

Communication and Cooperation among Agents

Guido Boella and Rossana Damiano

Dipartimento di Informatica and
Centro di Scienza Cognitiva - Università di Torino
C.so Svizzera 185 10149 Torino ITALY
email: {guido,rossana}@di.unito.it

Abstract

In this paper we consider the consequences for dialog modeling of a new proposal of cooperation based on the concepts of group utility, goal adoption and anticipatory coordination. On the one hand, the model accounts for the contextual occurrence of communicative acts during cooperative activity; on the other hand, it models grounding phenomena in dialog, seen as a cooperative activity, without explicitly prescribing them.

1. Introduction

One of the main aims of the definitions of cooperation is to explain when and why the members of a group have to communicate with each other. In particular, the attention has been focused on those parts of dialogs which have the function of coordinating helpful behavior (Grosz and Kraus, 1996) and the conclusion of the group's activity (Cohen and Levesque, 1991).

In (Cohen and Levesque, 1991), the definition of cooperation is built out of a number of examples which show how communication can be used to establish the right mutual beliefs among the agents. The role of communication in those examples is to ensure that the activity of the group is not jeopardised by a relevant divergence among the participants' mental states, especially for what the achievement of the joint goal is concerned. In order to ensure that the group ends its activity in a felicitous way, (Cohen and Levesque, 1991) add to the definition of joint intention a number of subsidiary goals: in particular, when the joint goal has been achieved, each agent has the goal that the other partners be aware of this situation; typically, but not necessarily, these goals are achieved via communication, thus explaining dialogic phenomena like notifications.

In this work we describe the consequences for dialog modeling of extending and generalizing the intuition of (Cohen and Levesque, 1991) that the lack of coordination leads to a worse performance of the group. We claim that coordination via communication can be explained by a sort of means-ends reasoning which aims at not decreasing the group performance.

In this way, it is possible to achieve the separation between the definition of cooperation and communication claimed by (Castelfranchi, 1998): communication and the related goals of making the knowledge mutual among the group members are just instrumental to the group joint action and not prescribed by it.

In order to do so we need two conceptual instruments for measuring the expected result of the group's action and for predicting what the partners are going to do. These two forms of reasoning are becoming increasingly exploited in the multiagent field; in particular, decision theory is used for making an agent choose the most promising course of

action while having a look at the benefit of the whole group (Hogg and Jennings, 2000); on the other hand, game theoretic concepts are used for predicting the behavior of the other agents and choosing a suitable action accordingly (Ndiaye and Jameson, 1996), (Gmytrasiewicz and Durfee, 1995), (Boella et al., 2000).

The main ideas underlying the new definition of cooperation are the following ones: first, when a member of the group chooses an action, he has to take into account the benefit of the whole group and not only the utility achieved for himself. Second, the benefit for the group must be computed by taking into account how the agent's choices affect the choices of other agents; that is, an agent has to predict which actions the partners will choose after the execution of the selected actions (anticipatory coordination); he can do so, since in the contest of a group he knows (at least partially) the tasks assigned to the other partners of the group. Finally, an agent should consider whether to adopt the goals of his partners in order to provide them with help (not only material help, but also information); see (Castelfranchi, 1998) for the importance of adoption in modeling interaction among agents.

From this sketch of the model it is possible to understand why it is not necessary to include into the definition of cooperation goals concerning the achievement of a mutual belief, nor the goal to communicate with the partners. However, the model can explain the occurrence of communication during cooperative interaction, even if it is not explicitly prescribed. In particular, communicative acts aiming at enhancing the coordination among the members of a group are performed depending on the utility they yield with respect to their cost. Notifications that a partner's goal has been achieved (or is unachievable) are motivated by the adoption of his goal to know whether these conditions apply to the goal he is committed to: under some circumstances, this behavior can be useful from the group's perspective, because it prevents the partner from wasting resources. At the same time, agents can communicate to signal the intention to adopt a goal of a partner, or to avoid conflicts in the use of shared resources: in the following we will detail the utility-driven reasoning underlying these phenomena.

On the other hand, the model can be applied as well to

communication itself, seen as a cooperative activity: once two agents are involved in a communicative interaction, they choose subsequent dialogue moves in the light of the group utility they are expected to yield. Under this perspective, acknowledgements of understanding, repetitions, repairs and other dialog management actions are the result of a trade-off between the joint goal to ensure mutual understanding among the participants and the cost of these actions in terms of the resources they require.

2. The Definition of Cooperation

A group of agents GR composed of agents G_1, \dots, G_n cooperates to a shared plan¹ for α with an associated recipe R^x composed of steps $\beta_1^{x,i_1}, \dots, \beta_m^{x,i_m}$ when:²

1. each step $\beta_r^{x,i}$ has been assigned to an agent G_i in GR for its execution;
2. each agent G_i of the group GR has the single agent intention to perform his part $\beta_r^{x,i}$ of the shared plan for α formed on the basis of the recipe R^x ;
3. the agents of GR have the mutual belief that each one (G_i) has the intention to perform his part $\beta_r^{x,i}$ of the shared plan for α ;
4. all agents mutually know that they share a utility function GF based on a weighted sum of the utility of the goal which the shared plan aims at and of the resource consumption of the single agents; each agent, when he plans his own part of the shared plan, has to consider also this global utility as part of the his individual utility function F_i ;
5. when an agent G_i becomes aware that a partner G_j has a goal ϕ that stems from his intention to do his part $\beta_p^{x,j}$, G_i will consider whether to adopt it; if G_i believes that the adoption of ϕ produces an increase of the utility GF of the whole group, then he adopts that goal;
6. each agent remains in the group as long as the value of the utility function GF can be increased by executing his part of the shared plan for α or by adopting some of the goals of the partners.

Here, we don't address the problem of group formation: we assume that a group is already at work, with different tasks assigned to the various members of the group. Moreover, the definition of cooperation presented above presupposes an appropriate model for planning, which includes the mechanism of anticipatory coordination.

¹Actually, we assume here that the initial shared plan has a particular structure; i.e., it is a one-level plan composed of a top-level action (α) decomposed into a sequence of steps. In a general plan, each step could in turn have been expanded into substeps, and so on recursively.

²The notation $\beta_l^{x,i}$ refers to the l -th step of the recipe R^x , a step which has to be executed by agent G_i .

3. The Anticipatory Coordination Planning

In order to plan and select actions, we exploited a decision theoretic planner, DRIPS (Haddawy and Hanks, 1998). DRIPS is a hierarchical planner which merges some ideas of decision theory with probabilistic planning techniques. Plans are organized along decomposition and abstraction hierarchies and have non-deterministic effects which produce a probability distributions over the action outcomes. A utility function is used for evaluating how promising a given plan is. The utility function does not compute just the payoff of the possible outcomes of the plan: it is computed starting from simpler utility functions associated with the goals of the agent.³

However, the planner itself is not sufficient for our purposes. In fact, since the agent's world is populated by other agents, the consequences of an action may affect the subsequent behavior of other agents. So, in case of interaction, an agent has to consider the consequences of his behavior with respect to what the other agents will do afterwards: in order to evaluate the real expected utility of the plan, he must explore the outcomes that may result from the different reactions of other agents.

Moreover, as stated above, when an agents comes to know a goal of a partner, the agent considers whether to adopt it as part of his intentions: he does so, if the adoption of the partner's goal as a new intention amounts to an increase of the global utility notwithstanding the decrease of his partial utility due to the overhead of helping the partner.

The construction of a plan is carried out by an agent G_i in a stepwise fashion: if G_i is in charge of step $\beta_m^{x,i}$ of the recipe R^x shared with G_j ⁴ for achieving goal α , he first has to find the best recipe for $\beta_m^{x,i}$ (let's say R^y , with steps $\gamma_{m,1}^{y,i}, \gamma_{m,2}^{y,i}, \dots, \gamma_{m,n_m}^{y,i}$), and then he can start refining $\gamma_{m,1}^{y,i}$. The approach of DRIPS to this process is to expand $\beta_m^{x,i}$ in all possible ways (i.e. applying to the current state S all existing recipes); then, it proceeds onward and expands the new partial plans. The search goes on in parallel, but the search tree is pruned using the utility function (applied to the state resulting from the potential execution of the recipe): so, the utility function acts as a heuristic able to exclude some possible ways (recipes) to execute an action.

In order to implement the ideas presented in the previous section, we had to make the evaluation of the heuristics somewhat more complex. Therefore, we have modified the DRIPS algorithm for allowing anticipatory coordination. In particular, since the planning agent has to predict the reactions of the partners, he must be endowed with the (assumed) knowledge concerning his partner's beliefs and utility functions. Moreover, we assume that the knowledge about actions is shared among all the group members.

The method we employed is the following:

- Using DRIPS (playing the role of G_i), we expand the current state S according to all alternative recipes for $\beta_m^{x,i}$, thus producing the states S_1, S_2, \dots, S_r (where r is the number of different recipes for $\beta_m^{x,i}$).

³This aggregation of simpler utility functions in a global one is possible only if some independence assumptions hold (see (Haddawy and Hanks, 1998)).

⁴For simplicity we have assumed a single partner G_j .

- This set of states is transformed in the set of the same states as viewed by $G_j, S'_1, S'_2, \dots, S'_r$.
- On each state S'_m ($1 \leq m \leq r$), we restart the planning process from the perspective of his partner G_j (i.e. trying to solve his current task $\beta_h^{x,j}$).
- This produces a set of sets of states $SS' = \{\{S'_{1,1}, \dots, S'_{1,n_1}\}, \{S'_{2,1}, \dots, S'_{2,n_2}\}, \dots, \{S'_{n_r,1}, \dots, S'_{n_r,n_r}\}\}$.
- The group utility function is applied to these states, and the best state of each subset is identified: $SS'_{best} = \{S'_{1,best(1)}, S'_{2,best(2)}, \dots, S'_{n_r,best(n_r)}\}$. These states are the ones assumed to be reached by G_j 's best action, for each of the possible G_i initial moves.
- The group utility function is applied to the states $S_{k,best(k)}$ ($1 \leq k \leq r$) from G_i 's point of view. This models the perspective of G_i on what could happen next.
- The best one of these states is selected ($S_{max,best(max)}$). This corresponds to the selection of the best recipe for $\beta_m^{x,i}$ of G_i (i.e. R^{max}).

Note that the algorithm above is just a modification of a two-level min-max algorithm: actually, it is a max-max, since at both levels the best option is selected, although at the second level it is evaluated from G_j 's perspective. As in min-max, G_i , when predicting G_j behavior, assumes that his partner is a rational agent, i.e. that he will choose the plan that gets the highest utility for the group.

The problem of simulating another agent's planning is very difficult. For instance, in some situations, G_j could not be aware of R^x effects. In our implementation, we adopted the simplification that the initial state is shared by the agents, while, during the planning phase, G_j 's knowledge of a state is updated in G_i 's beliefs by an action of G_i only with the effects which are explicitly mentioned as believed by G_j (e.g., the result of a communicative action having G_j as receiver). However, the treatment of the changes of the beliefs of the partner would deserve a more accurate model, as the one proposed for multiagent systems by (Hideki and Hirofumi, 2000).

Some more words must be devoted to the probability that an effect holds after the execution of a recipe R^x . Note that if a recipe R^x of G_i makes a proposition *Prop* true only with probability $p(Prop)$ the simulation of G_j 's planning phase must be carried on starting from both "possible" worlds resulting from the execution of R^x (i.e. one where *Prop* is true and one where *Prop* is false).⁵

Therefore, we simulate separately (see figure 1) what G_j would plan if *Prop* were true and if *Prop* were false; since also G_j 's recipes may involve uncertain effects, we adopted the solution of multiplying the probability of the different outcomes of G_j 's actions with the probability of G_j 's initial states in order to obtain the set of worlds representing the possible outcomes of G_j 's reactions to the plan R^x .

⁵Using as G_j 's initial world one where *Prop* has $p(Prop)$ probability to be true, would correspond to the situation in which G_j is planning with uncertainty about *Prop*.

4. Consequences for Dialog

4.1. Communication as a Consequence of Cooperation

Consider the situation where an agent has just discovered that the goal of the group is impossible to achieve: he has two alternatives, banned the idea to go on with the now impossible shared plan; he can choose to give up the shared plan and do something else. Otherwise, he can consider adopting some of his partners' goals; banned again the idea of helping the partners in doing their part (remember that the shared plan is impossible to achieve, so no help can be fruitful), there remain, however, other goals available for adoption: if the partners still have the intention to achieve the joint goal, they consequently have the goals of knowing whether they succeeded or are prevented from achieving their aim (as it follows from the definition of (individual) intention in (Cohen and Levesque, 1991)). In order to choose among the alternatives of leaving the group and of communicating to the partners the impossibility of achieving their aims, the agent resorts to anticipatory coordination. This consists of simulating what the partners will (presumably) do in both situations: in the first case, being not aware of the new state of affairs, the partners will go on in doing their part; but if the agent chooses to inform them that the shared plan has become impossible, he predicts that they will give up the (now useless) joint action (provided that communication is successful).

The same holds for the information that the joint goal has been achieved. Consider, for example, the situation where two agents, G_i and G_j are looking for a particular piece of a jigsaw: in order to speed up the search, they divide the set of pieces into two parts, and each one searches one part. If the agent G_i finds the piece, he will inform the partner G_j , to prevent him from searching all his part unsuccessfully. Or, if he comes to know that the piece has gone lost, the model predicts as well that he will inform the partner (provided, in both cases, that the cost of communicating for G_i does not override the resources wasted by G_j in his useless search).

The same reasoning applies to the goals that the partners are in charge of, and to the goals that are instrumental to the achievement of these goals. Each time an agent knows that a partner has a goal, he can infer that the partner will also have the subsidiary goals of knowing if he succeeded. If he comes to know (without any further cost) that the partner's goal holds, in his next planning phase he has to consider whether to adopt the partner's goal of knowing if the goal holds; if he adopts this goal, he will communicate to the partner that the goal holds, an action that adds a little overhead to the group utility.

For example, if G_i and G_j have the shared goal of preparing tea, and G_j is in charge of boiling the water, G_i may come to know that the water is boiling before G_j does. At this point, G_i can inform the partner that his goal holds, by performing a communicative act ("The water is boiling"), or he can go on with his activity, without informing G_j .

In terms of utility, when the agent considers the alternative of going on with his own plan and not adopting the partners' goal, he may discover that the group utility would be lower,

```

/* in input the one-level plan of agent  $G_i$  (gi) for  $\beta_k^{x,i}$  (plan-x-i-k), the identifier of
agent  $G_j$  (gj), the step in charge of  $G_j$ , i.e.  $\beta_m^j$  (action-j-m) and an initial world*/

plan-shared-actions(Gi, Plan-x-i-k, Gj, action-j-m, initial-world)
begin
  /* refinement of plan-x-i-k by selecting an alternative or adding
the decomposition of an action belonging to the plan */
  refined-plans := refine-plan (plan-x-i-k, gi, initial-world);
  final-worlds := nil;
  /* for each possible outcome of each possible alternative */
  for-each plan in refined-plans
  begin
    /* outcomes of a plan of  $G_i$  from the initial worlds
(their probability sums to one)*/
    for-each world in resulting-worlds(plan, initial-world)
    begin
      /* save the probability of the outcome of plan */
      prob := world.prob;
      /* simulate  $G_j$  planning from an outcome of plan as it were the only
possible one */
      world.prob := 1;
      primitive-plans-j := plan(action-j-m, gj, world);
      /* select best plan from  $G_j$ 's point of view:  $G_i$  considers only
 $G_j$ 's best alternative */
      chosen-plan-j := best-plan-EU(primitive-plans-j, gj, world);
      resulting-worlds := resulting-worlds(chosen-plan-j, gj, world);
      /* restore the probability of the outcomes w that come after world */
      for-each w in resulting-worlds
      begin w.prob := w.prob * prob; end
      /* the probability of worlds in final worlds will sum to one */
      final-worlds := final-worlds + resulting-worlds;
    end
    /* assign to each  $G_i$ 's alternative the expected utility from  $G_i$ 's perspective */
    plan.EU := compute-EU(plan, final-worlds, gi);
  end
  /* eliminate plans that are not promising */
  return(eliminate-plans(refined-plans, gi));
end

```

Figure 1: The function of the planner that given a plan, performs a step of refinement and discharges unpromising alternatives.

even if his own utility would be greater: the partners, being not aware of the relevant conditions, would waste their time going on in their activity or, at best, looking whether they succeeded or not.

In the previous example, if G_i does not warn G_j that the water is boiling, the consequence is that G_j will have to check again, while the water could evaporate and time could be wasted.

Communication is related to adoption also in a different sense; if an agent decides to adopt a goal of a partner, it can be useful for the group to communicate his intention to the partner. Although in many cases this intention can be easily inferred by the beneficiary by observing the adopting agent's actions, under some conditions it could be impossible for the adopting agent to display his adoption other than communicating it. If this is the case, the utility of communicating becomes higher than the utility of not

communicating, since, even if communication adds a little overhead, the second alternative could result in a most disadvantageous situation where two agents (the adopting one and the beneficiary) waste resources - or even conflict - by trying to achieve the same task.

Consider the example where G_i and G_j cooperate to change a light bulb: one of them, G_i , will climb the ladder, while the other, G_j , is in charge of fetching a new light bulb. If G_i switches off the light (because he is nearer to the switch, for example) while G_j is away, and he knows that it was G_j 's duty to do it, he should warn the partner ("I have already switched off the light for you!"). In fact, if G_j is not aware that G_i has already switched off the light, he may repeat the action, with the result of closing the circuit again, with potentially bad consequences for G_j : in this case the overhead added by the communicative action is clearly irrelevant if compared to the consequence it could provoke,

for both the beneficiary and the group.

Since agents share a group utility function, we can predict that they will (try to) avoid conflicts with other agents' intentions; in fact, performing an action that interferes with the plans of other team members decreases the utility of the whole team. But if they are aware of this possibility, they may resort to communication for coordinating the choice of single agents' plans ("Don't start painting the ceiling just now, I need the ladder to change the light bulb!"). The trigger of communication relies again on the utility function: if the partner is likely to choose an action that leads to the shared plan failure due to a conflict with the agent's intentions, the best strategy is to add a little overhead to communicate to the partner what he has planned to do.

In all the cases we have examined, the agent has to trade off the cost of communicating against the potential consequences of not communicating: i.e., communication is not compulsory, but that it is produced only if it is convenient to do so. In this way, the requirement of a more flexible communication posed by (Tambe, 1997) is solved by referring to the notion of utility for the whole group. If communication is expensive, it is not convenient for the group to waste resources in communicating. The same holds if communication is not reliable (the message gets lost) or slow: there is a probability that communication has not the desired effect (or it gets it too late). An important consequence is that, if an agent decides that is better (for the group) not to communicate, his choice does not disrupt the group: in fact, communication is not explicitly mentioned in our definition of cooperation.

Moreover, the utility-based model of cooperation allows to release the assumption of perfect communication: when evaluating the expected utility of a communicative action, the agents takes into account the probabilities that it doesn't bring about the expected result. If the action fails, however, its repetition is not automatic: again, repetition is subject to utility evaluation and this evaluation cannot but take the previous failure into account.

4.2. Communication as Cooperation

The model of cooperation we propose can predict the occurrence of grounding phenomena when the group shared goal is constituted by the exchange of information itself. Once the shared goal of communicating has been established, the goal of reaching the mutual understanding is assumed as a natural consequence of it, stemming from the communicative competence of the group members: for the dialog to proceed coherently, in fact, the interactants need to share the same interpretation of the previous part of the interaction, or better, the mutual belief must hold that their respective interpretations are reasonably aligned.

The requests for acknowledgement of understanding, for example, are useful for the group only when the content to be conveyed is sufficiently relevant for the success of the exchange and there are not irrelevant probabilities that it has not been correctly conveyed. Otherwise, if the cost of the action of requesting for an acknowledgement - and the related cost of acknowledging by the interlocutor, in the speaker's anticipation of his reaction - is expected to be higher than the benefit it produces, the action is not exe-

cuted, in favor of other, implicit grounding means (see the Principle of least effort in grounding presented in (Clark, 1996)).

The following exchange could seemingly take place in the light bulb example discussed above, due to the importance of the content conveyed by G_i in the first turn:

G_i : Be careful, I have already switched off the light

G_j : Pass me the light bulb, please

G_i : Did you understand what I said?.

Here, the explicit request for acknowledgement by G_i in the third turn is motivated by the relevance of the content of the turn to be conveyed, which has not been explicitly acknowledged by G_j in his reply (second turn).

Spontaneous acknowledgments of understanding by the interlocutor constitute the interlocutor's adoption of the speaker's goal of knowing if he succeeded in his communicative act (e.g. "I have already switched off the light" "I see"). Again, if the cost of the interlocutor adopting this goal is not balanced - according to the interlocutor's utility evaluation - by the utility of the outcomes it may produce, the acknowledgment is disregarded in favour of other, implicit grounding means and the conversation proceeds smoothly (compare the previous example with the following exchange, that we take to happen during an everyday conversation: "Do you know that John got married?" "What is the name of the wife?").

At the same time, the utility function, since it accounts for the reliability of communication, allows to model the frequency of acknowledgments and request for acknowledgments in certain interaction modalities, like telephone conversations.

The explanation of grounding that we have introduced so far is in line with the fact that interactants rarely acknowledge (or request to acknowledge) acknowledgments and notifications themselves: while in nested acknowledgments the relevance of the communicative acts decreases or at least remains the same, the probability that it doesn't succeed becomes lower, and the resource consumption increases.

Moreover, the utility function allows for accounting for the choice of actions that satisfy several goals at the same time, to different degrees: for this reason, the contextual preference for a smoother continuation of the dialog with respect to explicit grounding can be modeled.

In general, the lack of interpretation problems is, by itself, a symptom that there is a common interpretation, i.e., that participants' interpretations are reasonably aligned. This does not mean that the interpretation is really the same, but only that the potential differences fall within the standard individual differences. In absence of specific signs of misalignment, this normally makes the intrinsic utility of any active effort to check the interpretations' alignment very low, because this effort would not be compensated by the low risk of having to engage in a negotiation phase in the following, in order to restore the lost alignment.

On the contrary, the loss of dialog coherence normally means that a misunderstanding has occurred, i.e., that at least one of the two speakers has chosen a wrong interpretation of a turn (see (Ardissono et al., 1998) for a deeper

analysis). If a participant realizes that a misunderstanding has occurred, he will compare the alternative of performing a repair action for addressing and solving the misunderstanding (Schegloff, 1992) with the alternative of going on with the next turn, without executing any extra action. If the comparison between the two alternatives ends in favour of the former one, the participant will act in order to find out who is the misinterpreting agent, and correct the wrong interpretation. Intuitively, the utility function here embodies the idea that a misunderstanding is addressed if it is deemed relevant, and is at risk of posing difficulties for the subsequent interaction.

1. if G_i realizes that his G_j is the misinterpreting agent, he will plan a request for repair (“No, I mean that...”); G_j , in turn, is expected to repair his interpretation as requested, and notify the execution of the repair to G_i (“Oh, ok ...”).
2. if G_i realizes that he has misinterpreted G_j , he will plan a repair and the notification to B that now he holds the right interpretation (“Oh, you meant that ...”).

While it is outside the scope of this work to explain how shared goals are established in a general way, the model of cooperation that we propose provides a framework where the establishment of the shared goal of communicating can be explained. When communicative acts are performed within the context of a group cooperating to a non linguistic shared goal, their occurrence is motivated by the utility they yield with respect to the shared goal. Under these circumstances, we postulate that the performance of a communicative act by one of the group participants normally sets up the shared goal of communicating.

Finally, we exploit a utility-based reasoning to explain the establishment of communicative cooperation outside the context of a group cooperating to a shared goal (Boella et al., 1999): when an agent is requested to cooperate, he is relatively free to refuse cooperation, but, even if he does so, he normally cooperates at the communicative level, by informing the partner about his refusal. In fact, an agent usually has the goal of not offending his partner and the refusal of communicative cooperation is interpreted as an offensive behavior: a notification is a low-cost action, and its omission could result in a repetition of the request or in a solicitation of feedback (rather expensive actions for the partner).

For example, if an agent G_i is asked by another agent G_j to tell him the time, and G_i 's clock is broken, he will probably cooperate, at least conversationally, and reply with a communicative act of justification such as “I'm sorry, my watch is broken”, instead of ignoring the request and letting G_j believe that his request has been ignored or refused.

The anticipation of the hearer's reaction, even if a logical and not decision-theoretic framework, has been proposed in (Ardissono et al., 1999) for dealing with the choice of polite forms of speech-acts.

5. Related Work and Conclusions

As we stated in the introduction, decision and game theoretic concepts are being increasingly used in modeling

multi-agent situation, even if they still present some limitations, like, for example, the difficulty of estimating the utility functions. In particular, the idea of taking into account the benefit of the group has been advanced in (Hogg and Jennings, 2000), but in that work the advantage of the group does not take into account anticipatory coordination. On the other hand, (Gmytrasiewicz and Durfee, 1995) exploit the prediction of what the other agents will do in a manner which is very similar to our approach, but they do not have a notion of group utility and goal adoption. In the subsequent paper (Gmytrasiewicz and Durfee, 1997) the consequences of their approach for dialog is considered.

Grosz and Kraus (Grosz and Kraus, 1996) propose a formal specification of the notion of sharing a plan. They introduce the operator *Intend-that* in order to account for the commitment of each group member to the shared plan; from the intention-that the plan be performed, the agents derive that they have to avoid conflicts and to coordinate the group's behavior through communication. In addition, the definition of shared plans prescribes that agents intend that their partners are able to do their part in the plan. In our model, we have tried to obtain a similar effect by means of the interaction of the shared utility function with the mechanism of goal adoption. In particular, conflicts are avoided since a plan that interferes with the partners' actions normally makes the utility of the group decrease. The goal adoption mechanism makes an agent consider whether, by adopting the partners' goals, a gain for the group is achieved.

Finally, the ideas of goal adoption and anticipatory coordination have been put forth by (Castelfranchi, 1998) and exploited for the definition of cooperation in (Boella et al., 2000) and (Boella, 2000). A similar approach is used for modeling obligations in multi-agent systems in (Boella and Lesmo, 2000). The role of goal adoption and cooperation in communication has been analysed in (Ardissono et al., 2000); similar concepts have been used in (Boella et al., 1999) for dealing with obligations in dialog.

Acknowledgements

This work has been partially supported by the project “Conoscenza, intenzioni e comunicazione” (Knowledge, intentions, and communication) of the National Research Council and by the project “Approcci modulari all'analisi linguistica” (Modular approaches to linguistic analysis) of the Ministry for University and Scientific Research.

6. References

- L. Ardissono, G. Boella, and R. Damiano. 1998. A plan-based model of misunderstandings in cooperative dialogue. *International Journal of Human-Computer Studies*, 48:649–679.
- L. Ardissono, G. Boella, and L. Lesmo. 1999. The role of social goals in planning polite speech acts. In *Workshop on Attitude, Personality and Emotions in User-Adapted Interaction at UM'99 Conference*, pages 41–55, Banff.
- L. Ardissono, G. Boella, and L. Lesmo. 2000. Plan based agent architecture for interpreting natural language dialogue. *International Journal of Human-Computer Studies*, (52):583–636.
- G. Boella and L. Lesmo. 2000. Deliberate normative agents. In *Proc. of Autonomous Agents 2000 Workshop on Norms and Institutions.*, Barcelona.

- G. Boella, R. Damiano, L. Lesmo, and L. Ardissono. 1999. Conversational cooperation: the leading role of intentions. In *Amstelogue'99 Workshop on Dialogue*, Amsterdam.
- G. Boella, R. Damiano, and L. Lesmo. 2000. Cooperation and group utility. In N.R. Jennings and Y. Lespérance, editors, *Intelligent Agents VI — Proceedings of the Sixth International Workshop on Agent Theories, Architectures, and Languages (ATAL-99, Orlando FL)*, pages 319–333. Springer-Verlag, Berlin.
- G. Boella. 2000. *Cooperation among economically rational agents*. Ph.D. thesis, Università di Torino, Italy.
- C. Castelfranchi. 1998. Modeling social action for AI agents. *Artificial Intelligence*, 103:157–182.
- H.C. Clark. 1996. *Using Language*. Cambridge University Press.
- P. R. Cohen and H. J. Levesque. 1991. Teamwork. *Noûs*, 25:487–512.
- P. J. Gmytrasiewicz and E. H. Durfee. 1995. Formalization of recursive modeling. In *Proc. of first ICMAS-95*.
- P. J. Gmytrasiewicz and E. H. Durfee. 1997. Rational interaction in multiagent environments: Communication. In *Submitted for publication*, available at <http://www-cse.uta.edu/~piotr/www/piotr.html>.
- B. Grosz and S. Kraus. 1996. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357.
- P. Haddawy and S. Hanks. 1998. Utility models for goal-directed, decision-theoretic planners. *Computational Intelligence*, 14:392–429.
- I. Hideki and K. Hirofumi. 2000. Observability-based nested belief computation for multiagent systems and its formalization. In N.R. Jennings and Y. Lespérance, editors, *Intelligent Agents VI — Proceedings of the Sixth International Workshop on Agent Theories, Architectures, and Languages (ATAL-99)*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin.
- L. Hogg and N. Jennings. 2000. Variable sociability in agent-based decision making. In N.R. Jennings and Y. Lespérance, editors, *Intelligent Agents VI — Proceedings of the Sixth International Workshop on Agent Theories, Architectures, and Languages (ATAL-99)*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin.
- A. Ndiaye and A. Jameson. 1996. Predictive role taking in dialog: global anticipation feedback based on transmutability. In *Proc. 5th Int. Conf. on User Modeling*, pages 137–144, Kailua-Kona, Hawaii.
- E.A. Schegloff. 1992. Repair after the next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 7(5):1295–1345.
- M. Tambe. 1997. Towards flexible teamwork. *Journal of Artificial Intelligence Research*, 7:83–124.