoning. *Computational Intelligence*, 4:349–355.

Brown, Gretchen P. 1980. Characterizing indirect speech acts. *Computational Linguistics*, 6:150–166.

Carletta, Jean. 1996. Assessing agreement on classification tasks: the kappa statistic. *Computational Lingustics*, 22(2):249–254.

Carletta, Jean, Amy Isard, Stephen Isard, Jacqueline C. Kowtko, Gwyneth Doherty-Sneddon, and Anne H. Anderson. 1997. The reliability of a dialogue structure coding scheme. *Computational Lingustics*, 23(1):13–31.

Chu-Carroll, Jennifer and Sandra Carberry. 1998. Collaborative response generation in planning dialogues. *Computational Linguistics*, 24(3).

Clark, Herbert H. 1992. *Arenas of Language Use.* The University of Chicago Press.

Clark, Herbert H. 1996. *Using Language.* Cambridge University Press.

Clark, Herbert H. and Catherine Marshall. 1981. Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber, and Ivan A. Sag, editors, *Elements of discourse understanding.* New York: Cambridge University Press, pages 10–63.

Clark, Herbert H. and Edward F. Schaefer. 1987. Collaborating on contributions to conversations. *Language and Cognitive Processes*, 2:1–23.

Clark, Herbert H. and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 27:181–21.

Computer-Mediated Communication Magazine. [On-line] Available: http://www.december.com/cmc/mag. Published May 94-January 99.

Cohen, Philip and Hector Levesque. 1990. Rational Interaction as the Basis for Communication. In P. Cohen, J. Morgan, and M. Pollack, editors, *Intentions in Communication.* MIT Press.

Cohen, Philip and Hector Levesque. 1991. Teamwork. *Noûs*, 25(4):487–512.

Condon, Sherri L. and Claude G. Čech. 1996. Functional comparisons of face-to-face and computer-mediated decision making interactions. In Susan C. Herring, editor, *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*. John Benjamins Publishing Company.

Core, Mark G. and James Allen. 1997. Coding dialogues with the DAMSL annotation scheme. In David Traum, editor, *Working Papers of the AAAI Fall Symposium on Communicative Actions in Human and Machines*. American Association for Artificial Intelligence.

1997. DAMSL: Dialog act markup in several layers. Available under *Tools and resources*, at http://www.georgetown.edu/luperfoy/Discourse-Treebank/dri-home.html.

Di Eugenio, Barbara, Pamela W. Jordan, Johanna D. Moore, and Richmond H. Thomason. 1998. An empirical investigation of proposals in collaborative dialogues. In *ACL/COLING 98, Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics* (joint with the 17th International Conference on Computational Linguistics), Montreal, Canada.

Di Eugenio, Barbara, Pamela W. Jordan, and Liina Pylkkänen. 1998. The COCONUT project: Dialogue annotation manual. Technical Report ISP Technical Report 98-1, University of Pittsburgh, December. [On-line] Available: http://www.isp.pitt.edu/~intgen/research-papers.

Ephrati, Eithan, Martha E. Pollack, and Sigalit Ur. 1995. Deriving multi-agent coordination through filtering strategies. In *14th International Joint Conference on Artificial Intelligence*.

Etzioni, Oren and Daniel Weld. 1994. A softbot-based interface to the internet. *Communications of the ACM*, 37(7):72–76.

Gross, Derek, James Allen, and David Traum. 1993. The TRAINS 91 dialogues. Technical Report TRAINS 92-1, University of Rochester.

Grosz, Barbara and Candace Sidner. 1990. Plans for Discourse. In P. Cohen, J. Morgan, and M. Pollack, editors, *Intentions in Communication*. MIT Press.

Grosz, Barbara J. 1996. Collaborative Systems. *AI Magazine*, 17(2):67–85, Summer.

Grosz, Barbara J. and Sarit Kraus. 1996. Collaborative Plans for Complex Group Actions. *Artificial Intelligence*, 86/2:269–357, October.

Grove, William M., Nancy C. Andreasen, Patricia McDonald-Scott, Martin B. Keller, and Robert W. Shapiro. 1981. Reliability Studies of Psychiatric Diagnosis. Theory and Practice. *Archives of General Psychiatry*, 38:408–413.

Heeman, Peter and James Allen. 1995. The TRAINS 93 dialogues. Technical Report TRAINS 94-2, University of Rochester.

Heeman, Peter A. and Graeme Hirst. 1995. Collaborating on referring expressions. *Computational Linguistics*, 21(3):351–382.

Journal of Computer-Mediated Communication. [On-line] Available: http://www.ascusc.org/jcmc/. Published since June 1995.

Joslin, David. 1996. *Passive and Active Decision Postponement in Plan Generation*. Ph.D. thesis, University of Pittsburgh, Intelligent Systems Program.

Krippendorff, Klaus. 1980. *Content Analysis: an Introduction to its Methodology*. Beverly Hills: Sage Publications.

Lambert, Lynn and Sandra Carberry. 1991. A Tripartite Plan-Based Model of Dialogue. In *ACL91, Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 47–54.

Lambert, Lynn and Sandra Carberry. 1992. Modeling Negotiation Subdialogues. In *ACL92, Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*, pages 193–200.

Litman, Diane. 1985. *Plan Recognition and Discourse Analysis. An Integrated Approach for Understanding Dialogues*. Ph.D. thesis, University of Rochester.

Lochbaum, Karen. 1995. The Use of Knowledge Precondition in Language Processing. In *IJCAI95, Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pages 1260–1266.

Lochbaum, Karen E. 1994. *Using Collaborative Plans to Model the Intentional Structure of Discourse*. Ph.D. thesis, Harvard University. Technical Report TR-25-94.

Lochbaum, Karen E., Barbara J. Grosz, and Candace L. Sidner. 1990. Models of plans to support communication. In *AAAI90, Proceedings of the Eighth National Conference on Artificial Intelligence*, Boston, CA.

Lottaz, Claudio. 1996. Constraint Solving, Preference Activation and Solution Adaptation in IDIOM. Technical Report 96/204, Artificial Intelligence Laboratory, Swiss Federal Institute of Technology in Lausanne, Switzerland.

Lottaz, Claudio and Ian Smith. 1997. Collaborative design using constraint solving. From Swiss Workshop on Collaborative and Distributed Systems, Lausanne Switzerland. See http://liawww.epfl.ch/ lottaz/ICCS/Design_and_CSP/design_and_CSP.html and http://liawww.epfl.ch/ lottaz/ICCS/Collaboration/index.html, May.

Lyons, Kevin W. 1995. Collaborative design for assembly of complex electro-mechanical products. Presentation abstract for NCMS Manufacturing Technical Conference. See http://elib.cme.nist.gov/made/presentations/ncms.html, May.

Maes, Patti. 1994. Agents that reduce work and information overload. *Communications of the ACM*, 37(7):31–40.

Moore, Robert C. 1985. A formal theory of knowledge and action. In Jerry R. Hobbs and Robert C. Moore, editors, *Formal Theories of the Commonsense World*. Ablex Publishing Corporation, Norwood, New Jersey, pages 319–358.

Moser, Megan, Johanna D. Moore, and Erin Glendening. 1996. Instructions for Coding Explanations: Identifying Segments, Relations and Minimal Units. Technical Report 96-17, University of Pittsburgh, Department of Computer Science.

Nakatani, Christine H., Barbara J. Grosz, David D. Hahn, and Julia Hirschberg. 1995. Instructions for annotating discourses. Technical Report TR-25-95, Harvard University Center for Research in Computing Technology.

Novick, David G., Catherine R. Marshall, Brian Hansen, and Karen Ward. 1998. Implications of co-presence and acceptance for cooperative systems. In *Working Notes of the COOP 98 Workshop on the use of Herbert Clark's models of language use for the design of cooperative systems.*

Novick, David G. and Karen Ward. 1993. Mutual beliefs of multiple conversants: A computational model of collaboration in air traffic control. In *AAAI93, Proceedings of the 12th Conference of the American Association for Artificial Intelligence*, pages 196–201.

O'Conaill, Brid, Steve Whittaker, and Sylvia Wilbur. 1993. Conversations over video conferences: An evaluation of the spoken aspects of video-mediated communication. *Human-Computer Interaction*, 8:389–428.

Passonneau, Rebecca J. 1994. Protocol for coding discourse referential noun phrases and their antecedents. Technical report, Columbia University.

Pollack, Martha E. 1992. The uses of plans. *Artificial Intelligence*, 57:43–68.

Qu, Yan. 1997. A constraint-based model for cooperative response generation in information systems. PhD Proposal, Computational Linguistics Program, Carnegie Mellon University.

Ramshaw, Lance A. 1991. A Three-Level Model for Plan Exploration. In *ACL91, Proceedings of the 29th Meeting of the Association for Computational Linguistics*, pages 39–46.

Rich, Charles and Candace Sidner. 1997. Collagen: when agents collaborate with people. In M. Huhns and M. Singh, editors, *Readings in agents*. Morgan Kaufmann.

Rietveld, T. and R. van Hout. 1993. *Statistical Techniques for the Study of Language and Language Behaviour*. Mouton de Gruyter.

Rosenschein, J. and G. Zlotkin. 1994. *Rules of Encounter: Designing Conventions for Automatic Negotiation among Computers*. The MIT Press, Cambridge, Mass.

Schober, Michael F. and Herbert H. Clark. 1989. Understanding by addressees and overhearers. *Cognitive Psychology*, 21:211–232.

Searle, John R. 1965. What Is a Speech Act. In Max Black, editor, *Philosophy in America*. Cornell University Press, Ithaca, New York, pages 615–628. Reprinted in *Pragmatics. A Reader*, Steven Davis editor, Oxford University Press, 1991.

Searle, John R. 1975. Indirect Speech Acts. In P. Cole and J.L. Morgan, editors, *Syntax and Semantics 3. Speech Acts*. Academic Press. Reprinted in *Pragmatics. A Reader*, Steven Davis editor, Oxford University Press, 1991.

Sidner, Candace. 1992. Using discourse to negotiate in collaborative activity: An artificial language. In *AAAI Workshop on Cooperation among Heterogeneous Agents*. American Association for Artificial Intelligence.

Sidner, Candace. 1994. An artificial discourse language for collaborative negotiation. In *AAAI94, Proceedings of the Twelfth National Conference on Artificial Intelligence.*

Simon, Herbert. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69:99–118.

Thomason, Richmond. 1990. Accommodation, Meaning, and Implicature: Interdisciplinary Foundations for Pragmatics. In P. Cohen, J. Morgan, and M. Pollack, editors, *Intentions in Communication.* MIT Press.

Thomason, Richmond and Jerry Hobbs. 1997. Interrelating interpretation and generation in an abductive framework. In David Traum, editor, *Working Papers of the AAAI Fall Symposium on Communicative Action in Humans and Machines*, pages 97–105. American Association for Artificial Intelligence.

Thomason, Richmond and Johanna Moore. 1995. Discourse context. In Sasa Buvač, editor, *Formalizing Context*, pages 102–109, Menlo Park, California. American Association for Artificial Intelligence.

Traum, David. 1994. *A Computational Theory of Grounding in Natural Language Conversation.* Ph.D. thesis, University of Rochester, December. Technical Report 545.

Traum, David and James Allen. 1994. Discourse Obligations in Dialogue Processing. In *ACL94, Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, pages 1–8.

Tuomela, Raimo. 1995. *The Importance of Us.* Stanford University Press.

Walker, Marilyn A. 1993. *Informational Redundancy and Resource Bounds in Dialogue.* Ph.D. thesis, University of Pennsylvania, December.

Walker, Marilyn A. 1995. Efficiency tradeoffs for language and action in collaborative tasks. Technical Report FS-95-05, American Association for Artificial Intelligence. Working Papers of the AAAI Fall Symposium on Embodied Language and Action.

Walker, Marilyn A. 1996. Inferring acceptance and rejection in dialogue by default rules of inference. *Language and Speech*, 39(2).

Walker, Marilyn A., Diane J. Litman, Candace Kamm, and Alicia Abella. 1997. Paradise: A framework for evaluating spoken dialogue agents. In *35th Meeting of ACL and 8th Conference of the EACL.*

Webber, Bonnie Lynn. 1999. Computational aspects of discourse and dialogue. In Deborah Schiffrin, Deborah Tannen, and Heidi Hamilton, editors, *The Handbook of Discourse Analysis.* Blackwell Publishers.

Whittaker, Steve. 1995. Rethinking video as a technology for interpersonal communications: theory and design implications. *International Journal of Human-Computer Studies*, 42:501–529.

Whittaker, Steve, Erik Geelhoed, and Elizabeth Robinson. 1993. Shared workspaces: How do they work and when are they useful? *International Journal of Man-Machine Studies*, 39:813–842.