

Conversations are Joint Activities

Brahim Chaib-draa

Laval University, Computer Science Department
Pavillon Pouliot, Ste-Foy, PQ, Canada, G1K 7P4
chaib@ift.ulaval.ca

Abstract. This paper proposes to see agent communication language (ACL) as a joint activity and not as the sum of the speaker's and hearer's (speech) acts. In this paper, a conversation in the context of ACL is viewed as a *joint activity* which can be realized as sequences of smaller actions, many of which are themselves joint actions. Social agents which participate to this joint activity have to coordinate their joint actions. In each joint act, the participants face a coordination problem: which actions are expected? The answer to this question proposed here, is based on complex notions as collective intention, joint plan, joint commitments and the notion of common ground.

1 Introduction

Multi-agent systems (MAS) are the subject of an area of research studying systems made up of multiple heterogeneous intelligent software entities (called agents) where competition coexistence or cooperation is possible between them. MAS differs from *distributed problem solving* in the sense that there is no common global goal to be solved which is known at design time; on the contrary, a multi-agent system is generally peopled by different agents having different purposes.

In recent years the interest in multi-agent systems (MAS) has grown tremendously, and today multi-agent technology is being used in a large range of important industrial application areas. These applications ranges from information management through industrial process control to electronic commerce. All these applications have one thing in common. Agents must be able to talk to each other to decide what action to take and how this action can be coordinated with others' actions. The language used for this exchange is the *agent communication language* (ACL).

Traditional approaches of ACL view conversations as the sum of speaker's and hearer's speech acts. This approach is counterintuitive in the sense it does not reflect our intuition about conversations which are *social activities*. In this paper we consider a conversation as a *joint activity* which can be realized as sequences of smaller actions, many of which are themselves joint actions. Social agents which participate to this joint activity have to coordinate their joint actions. In each joint act, the participants face a coordination problem: which actions are expected? To achieve this, agents should base their reasoning on

notions as collective intention, joint plan, joint commitments and the notion of common ground. This is the approach that we have adopted for ACL, and that we detail here.

2 Agent Communication Language (ACL): An Overview

A first attempt to come to a standardized agent communication language (ACL) came forth from the ARPA knowledge sharing project and produced KQML. In the context of this project, researchers developed two main components: (1) a representation language for the contents of messages (called Knowledge Interchange Format–KIF), which is an extension of first-order logic; and (2) a communication language KQML (Knowledge Query and Manipulation Language) which consists of a set of communication primitives aiming to support interaction among agents in MAS. KQML includes many performatives of speech acts, all assertives (i.e. when it states a fact) or directives (i.e. when it reflects command or request), which agents use to assert facts, request queries or subscribe to services. A sample KQML message has the following syntax (`tell :sender A :receiver B :content "snowing"`), that is the agent A tells to agent B that the proposition "it is raining" is true. The semantics of KQML presupposes that each has its own virtual KB (knowledge base). In these conditions, telling P corresponds to reporting that P is in its KB; asking for P is attempting to extract P from the addressee's KB, etc. Up till now KQML is the only ACL that is implemented and (widely) used (at least in the academic world). Recently, the authors of KQML gave it a semantics issued from the theoretical foundation of Searle and Vanderveken [11].

More recently another effort to come to a standard ACL has started through the FIPA initiative. This effort brings Arcol to bear on ACL, a language developed by France Télécom [2]. In Arcol the set of primitives is smaller than in KQML (primitives in Arcol can be composed) and this set also includes assertives or directives as in KQML. Arcol has a formal semantics based on Cohen and Levesque's approach on speech acts [4]. Conversely to KQML, in Arcol, agent A can tell agent B that P only if A believes that P and believes that B does not believe P. Thus, Arcol gives preconditions on communicative acts as specified by its semantics.

Although some work has been done on the semantics of individual speech acts in KQML and Arcol, little is known about the semantics of conversations and the relations between speech acts and the conversations of which they form a part. In fact, ACL must be viewed as a sort of conversation between software agents and not as a set of speech acts. In this sense, the semantics of ACL is the semantics of a conversation that cannot be reduced to the conjunction or composition of semantics of its speech acts.

In the context of multi-agent systems, conversations between agents are used to share tasks and results as in the Winograd and Flores' conversation for action (WFcfa) (shown in Fig.1). Although we have concerns about the adequacy of this for human interactions (as other researchers, it may be an adequate model for a

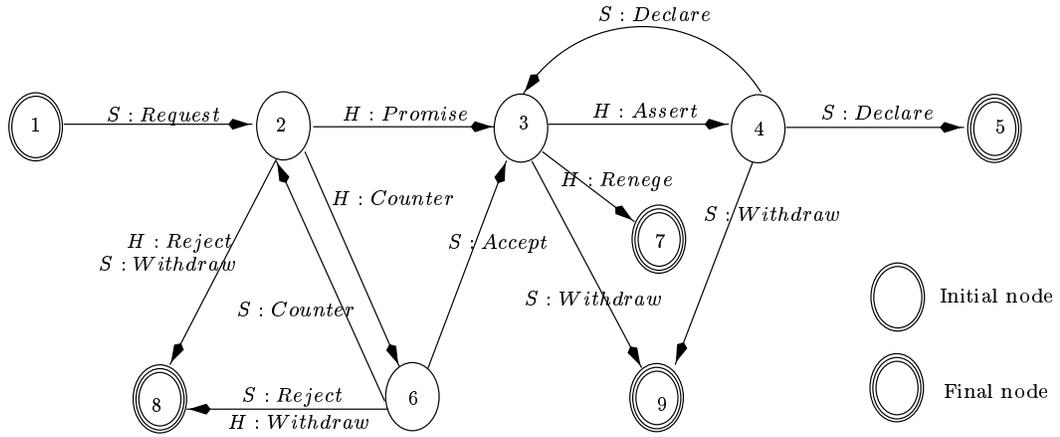


Fig. 1. Winograd and Flores' Conversation for action [18].

communication protocol among software agents. Indeed, this model is adequate for requests that software agents make of each other and that humans make of agent systems [1].

In the WFcfa model (see Fig.1), the circles (nodes) represent the states of the complex speech act, and the arcs represent speech acts that cause transitions from state to state in the complex request act. The authors of this model assert that (i) states 5, 7, 8 and 9 represent final states of the complex act, (ii) state 1 initiates the conversation and, (iii) other states (2, 3, 4 and 6) represent an intermediate state of the complex request.

Conversations between agents are also used in the contract Net where agents coordinate their activities through contracts to accomplish specific tasks. An agent acting as a *manager*, decomposes its contract (the task or problem it was assigned with) into sub-contracts to be accomplished by other *potential contractor* agents. For each subcontract the manager announces a task to the group of agents. These agents receive and evaluate the announcement and those with the appropriate resources, expertise, and knowledge reply to the manager with *bids* that indicate their ability to achieve the announced task. The manager evaluates the bids it has received and awards the task to the most suitable agent, called the *contractor*. Finally, manager and contractor exchange information together during the accomplishment of the task. In these conditions, we see in this protocol the following key steps: 1) the manager announces a task; 2) agents reply with bids; 3) the manager awards the task to the contractor; 5) the manager and the contractor exchange information about the results of the task.

Market mechanisms constitute another aspect of coordination between agents using conversations. Here, tasks are matched to agents by generalized agreement,

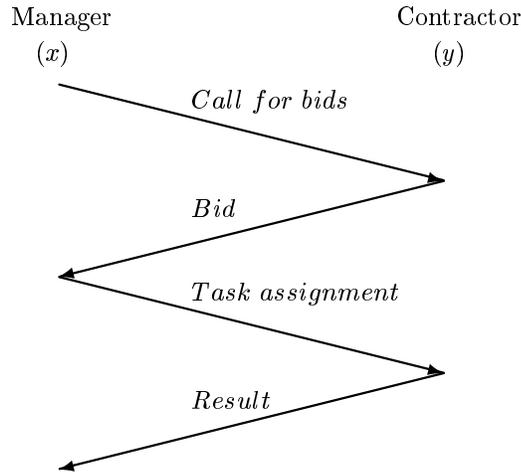


Fig. 2. Messages Exchanged in the Contract Net (after [13])

mutual selection or by using some auction mechanisms as English, Dutch, double-auction, Vickery, etc.

Another important aspect of conversations between agents turns around the “matchmaking problem”. Matchmaking is based on a cooperative partnership between information providers and consumers, assisted by an intelligent facilitator utilizing a knowledge sharing infrastructure [10]. Information providers take an active role in finding specific consumers by *advertising* their information capabilities to a matchmaker. Conversely, consumers send *requests* for desired information to the matchmaker, which in turns, attempts to identify any advertisements that are relevant to the requests and notify the providers and consumers as appropriate.

From a larger perspective, interactions between social agents may be considered as taking place within *conversations*, and conversation may be viewed as a *joint activity* [3] which can be realized as sequences of smaller actions, many of which are themselves joint actions. Social agents which participate to this joint activity have to coordinate their joint actions. In each joint act, the participants face a coordination problem: which actions are expected?

To sum up, ACLs are conversations and therefore

1. their analyze must start from conversations to speech acts;
2. their formal semantics should emphasize on social concepts (as joint activity, joint actions, collective intention) and the “commun ground” reflecting their common background;

This can help to elaborate a semantics of ACL, since there is a lack of consensus on the semantics of the communication between agents. This also can help to clarify relation between speech acts and conversations since little is known

about the semantics of conversations and the relations between speech acts and the conversations of which they form a part.

In our approach, communications between agents are considered as conversation types with goal-defined with constraints on the allowable contributions and fairly fixed turn order. These conversation types are more *structure* oriented and in this sense they contrast with conversations in general, which are more *process* oriented. The interactions in this later type of conversations are not predictable because, on the one hand, they depend on the interactions' individual histories and experiences, and on the other hand, they are continually enlarged and modified during the process of conversation. In the next section, we will detail these aspects of structure oriented vs process oriented and explain why ACL is a structure oriented.

3 Agent Communication Language is Structure Oriented

Discourses, and in particular conversations, only come into existence through the *interaction* of two or more people who may have completely diverging goals and intentions. At the beginning of a conversation it is very often not clear to any of its par participants how long it is going and what its final outcome will be. A lot of work that has appeared during the last twenty years or so in fact assumes that both text and discourses are structural entities requiring analytical tools borrowed from sentence grammar. Proponents of approaches that fall within this category claim that discourses consist of well-defined units of a lower level in very much the same way as sentences consist of clauses of phrases, phrases of words and so on. In addition there are rules which govern possible and impossible sequences of these units, just as sentence grammar distinguishes between well-formed and ill-formed sequences of words or phrases. In this case, there must be a mapping procedure which relates actual utterance, as they are produced in discourses, to their underlying actions. In fact, such a mapping procedure cannot be found because utterances may have more than one function (if they are ambiguous or more or less intentionally indeterminate as to their illocutionary force).

It should follow, then, that the structural tools which are used in sentence grammars cannot be adopted for the analysis of conversations. In general, a conversation is a complex process based on the interaction between participants. We now detail this view.

Participants in a conversation have in their brains a large set of assumptions, knowledge and beliefs. Assumptions, knowledge and beliefs about the world they live in, about their particular society, about themselves, and about their experiences. They also have assumptions and beliefs about each other. They use these assumptions and beliefs according to the relevance theory as established by Sperber and Wilson (e.g. [14,15]). In this theory, every utterance comes with a tacit guarantee of "its own optimal relevance", that is to say, the speaker always assumes that what he or she has to say will be worth the addressee's while to process. Thus, participants constantly use, enlarge and modify their assumptions,

knowledge and beliefs during the process of a conversation. Conversationalists' assumptions, knowledge and beliefs correspond to some extent to the notions of "background + network" as postulated by Searle [12] or to the notion of "commun ground" as introduced by Clark [3]. According to Searle, the network comprises the propositional beliefs that are directly necessary to understand an utterance, whereas the background comprises those assumptions and beliefs that are too basic to be needed directly for the interpretation of utterances but which are necessary if the network propositions are spelt out in details. In fact, the distinction between these two notions is not important here since both reflect assumptions and beliefs and they only differ by their degrees of manifestness. Similar to these two notions, is the "commun ground" notion introduced by Clark [3]. According to Clark, common ground is a *sine qua non* for everything we do with others as the activity of conversations that he considers as a *joint activity*.

In light of these considerations, relevance theory can account for conversation if they are taken to be processes rather than structures, processes in which the aims of the participants can change or shift. Thus, what is relevant to the participants changes continually along with the changing set of background or commun ground assumptions.

Levinson [9] introduced the useful notion of "activity type", which is more general than "discourse type" because it also includes social events to which verbal exchanges are only incidental, as for instance a game of soccer or the task of repairing a computer in a lab.

I take the notion of an activity type to refer to a fuzzy category whose focal members are goal-defined, socially constituted, bounded, events with *constraints* on participants, setting, and so on, but above all on the kinds of allowable contributions. Paradigm examples would be teaching, doctor-patient consultation, a round-table debate, a job interview, a jural interrogation, a football game, a task in a workshop, a dinner party, and so on. ([9]p. 368, his emphasis).

The notions "goal-defined" and "constraints" do not apply to all activity types or discourse type in the same manner. For instance, a doctor-patient consultation certainly has a well-defined dominant goal and there are a great number of constraints on allowable contributions. On the other hand, conversations in general have few constraints, and their dominant goal are ill-defined. There are further variables, concomitant with these as summarized in Tab. 1, taken from [7].

As we see, conversations naturally tend very much towards the left hand side of the scales in Tab. 1, whereas exchanges for an auction, interviews for jobs, or courtroom examinations are situated towards the opposite end. One obvious difference between the two is the roles of the participants, and concomitant with this, the turn order and the types of turn that may occur.

In the case of agent communication language and as we can see through the examples of section 2:

1. there are many constraints on allowable agents' contributions;
2. there are few goals and almost the time, there is a dominant goal;

Table 1. Scales characterizing unconstrained and constrained discourses (after [7]).

Process Oriented	Structure Oriented
Few constraints on allowable contribution	Many constraints on allowable contributions
Multiple and mainly ? goals	Few and mainly dominant goals
Relatively free turn order	Relatively fixed turn order
Roles of participants not clearly defined	Roles of participants clearly defined
Contribution largely determined by previous contributions	Contribution often not determined by previous contributions
Local organization principles	Global organization principles

3. all agents present in interaction are assigned clearly defined roles within the proceedings and the right to communicate and the allowable contributions are inseparably linked with the individual roles;
4. relevance can't account since the aims of participants do not change or shift.

4 Agent Communication Language is a joint activity

Searle [12] suggests that for a minimal understanding of conversation as a theoretical entity, it must be seen as an expression of *shared intentionality*. This kind of intentionality transcends the *conjunction* of the speaker's individual intentional states, because in fact it is the intentionality of an inclusive *We* (such as when pushing a table together, an act in which I push the table as an intrinsic contribution to the fact of pushing the table together). Searle puts forward the idea that *shared intentionality* (i.e., joint or collective intention, or we-intention) is an important factor in the explanation of any form of social behavior.

More precisely, Searle claims that all conversations are forms of collective intentionality. He considers this collective intentionality as a *primitive* concept which does not reduce to individual intentionality plus mutual knowledge. The high level of the collective intentionality "we are doing such and such" allows within it for individual intentionality of the form "I am doing so and so as part of our doing such and such". Thus, the individual intentionality is part of the collective intentionality. In this case, conversations are forms of collective intentionality and the "We-intention" of "We are talking about the price of X" allows for differing "I intentions", e.g. for my "I intention": "I offer you \$5 for X" and for your I-intention "I refuse your offer".

The Clark's notion of *joint project* [3] can serve to make clear this notion of joint intention. Clark started from the notion of an activity type of Levinson as previously introduced and considers language as one type of *joint activity*, one in which language plays an especially prominent role. In conversations, according to Clark, people accomplish one piece at time. They do that via joint projects: one participant projects a joint activity for all the participants to complete, and

the others take it up and complete it. Generally, *a joint project is a joint action projected by one of its participants and taken up by the others*. Joint projects require participants in conversations to commit to doing things with each other. Thus, conversely to autonomous actions where individuals have to be willing and able to do, joint project need the commitment of *all* the participants.

In these conditions, we can state:

Thesis 1: Joint intentions are joint projects + joint commitments

Thesis 2: Joint projects are plans to perform some activity jointly

Joint intentions, Joint plans (or joint projects) and joint commitments can be linked through the plan terminology by the following Tuomela's thesis [17]:

Some agents (say $A_i, \dots, A_j, \dots, A_m$) have formed the (agreement-based) joint intention to perform X iff each of them (a) has a plan to perform X jointly; (b) has communicated this acceptance to the others, and (c) because of (a) and (b) it is a true mutual belief among $A_i, \dots, A_j, \dots, A_m$ that they are jointly committed to performing X and that there is or will be a part or share (requiring at least potential contribution) of X for each agent to perform that he accordingly is or will be committed to performing

In fact, Tuomela distinguished between the joint intention and the notion of accepting and endorsing a plan to act jointly. For him, having a joint intention amounts to accepting and endorsing a plan, provided the agents' have in the simple core cases at least, communicated their acceptance to each other and have become jointly committed to carrying it out [16].

Notice that in the case of conversations process oriented the joint intention can take time to "emerge" since this joint intention is not given in advance but built up in the course of and because of the conversation. It is in fact the object of continuous negotiation and evaluation by the interlocutors. In conversation structure oriented in contrast, there is a purpose and topic from which agents can enter under the umbrella of a joint intention very easily. In the case where protocols are used in ACL, we can assume that each protocol is some sort of "joint project" on which agents have to be committed if they want to use it. In this case it is easy for agents to know which communication protocol to use.

When agents take part in joint plan (or activity), they perform a variety of joint actions. Many of these joint actions, or their part, are *communicative acts* through which they get others to understand what they mean. As these communicative acts are linked to speech acts, we should consider here a new perspective of speech acts, a perspective which is not concerned with only the sender's perspective as is the case of traditional speech acts. As new perspectives of speech acts, we should consider (1) the role of any speech act in a conversation by knowing how it is related to other acts in the conversation (particularly relations acts between acts in what speaker (or hearer) say, and relations between acts in what speaker and hearer say); (2) the generation of *all* potential illocutionary acts needed for ACL; (3) the way that speakers get their addresses to recognize the

illocutionary act they are performing; (4) the way that hearers recognize the speaker's meaning.

As our approach views ACL as a joint activity, we suggest to consider the following thesis:

Thesis 3: speaker's meaning is a type of intention that can be discharged only through joint actions.

Thesis 4: Illocutionary acts and perlocutionary acts are accomplished only as joint actions.

Searle's second constructive proposal concerns the importance for a theory of conversation of the notion of *background*. This notion is necessary for the mutual understanding by the interlocutors of conversational sequences. In order to understand a sequence, one has to embed the intrinsic semantics of utterances in a network of fundamental beliefs and knowledge which serves as the bedrock on which understanding is based. This background functions as a necessary context in any communicative interaction, and conversational relevance is necessary measured with reference to it.

Here also Clark offers similar notion that he called *commun ground*. This notion has a clearer profile, and specially makes the background notion empirically operational. For Clark, Most joint activities get realized as sequences of smaller actions, many of which are themselves joint actions. This is also the case of conversations where participants have to coordinate their joint actions. In each joint act, the participants face a coordination problem: which actions are expected? To solve this problem, Clark proposes a coordination device based on the solution that is most salient, prominent, or conspicuous with respect to the commun ground of the participants.

According to Clark, common ground is a *sine qua non* for everything we do with other—from the broadest joint activity to the smallest actions that comprise them. For two agents *A* and *B* to act jointly, they have to coordinate what they do and when they do it. And to coordinate, they have to appeal, ultimately, to their current common ground. At the same time, with every joint action *A* and *B* perform, they add to their common ground. This is how a flexible communication between human progress.

We can do the same for communication between agents, in the sense where agents enter a conversation, by presupposing certain common ground and with each joint action—each speech acts, for example—they try to add to it. To do that they need to keep track of their common ground as it accumulates increment by increment. It is clear that in the case of ACL, a common ground should include (1) rules for the coordination of joint actions; (2) conventions, obligations [5], norms and social rules; (3) shared expertise; etc. In addition to the common ground, each agent has her own or personal ground which reflect her own knowledge, assumptions and expertise.

5 Conclusion

We have presented here a new perspective of Agent Communication Language, a perspective which views it as a conversation and therefore as a social activity. Evidently, other researchers have focused on conversation considered as a social activity: (1) Cohen and Levesque have proposed persistent goal and mutual belief [4]; (2) Grosz has proposed a shared plan [6]; (3) Singh has proposed Joint commitments [13]. We have taken the same road and introduced here some new ideas taken mainly from Searle and Clark work: collective intention, joint plan, joint commitments and common ground.

Acknowledgments:

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), by the Social Sciences and Humanities Research Council (SSHRC) of Canada, by the Fonds pour la Formation des Chercheurs et l'aide à la Recherche (FCAR) du Québec and in part, by the German Research Centre for AI (DFKI).

References

1. Bradshaw, J. M., S. Dufield, B. Carpenter, and Robinson, T. KAoS: A Generic Architecture for Aerospace Applications. *Proc. of the CIKM'95- Workshop on Intelligent Information Agents*, MD, USA, 1995.
2. Breiter, P. and M. D. Sadek. A rational agent as a kernel of a cooperative dialogue system: implementing a logical theory of interaction. in *Proc. ECAI-96 Workshop Agent Theories, Architectures, and languages*, Springer-Verlag, Berlin, 1996, pp. 261-276.
3. Clark, H. H. *Using Language*, Cambridge University Press, 1996.
4. Cohen, P. and H. Levesque. Performatives in a rationally based speech act theory, *Proc. 28th Annual Meeting of the Association for Computational Linguistics*, Pittsburg, PA. 1990. pp. 79-88.
5. Dignum, F. and Linder, B. Modelling social agents: communication as Action. in J. P. Müller, M. J. Wooldridge and N. R. Jennings (Eds). *Intelligent Agents III*, Springer Verlag, 1996, pp. 205-218.
6. Grosz, B. and S. Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, vol. 86, 1996, pp. 269-357.
7. Jucker, A. H. Conversation: structure or process? in H. Parret and J. Verschueren, (eds). *(On) Searle on Conversation*, Benjamins Pub., Philadelphia, 1992, pp.7-29.
8. Labrou, Y. and T. Finin. Semantics and conversations for an agent communication language. in M. Huhns and M. Singh (eds). *Reading in Agents*, Morgan Kaufman, San Mateo, Calif., 1988. pp. 235-242.
9. Levinson, S. C. Activity types and language. *Linguistics*, 17, 1979, pp. 365-399.
10. Patil, R. S., R. Fikes, P. F. Patel-Schneider, D. McKay, T. Finin, G. Thomas and R. Neches. The DARPA knowledge sharing effort: progress report. *Proc. of the Third Int. Conf on Principles of Knowledge Representation and Reasoning*, 1992, pp. 777-787.

11. Searle, J.R. and D. Vanderveken. *Foundations of Illocutionary Logic*. Cambridge Univ. Press. 1985.
12. Searle, J.: Conversation. in H. Parret and J. Verschueren, (eds). *(On) Searle on Conversation*, Benjamins Pub., Philadelphia, 1992, pp.7-29.
13. Singh, M. *Multiagent Systems: a Theoretical Framework for Intentions, Know-How, and Communications*. Springer Verlag. Heidelberg, 1994.
14. Sperber, D. and D. Wilson. Irony and the use-mention distinction. in P. Cole (ed), *Radical Pragmatics*, Academic Press, NY, 1981, pp. 295-318.
15. Sperber, D. and D. Wilson. *Relevance*. Cambridge MA: Harvard University Press, 1986.
16. Tuomela, Raimo. *The Importance of Us: A Philosophical Study of Basic Social Notions*. Stanford Univ. Press, Stanford, CA.
17. Tuomela, Raimo. Philosophy and distributed artificial intelligence: the case of joint intention. In G.M.P. O'Hare and N. R. Jennings, (eds). *Foundations of Distributed Artificial Intelligence*, Wiley & Sons, Inc. 1996, pp. 487-503.
18. Winograd, T. and F. Flores. *Understanding Computers and Cognition: A New Foundation for Design*. Ablex, Norwood, NJ, 1986.