A Topic-Based Framework for Rational Interaction

Andreas Herzig, Dominique Longin

Conférence TALN 2000, Lausanne, 16-18 octobre 2000

A topic-based framework for rational interaction

Andreas Herzig and Dominique Longin

Institut de Recherche en Informatique de Toulouse (IRIT)
(CNRS – Université Paul Sabatier – INPT)
118 Route de Narbonne, F-31062 Toulouse Cedex 4
mailto: {herzig,longin}@irit.fr
http://www.irit.fr/ACTIVITES/EQ_ALG/

1. Introduction

Cooperative dialogue is one of the most important challenges of computer science. The background of this work are the dialogue systems developed by France Telecom R& D as instantiations of its ARTIMIS (Sadek, 1999; Sadek *et al.*, 1997; Sadek *et al.*, 1996) generic rational agent technology. Such systems allow to manage real-time cooperative dialogues in natural language.

Our framework is what we call an "intentional approach" of dialogue (Cohen & Levesque, 1990a; Sadek, 1991; Sadek, 1992; Rao & Georgeff, 1992). This approach is based on theories of Intentionality (Searle, 1983; Bratman, 1987). Within these theories, an agent is represented by its "mental state", which is a set of informations. This set contains the different mental attitudes about the world the agent has: beliefs, goals, intentions... These theories are at the base of what is called "BDI-architectures" (for belief, desire and intention) in the literature.

Intentional approaches are defined within twofaced formal theories : *rational balance* and *rational interaction*. The first theory describes, through properties of mental attitudes and action, the relationships that must be maintained as true (the relationships between the different mental attitudes of an agent firstly, and between these mental attitudes, plans and actions secondly). The second theory characterizes the inter-agent relationship within a multiagent environment (communication, cooperation, ...). Agents built on these twofaced theories are called *rational agents*.

In this paper, we focus on the *belief change process*, viz. the ability to take into account the dynamics of the world.

The difficulties are highlighted by the following vending dialogue between the system s and a user u, where agents make mistakes, change their mind, and misinterpret.

s_1 : *Hello. What do you want?*
u_1 : *A first class train ticket to Paris, please.*
s_2 : *150 €, please.*
u_2 : *Oups... A second-class train ticket, please.*
s_3 : *100 €, please.*
u_3 : *Can I pay the 80 € by credit card?*
s_4 : *The price isn't 80 € but 100 €. Yes, you can pay by credit card.*
u_4 : ...

Firstly, we present our topic-based approach of belief change (Sect. 2). Secondly, we present related works (Sect. 3). And finally, we give some future works (Sect. 4).

2. A topic based approach

Focussing on the evolution of the system's beliefs, we aim at a semantics having both a complete axiomatization and an associated automated deduction procedure (Herzig & Longin, 2000). This has motivated several choices, in particular a Sahlqvist-type possible worlds semantics (1975), for which general completeness results exist, and a notion of intention that is primitive (contrarily to the complex constructions in the literature). Intentions have a non-normal modal logic, reflecting that they are not closed under conjunction and implication. They can nevertheless be reduced to the Sahlqvist framework.

In our approach, we proceed in two steps: the hearer always accepts the indirect and intentional effects, but not all of their consequences. Their acceptation is determined by the *speaker's competence*. For example, after the user's u_2, s accepts the new class, because s considers u to be competent at classes. And s rejects the price after the user's u_3, because s does not consider u to be competent at ticket prices.

Which mental attitudes of the hearer can 'survive' the performance of a speech act α? We consider that if there exists a *relation of influence* of α towards an attitude, then the latter cannot be preserved in the new mental state. In our example, the system's belief about the old transport class cannot be preserved through u_2, because informing about classes influences the hearer's beliefs about classes. On the other hand, the destination is not influenced by u_2, and can thus be preserved.

All this presupposes that we are able to determine the competence of an agent and the influence of a speech act. We base both notions on the notion of *topics*. This is a natural and intuitively appealing concept, and it will allow us to fine-tune the consommation of speech acts.

Topics are well studied in linguistics and philosophy. Epstein (1990) associates to a formula its *subject matter*, and defines two formulas as being related if they have some subject matter in common. Generalizing his idea, we associate a set of topics to every agent i, speech act α, and formula A. Then we consider that i is competent at a topic if and only if that topic is associated to i. And a speech act α influences a topic if that topic is associated to α.

3. Related works

Cohen&Levesque (1990a; 1990b) have laid the bases of a general theory of rational interaction where a conversation theory must explain the dialogue coherence by the way of spokesper-

sons' mental states. The speech act theory is viewed as a part of a (more general) theory of action. The acceptance criterion is the sincerity of the speaker from the point of view of the hearer. Their theory allows the agent to either reject the input (if the speaker is believed to be insincere), or change his beliefs and adopt it. But in the latter case, the theory suffers from the well-known frame problem: all previous beliefs are abandoned.

Perrault (1990) tries to solve the frame problem using Reiter's default logic. His theory suffers from some problems. In particular, Perrault's agents never question old beliefs, and can only expand their mental state (in the sense of AGM revision).

Appelt&Konolige (1989) advocate the use of a hierarchic autoepistemic logic in the aim to control the order of application of autoepistemic rules. Through a built-in qualitative notion of competence, an agent can accept some informations, and reject some others. Thus, an agent can change its beliefs, but the formalism is relatively complicate, and some non intuitive side effects follow from this framework.

Sadek's framework (1991; 1994) is based on a dynamic doxastic logic. A generic belief-change strategy allows the agent to choose among three strategies: accept the input, change its beliefs and accept the input, and change its belief and reject the input. But from a logical point of view, this is an open choice: it is no said how to determine which strategy must be chosen.

More details and formal analysis can be found in (cf. (Longin, 1999; Longin & Sadek, 2000)).

4. Perspectives

We are currently extending our framework towards the handling of indirect speech acts (Herzig *et al.*, 2000; Faure *et al.*, 2000b; Faure *et al.*, 2000a).

We want to study more deeply the effects of speech acts on the intentions of the speaker. Indeed, after the performance of a speech act, the speaker has three options: first, he might consider that the hearer has interpreted the act as intended; then he must abandon the intention which made him perform the act. Second, he might keep this intention and wait until he gets feedback confirming that the hearer has indeed made the right interpretation. The third option is that he is in an intermediate state of ignorance, in which case his intention leading to his act should be preserved, too.

References

APPELT D. & KONOLIGE K. (1989). A nonmonotonic logic for reasoning about speech acts and belief revision. In M. REINFRANK, J. DE KLEER, M. GINSBERG & E. SANDEWALL, Eds., *Proc. Second Int. Workshop on Non-Monotonic Reasoning*, vol. 346 of *LNAI*.

BRATMAN M. E. (1987). *Intention, Plans, and Practical Reason*. Harvard University Press.

COHEN P. R. & LEVESQUE H. J. (1990a). Intention is choice with commitment. *AI*, **42**(2–3).

COHEN P. R. & LEVESQUE H. J. (1990b). Rational interaction as the basis for communication. In P. R. COHEN, J. MORGAN & M. E. POLLACK, Eds., *Intentions in Communication*. MIT Press.

EPSTEIN R. L. (1990). *The Semantic Foundations of Logic*. Kluwer Academic Publishers.

FAURE R., HERZIG A. & LONGIN D. (2000a). Actes indirects et évolution des croyances. In *Actes des 5èmes journées nationales du PRC-GDR Intelligence Artificielle*. Poster session.

FAURE R., HERZIG A. & LONGIN D. (2000b). *Une formalisation de la détection des actes de langage indirects*. Tech. Rep. To appear. cf. `www.irit.fr/ACTIVITES/EQ_ALG`.

HERZIG A. & LONGIN D. (2000). Belief dynamics in cooperative dialogues. *Journal of Semantics*. To appear.

HERZIG A., LONGIN D. & VIRBEL J. (2000). Towards an analysis of dialogue acts and indirect speech acts in a BDI framework. In M. POESIO & D. TRAUM, Eds., *Proc. Fourth Int. Workshop on the Semantics and Pragmatics of Dialogue (Götelog-2000)*, Göteborg University, Sweden.

LONGIN D. (1999). *Interaction rationnelle et évolution des croyances dans le dialogue : une logique basée sur la notion de topique*. PhD thesis. `www.irit.fr/ACTIVITES/EQ_ALG/`.

LONGIN D. & SADEK D. (2000). Dialogue et dynamique des croyances. In *Actes des 5èmes journées nationales du PRC-GDR Intelligence Artificielle*.

PERRAULT C. R. (1990). An application of default logic to speech act theory. In P. R. COHEN, J. MORGAN & M. E. POLLACK, Eds., *Intentions in Communication*. MIT Press.

RAO A. S. & GEORGEFF M. P. (1992). An abstract architecture for rational agents. In B. NEBEL, C. RICH & W. SWARTOUT, Eds., *Proc. Third Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'92)*, p. 439–449: Morgan Kaufmann Publishers.

SADEK D. (1999). Design considerations on dialogue systems: From theory to technology - the case of Artimis -. In *Proc. ESCA TR Workshop on Interactive Dialogue Systems (IDS'99)*, Kloster Irsee, Germany.

SADEK M. D. (1991). *Attitudes mentales et interaction rationnelle : vers une théorie formelle de la communication*. PhD thesis, Université de Rennes I, Rennes, France.

SADEK M. D. (1992). A study in the logic of intention. In B. NEBEL, C. RICH & W. SWARTOUT, Eds., *Proc. Third Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'92)*, p. 462–473: Morgan Kaufmann Publishers.

SADEK M. D. (1994). Towards a theory of belief reconstruction: Application to communication. *Speech Communication Journal'94, special issue on Spoken Dialogue*, **15**(3–4).

SADEK M. D., BRETIER P. & PANAGET F. (1997). ARTIMIS: Natural dialogue meets rational agency. In M. E. POLLACK, Ed., *Proc. 15th Int. Joint Conf. on Artificial Intelligence (IJCAI'97)*, p. 1030–1035: Morgan Kaufmann Publishers.

SADEK M. D., FERRIEUX A., COZANNET A., BRETIER P., PANAGET F. & SIMONIN J. (1996). Effective human-computer cooperative spoken dialogue: The AGS demonstrator. In *Proc. Int. Conf. on Spoken Language Processing (ICSLP'96)*, Philadelphia, USA.

SAHLQVIST H. (1975). Completeness and correspondence in the first and second order semantics for modal logics. In S. KANGER, Ed., *Proc. 3rd Scandinavian Logic Symposium*, vol. 82 of *Studies in Logic*.

SEARLE J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge Univ. Press.