KNOWLEDGE REPRESENTATION FOR HUMAN-MACHINE INTERACTION

Mare Koit, Tiit Roosmaa and Haldur Õim
Institute of Computer Science & Institute of Estonian and General Linguistics, University of Tartu, J. Liivi 2, Tartu, Estonia
{mare.koit, tiit.roosmaa}@ut.ee, haldur.oim@ut.ee

Keywords: Dialogue Model, Reasoning Model, Conversation Agent, Knowledge Representation.

Abstract: The paper describes a computational model that we are implementing in an experimental dialogue system. Conversation process is modelled where one participant is trying to influence his/her partner to agree to do an action. Our goal is to model natural dialogue where computer as a dialogue participant follows norms and rules of human-human communication. We have worked on different aspects of developing a model of dialogue, including its computer realisation in the lines of BDI model. The main specific traits of our model are: 1) taking into account the "naïve" common-sense reasoning as the basis of dialogue, 2) modelling dialogues where the goal of the initiator is to impose the partner to do a certain action. In the paper we concentrate on the use of frames as the knowledge representation formalism in the dynamic context of dialogue. As a practical realisation of the model we have in view a computer program which we call communication trainer.

1 INTRODUCTION

We are dealing with interactions where the goal of one of the participants is to get the partner to carry out a certain action. Such dialogue can be considered as rational behaviour of conversation agents which is based on beliefs, desires and intentions of agents, at the same time being restricted by their resources (Webber, 2001; Jokinen, 2009).

Because of this, we have modelled the reasoning processes that people supposedly go through when working out a decision whether to do an action or not. In a model of conversation agent it is necessary to represent its cognitive states as well as cognitive processes. One of the most well-known models of this type is the BDI model (Allen, 1994, Boella and van der Torre, 2003). A framework for argumentation-based negotiation is proposed in (Amgoud et al., 2007). In this paper, we will develop the model considered in (Koit and Õim, 2000, 2004).

2 MODELLING THE COMMUNICATION PROCESS

Let us consider conversation between two agents - A (he) and B (she). In the goal base of one participant (let it be A) a certain goal G^A related to B´s activities gets activated and triggers in A a reasoning process. In constructing his first turn A must plan the dialogue acts and determine their verbal form as a turn r_1. This turn triggers a reasoning process in B where two types of procedures should be distinguished: the interpretation of A´s turn and the generation of her response r_2. B´s response triggers in A the same kind of reasoning cycle in the course of which he has to evaluate how the realization of his goal G^A has proceeded, and depending on this he may activate a new sub-goal of G^A, and the cycle is repeated: A builds a new turn r_3. Dialogue comes to an end, when A has reached or abandoned his goal.

2.1 Model of Conversation Agent

A conversation agent is a program that consists of six (interacting) modules (cf. Koit and Õim, 2004):

(PL, PS, DM, INT, GEN, LP),

where PL - planner, PS - problem solver, DM - dialogue manager, INT - interpreter, GEN - generator, LP - linguistic processor. Conversation agent uses in its work goal base GB and knowledge base KB. A necessary precondition of interaction is existence of shared (mutual) knowledge of agents.

2.2 Reasoning Model

After A has expressed his intention (that B does D), B can respond with agreement or rejection, depending on the result of her reasoning. We want to model a "naïve" theory of reasoning that people themselves use when they are interacting with other people and trying to predict and influence their decisions.

The reasoning model consists of two parts: 1) a model of human motivational sphere; 2) reasoning schemes. In the motivational sphere three basic factors that regulate reasoning of a subject concerning D are differentiated. First, subject may wish to do D, if pleasant aspects of D for him/her overweight unpleasant ones; second, subject may find reasonable to do D, if D is needed to reach some higher goal, and useful aspects of D overweight harmful ones; and third, subject can be in a situation where (s)he must (is obliged) to do D - if not doing D will lead to some kind of punishment. We call these factors WISH-, NEEDED- and MUST-factors, respectively.

It is supposed here that the dimensions pleasant/unpleasant, useful/harmful have numerical values and that in the process of reasoning (weighing the pro- and counter-arguments) these values can be summed up. For examplee, for the characterisation of pleasant and unpleasant aspects of some action there are specific words which can be expressed quantitatively: *enticing, delightful, enjoyable, attractive, acceptable, unattractive, displeasing, repulsive* etc.

We have represented the model of motivational sphere of a subject by the following vector of weights:

w = (w(resources), w(pleasant), w(unpleasant), w(useful), w(harmful), w(obligatory), w(prohibited), w(punishment-for-doing-a-prohibited-action), w(punishment-for-not-doing-an-obligatory-action)).

In the description, w(pleasant), etc. means weight of pleasant, etc. aspects of D.

The second part of the reasoning model consists of reasoning schemes that supposedly regulate human action-oriented reasoning. The reasoning proceeds depending on the determinant which triggers it (WISH, NEEDED or MUST). As an example, let us present a reasoning procedure.

```
// Reasoning triggered by NEEDED-determinant
Presumption: w(useful) > w(harmful) //
1. Are there enough resources for doing D?
2. If not then do not do D.
3. Is w(pleasant) > w(unpleasant)?
4. If not then go to 10.
5. Is D prohibited?
6. If not then do D.
7. Is w(pleasant) + w(useful) > w(unpleasant) + w(harmful) + w(punishment-for-doing-a-prohibited-action)?
8. If yes then do D.
9. Otherwise do not do D.
10. Is D obligatory?
11. If not then do not do D.
12. Is w(pleasant) + w(useful) + w(punishment-for-not-doing-an-obligatory-action) > w(unpleasant) + w(harmful)?
13. If yes then do D.
14. Otherwise do not do D.
```

3 KNOWLEDGE REPRESENTATION

3.1 World Knowledge

We are using frames for representing world knowledge in our system. Let us consider the following situation: A makes B a proposal to do an action D. For example, Mary proposes John to make a potato salad for the party.

There is the frame ACTION in our system:

```
ACTION
     RESOURCES
     ACTOR
     ACT: a sequence of elementary acts
     SETTING: ACTOR has RESOURCES
     GOAL
     CONSEQUENCE
```

The frame ACTION has sub-frames, e.g.:

```
PREPARING-POTATO-SALAD
      SUP: ACTION
      RESOURCES:
           Components: boiled potato, boiled egg, pickled cucumber, hacked onion, sour cream, salt, bowl
           Skills: take, chop up, mix, decorate, add
           Time: 30 minutes
      ACT: take Components; chop up potato, egg, cucumber; mix in bowl; decorate with onion; add salt
      GOAL, CONSEQUENCE: potato salad
```

3.2 Communication Knowledge

We are using two kinds of knowledge about communication: 1) descriptions of dialogue acts (proposal, question, argument, etc.), and 2) communication algorithms - communicative strategies and tactics.

3.2.1 Dialogue Acts

The dynamic parts of dialogue acts work for a coherent dialogue - there are limited sets of dialogue acts that can be come after the current act.

3.2.2 Communicative Strategies and Tactics

A communicative strategy is an algorithm used by a participant for achieving his/her goal in interaction.

Communication takes place in a communicative space which is determined by a number of coordinates that characterize the relationships of participants. Communication can be collaborative or confrontational, personal or impersonal; it can be characterized by the social distance between participants; by the modality (friendly, ironic, hostile, etc.) and by intensity (peaceful, vehement, etc.).

The choice of communicative tactics depends on the point of the communicative space in which the participants place themselves. The values of the coordinates are again given in the form of numerical values.

The participant A can realize his communicative strategy in different ways: stress pleasant aspects of D (i.e. entice B), stress usefulness of D for B (i.e. persuade B), stress punishment for not doing D if it is obligatory (threaten B). We call communicative tactics these concrete ways of realization of a communicative strategy. The participant A, trying to direct B´s reasoning to the positive decision (to do D), proposes various arguments for doing D while B, when opposing, proposes counter-arguments.

There exist three tactics for A in our model connected with three reasoning procedures (WISH, NEEDED, MUST). By tactics of *enticing* the reasoning procedure WISH, by tactics of *persuading* the procedure NEEDED and by tactics of *threatening* the procedure MUST will be tried to trigger in the partner.

The participant A when implementing a communicative strategy uses a partner model - a vector w^{AB} - which includes his imagination about weights of the aspects of the action D for B. The more A knows about B the more similar is the vector w^{AB} with the vector w^B of the motivational sphere of the partner B. We can suppose that A has sets of statements for influencing the weights of different aspects of D for the partner B: $\{st^A_i\text{-}asp_j,\ i=1, \ldots, n^A_{asp\ j};\ j=1, \ldots, n\}$ where asp_j is the j-th aspect of D and n is the number of different aspects. All the statements have their weights as well.

For illustration, let us present a schematic description of the tactic of persuasion, based on the reasoning procedure NEEDED. B may verbalize her rejection to do D bringing out a certain statement about a certain aspect of D (e.g. if B says *I do not have enough time* then she indicates that resources are missing for doing D). We can suppose that B has a set of statements $\{st^B_i\text{-}asp_j,\ i=1,\ldots,n^B_{asp\ j};\ j=1,\ldots, n\}$ for indicating the aspect which weight caused her rejection where asp_j is the j-th aspect of D and n is the number of aspects.

```
// Persuasion: A persuades B to do D //
   WHILE B is rejecting AND A is not giving up DO
     CASE B´s answer OF
      stB-resources //no resources//:
   IF there are statements stA-resources THEN present a statement stA-resources_i in order to point at the possibility to gain the resources, at the same time showing that the cost of gaining these resources is lower than the weight of the usefulness of D. //The expected result:
     wB(resources):=wB(recourses)+w(stA-resources_i)//
      ELSE exit //there are no more statements, give up//
      stB-harmful //much harm//:
   IF there are statements stA-harmful THEN present a statement stA-harmful_i to decrease the value of harmfulness in comparison with the weight of usefulness
   //The expected result:
wB(harmful):=wB(harmful)-w(stA-harmful_i)//
     stB-unpleasant //much unpleasant//:
   IF there are statements stA-unpleasant THEN present a statement stA-unpleasant_i in order to downgrade the unpleasant aspects of D as compared to the useful aspects of D
   //The expected result:
wB(unpleasant):=wB(unpleasant)-w(stA-unpleasant_i)//
      stB-punishment-for-doing-a-prohibited-action //D is prohibited and the punishment is great//:
   IF there are statements stA-punishment-for-doing-a-prohibited-action THEN present a statement stA-punishment-for-doing-a-prohibited-action_i in order to downgrade the weight of punishment as compared to the usefulness of D
   //The expected result:
wB(punishment-for-doing-a-prohibited-action):=wB(punishment-for-doing-a-prohibited-action) - w(stA- punishment-for-doing-a-prohibited-action_i)//
      stB-pleasant //little pleasant//:
   IF there are statements stA-pleasant THEN present a statement stA-pleasant_i in order to stress pleasantness
   ELSE IF there are statements stA-unpleasant THEN present a statement stA-unpleasant_i in order to downgrade unpleasantness
      stB-obligatory //not obligatory; in such a case, B´s reasoning finished on the step 11, see above//
   IF there are statements stA-pleasant THEN present a statement stA-pleasant_i in order to stress the pleasant aspects of D
   ELSE IF there are statements stA-unpleasant THEN present a statement stA-unpleasant_i in order to downgrade the unpleasant aspects of D
     END CASE
   IF there are statements stA-useful THEN present a statement stA-useful_i in order to stress usefulness
   //The expected result:
wB(useful):=wB(useful)+w(stA-useful_i)//
   ELSE exit //give up//.
```

4 DISCUSSION

When A tries to influence B in order to bring her to a decision, A uses several statements to increase the weights of the positive aspects and to decrease the weights of the negative aspects of the action D under consideration.

If B indicates a certain aspect which does not allow her to do D then A simply can choose a statement for attacking this aspect. If B does not indicate a certain reason of rejection then A only can stress the usefulness.

Let us consider an interaction where A is the computer, and B - the user. When starting a dialogue the computer chooses a point in the communicative space and a communicative tactic and generates such a partner model w^{AB} (a set of weights) that a reasoning procedure will give a positive decision. Let us consider a brief example where the action D is "to prepare a potato salad". A chooses a co-operative and personal character of communication, a short distance between participants and the neutral intensity (that means that A and B are friends), and generates such a partner model that the reasoning procedure NEEDED will give a positive decision. A will implement the tactic of persuasion. The computer composes exemplars of the frames PREPARING-POTATO-SALAD and PROPOSAL.

A (computer): Please prepare a potato salad.

B (user): I do not have enough time.

The computer must correct the value of w(resources) in the partner model and chooses a dialogue act ARGUMENT.

A: I will help you.

B: It is very hot in the kitchen.

The user pointed out the harmfulness of the action. Thus the weight of w(harmful) will be corrected in the user model.

A: My kitchen has a good ventilation.

etc.

An experimental dialogue system is implemented which in interaction with a user can play the role of both A or B. At the moment the computer operates with semantic representations of linguistic input/output only, the surface linguistic part of interaction is provided in the form of a list of ready-made and classified utterances which are used both by the computer and user.

5 CONCLUSIONS

The main specific traits of our model are: 1) taking into account the "naïve" common-sense reasoning as the basis of dialogue, 2) modelling dialogues where the initiator´s goal is to impose the partner to do a certain action. We are continuing our work in the following directions: 1) refining the reasoning model, 2) developing linguistic knowledge, 3) analysis of human-human dialogues in the Estonian dialogue corpus in order to verify the model.

ACKNOWLEDGEMENTS

This work was supported in part by the Estonian Science Foundation (grant No 7503) and the Estonian Ministry of Research and Education (the projects No SF0180078s08 and EKKTT09-57).

REFERENCES

Allen, J., 1994. *Natural Language Understanding.* Second Edition. The Benjamins/Cummings Publ. Co. Redwood City etc., 678 p.

Amgoud, L., Dimopoulus, Y., Moraitis, P., 2007. A Unified and General Framework for Argumentation-based Negotiation. In *Proc. of AAMAS´07 - the 6th international joint conference on Autonomous agents and multiagent systems*, http://portal.acm.org/-citation.cfm?doid=329125.1329317

Boella, G., van der Torre, L., 2003. BDI and BOID Argumentation. In *Proc. of CMNA-03* http://www.computing.dundee.ac.uk/staff/creed/research/previous/cmna/finals/boella-final.pdf

Jokinen, K., 2009. *Constructive Dialogue Modelling. Speech Interaction and Rational Agents.* Wiley, 160 p.

Koit, M., Õim, H., 2004. Argumentation in the Agreement Negotiation Process: A Model that Involves Natural Reasoning. In *CMNA-04 - Proc. of the Workshop W12 on Computational Models of Natural Argument. 16th European Conference on Artificial Intelligence*, 53-56. Valencia, Spain.

Koit, M., Õim, H., 2000. Reasoning in interaction: a model of dialogue. In *TALN 2000 - 7th Conference on Automatic Natural Language Processing*, 217-224. Ed. E. Wehrli. Lausanne, Switzerland.

Webber, B., 2001. Computational Perspectives on Discourse and Dialogue. In *The Handbook of Discourse Analysis*, 798-816. Eds. D. Schiffrin, D. Tannen, H. Hamilton. Blackwell Publishers Ltd.