CONVERSATIONAL AGENT IN ARGUMENTATION
A Model and Evaluation on a Dialogue Corpus

Mare Koit

Institute of Computer Science, University of Tartu, 2 J. Liivi Str, Tartu, Estonia

Abstract: Communication between two participants, *A* and *B*, is considered, where *A* has a communicative goal that his/her partner, *B*, will make a decision to perform an action *D*. A computational model of argumentation is developed which includes reasoning. Communicative strategies and tactics used by participants for achieving their communicative goals are considered. A simple dialogue system (conversational agent) is implemented which can optionally play the role of *A* or *B* using classified sets of pre-defined Estonian sentences. For further evaluation of the model and with the aim to develop the dialogue system, the analysis of the Estonian Dialogue Corpus is carried out. Calls of sales clerks who persuade clients to take training courses of an educational company are analysed. The calls end mostly with the postponement of the decision therefore the sales clerks do not achieve their communicative goal.

1 INTRODUCTION

There are many dialogue systems (DS) that interact with a user in a spoken natural language and help him/her to solve practical problems, e.g. to book flights, to get information about bus or train timetables, to detect computer faults, etc (McTear, 2004). Usually, these tasks do not include argumentation. Rather practical dialogue is implemented in such systems. On the other hand, there are tasks and situations where not only information search, but also argumentation is required.

Analysis of human-human dialogues can provide information about their structure and linguistic features with the purpose of developing a DS. Some of the well-known dialogue corpora are the HCRC Maptask, TRAINS, VERBMOBIL (McTear, 2004).

Our research is based on the Estonian Dialogue Corpus (EDiC). We investigate the conversations where the goal of one participant, *A*, is to get another participant, *B*, to carry out a certain action *D*. This type of dialogue forms one kind of so-called agreement negotiation dialogues. In this paper, we consider the dialogues where sales clerks of an educational company call another institution (a manager or another responsible person) and offer courses of their company. We may expect that a sales clerk tries to influence the partner in such a way that s/he decides to book a course for the employees of his/her institution. We are looking for ways of argumentation both of the sales clerks and the clients. Our further goal is to develop a DS which participates in an agreement negotiation dialogue with a user in a natural language, optionally performing the roles of *A* or *B*. Because of this, we have modelled the reasoning processes that people supposedly go through when working out a decision whether to perform an action or not.

2 CONVERSATIONAL AGENT

Let us consider a conversational agent as a program that consists of six modules:

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

where PL – planner, PS – problem solver, DM – dialogue manager, INT – interpreter, GEN – generator, LP – linguistic processor. PL directs the work of both DM and PS, where DM controls communication process and PS solves domain-related tasks. The task of INT is to make semantic analysis of partner's utterances and that of GEN is to generate semantic representations of agent's own contributions. LP carries out linguistic analysis and generation.

Conversational agent uses in its work goal base GB and knowledge base KB which consists of four

components: (KB_W, KB_L, KB_D, KB_S), where KB_W contains world knowledge, KB_L – linguistic knowledge, KB_D – knowledge about dialogue and KB_S – knowledge about interacting subjects. KB_D contains definitions of dialogue acts and algorithms that are applied to reach communicative goals – communicative strategies and tactics. KB_S contains knowledge about evaluative dispositions of participants towards the action(s) (e.g. what do they consider as pleasant or unpleasant, useful or harmful), and, on the other hand, algorithms that are used to reason about actions.

In this paper, we concentrate on the parts KB_S and KB_D of the knowledge base – the reasoning model which uses a model of the motivational sphere of an agent who is reasoning to do an action or not, and communicative strategies and tactics used by agents in order to achieve their communicative goals.

2.1 Reasoning Model

If a conversation in a natural language takes place between two agents – A and B – then 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 a reasoning process in A. In constructing his/her first utterance A must plan the dialogue acts and determine their verbal form as an utterance r_1. This utterance triggers a reasoning process in B where two types of procedures should be distinguished: the interpretation of A's utterance and the generation of his/her response r_2. B's response triggers in A the same kind of reasoning cycle in the course of which s/he has to evaluate how the realization of his/her goal G^A has proceeded. Depending on this s/he may activate a (new) sub-goal of G^A, and the cycle is repeated. A dialogue comes to an end, when A has reached his/her goal or abandoned it.

In general, our reasoning model follows the ideas realised in the BDI model (Allen, 1994). Our model consists of two functionally linked parts (Koit and Õim, 2004): (1) a model of a motivational sphere of a subject who is reasoning to perform an action D or not, and (2) reasoning procedures. We represent the model of motivational sphere by the vector of "weights" of different aspects of the action (these are e.g. presence of the resources for doing D, pleasantness of D, etc.): $w = (w(are\text{-}resources), w(pleasantness), w(unpleasantness), w(usefulness), w(harmfulness), w(is\text{-}obligatory), w(is\text{-}prohibited), w(punishment\text{-}for\text{-}doing\text{-}a\text{-}prohibited\text{-}action), w(punishment\text{-}for\text{-}not\text{-}doing\text{-}an\text{-}obligatory\text{-}action))$.

In the vector, the components $w(pleasantness)$, $w(usefulness)$, etc. mean weights of different aspects of D. For simplicity, it is supposed that the aspects have numerical values and that in the reasoning process their values can be summed up. In this way, the model considers a conversational agent from the standpoint of an action. Here $w(resources) = 1$, if the agent has resources necessary to do D (otherwise 0); $w(obligatory) = 1$, if D is obligatory for the reasoning subject (otherwise 0); $w(prohibited) = 1$, if D is prohibited (otherwise 0). The values of other weights are non-negative natural numbers.

The second part of the reasoning model consists of reasoning schemes that regulate human action-oriented reasoning. A reasoning scheme represents steps that the subject goes through in his/her reasoning process; these consist in computing and comparing the weights of different aspects of D; and the result is the decision – to do D or not.

Three basic factors that regulate reasoning of a subject concerning D are differentiated: his/her wishes, needs and obligations (Õim, 1996): (1) subject may wish to do D, if pleasant aspects of D for him/her outweigh unpleasant ones; (2) subject may find useful to do D, if D is needed to reach some higher goal, and usefulness of D outweighs harmfulness; and (3) 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. Respectively, there are three reasoning procedures (WISH, NEEDED and MUST) in our model depending on the factor that triggers the reasoning. Each procedure represents the steps that a subject goes through in the reasoning process, computing and comparing weights of different aspects of D (Koit et al., 2009).

2.2 Communicative Strategies and Tactics

Communication takes place in so-called communicative space. The communicative space is determined as an N-dimensional space where coordinates characterize the relationships of participants (in our model, $N=5$). For example, communication can be measured on the scales personal-impersonal, collaborative-confrontational.

A communicative strategy is an algorithm which is used by a communication participant to achieve his/her communicative goal. A communicative strategy for A (who has the goal that the partner B decides to perform D) can be represented as the following algorithm.

```
1. Choose an initial point in the
   communicative space.
2. Choose a communicative tactic.
3. Implement the tactic to generate a
   utterance: inform the partner of the
   communicative goal (decision to do
   an action D).
```

```
4. Did the partner agree to do D? If
   yes then finish (the communicative
   goal has been achieved).
5. Give up? If yes then finish (the
   communicative goal has not been
   achieved).
6. Change the point in the
   communicative space? If yes then
   choose a new point.
7. Change the communicative tactic? If
   yes then choose a new tactic.
8. Implement the tactic to generate a
   utterance (an argument) for doing D.
9. Go to 4.
```

The participant *A* can realize his/her communicative strategy in different ways (using different arguments for doing *D*): stress pleasantness of *D* (i.e. *entice B*), stress its usefulness (i.e. *persuade B*), or stress punishment for not doing *D* if it is obligatory (i.e. *threaten B*). We call these certain ways of realization of a communicative strategy communicative tactics. That can be considered as argumentation: *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.

These three tactics are connected with the reasoning procedures WISH, NEEDED, and MUST, respectively. The general idea underlying the tactics is that *A* proposes arguments for pleasantness of *D* (when enticing), usefulness of *D* (when persuading) and punishment of not doing *D* (when threatening) trying to keep the weight of pleasantness, usefulness of doing *D* or punishment of not doing an obligatory *D* high enough and the possible values of other aspects brought out by *B* low enough so that it would bring *B* to the decision to do *D* (cf. Koit et al., 2009).

The tactics for *B* are *collaboration* and *antagonism*. In the first case, *B* is interested in doing *D* and, in collaboration with *A*, is looking for arguments that support his/her positive decision. In the second case, *B* only uses arguments against *D*, his/her goal is opposite with *A*'s (like in two player games). Both *A* and *B* can implement a *mixed strategy* – change their communicative tactics during a conversation.

3 CORPUS ANALYSIS

In the following, we carry out corpus analysis, in order to evaluate the communicative tactics in our model. For that, 30 phone calls are taken from the EDiC where sales clerks of an educational company offer different courses to clients. The action *D* is 'to take the offered course'. In the case of institutional communication, both of enticing and threatening

should be excluded because a clerk is an official person and s/he is obliged to communicate cooperatively, impersonally, peacefully, etc (i.e. to stay in a fixed point of the communicative space). S/he can only persuade a client.

3.1 Sales Clerks' Tactics

All the dialogues are recorded in the beginning phase of negotiations, therefore, *B* takes a final decision only in few cases. A typical dialogue starts with *A*'s introduction and a question whether *B* does know the education company. Then a short overview of the company is given (e.g. *we are an international company, we are acting six years in Estonia, we are dealing with sale, service, management, marketing*). All the statements can be considered as arguments for taking a training course. Then a proposal is made by *A* to take some courses. *A* points the activities of *B*'s organisation which demonstrates that s/he has pre-knowledge about the institution (e.g. *your firm is dealing with retail and whole sale, therefore you could be interested in our courses*). If *B* does not make a decision then *A* asks *B* to tell more about *B*'s institution in order to get more arguments for usability of the courses for *B*, and offers them again. The dialogues end with an agreement to keep the contact (*A* promises to send information materials to *B*, to call *B* later). *B* does not decide to accept nor reject a course but postpones the decision.

If *A* and *B* have been in contact before, then *A* always starts the conversation with pointing to a previous contact. *B* has had the time to evaluate the information about the courses in order to make a decision. *B* agrees to take a course only in one conversation, s/he agrees with reservations in two dialogues, and does not agree in one dialogue. In the remaining dialogues, *A* and *B* come to the agreement to keep the contact like in the case of the first communication. Therefore, *B* typically postpones the decision.

3.2 Clients' Tactics

A's final goal is that *B* decides to do *D* (to take a course). In the case of **collaboration**, *B* actively looks for arguments for doing *D*.

In a typical dialogue, *A* introduces himself, gives an overview of his company (*it offers courses of management, marketing, sale, customer service, secretary training*), and asks whether *B* has made training plans for his employees (i.e. an indirect proposal to take a course). *B* argues that his staff is small, and he has got many offers from other training companies (i.e. a refusal with two

arguments). Then *A* tries to awake *B* to a certain course by asking about customers of *B*'s firm. After that an offer is made to send a catalogue. Now, *B* takes the initiative starting to check the presence of resources and usability of performing *D*. At the end of conversation, *A* and *B* agree that *A* sends a catalogue and calls *B* again a week later.

Pure **antagonism** is expressed in one dialogue. *B* has studied the catalogue, and made the negative decision (*but yes, I have studied it and unfortunately, I'll say that you are not able to teach what I want*). *A* is looking for new arguments and asks questions about activities of *B*'s company trying to show that the courses are useful for *B*. Anyway, *B* does not give up, and the dialogue ends with a resolute refusal. Here, both participants try to take initiative. *A* implements the tactic of persuasion but *B* does not capitulate.

In most cases, *B* having studied a catalogue, starts a conversation with antagonism but goes over to collaboration, i.e. uses a **mixed tactic**.

4 CONCLUSIONS

When communicating in a natural language, where *A* tries to influence *B* in order to bring him/her to a decision to perform an action *D*, *A* uses several arguments in order to increase the weights of the positive aspects and to decrease the weights of the negative aspects of the action under consideration. When *B* has started the reasoning process, s/he considers various positive and negative aspects of *D*. If the positive aspects weigh more, *B* will make the decision to do *D*, otherwise s/he will make the decision not to do *D*. If *B* indicates a certain aspect which does not allow him/her to do *D* then *A* simply can choose an argument for attacking this aspect until there are arguments at his/her disposal. When reasoning, *B* can make his/her negative decision on different steps. For example, if *B* says that resources are missing and *A* indicates that resources can be obtained then *B* has to start his/her reasoning again from the beginning. If *B* does not indicate a certain reason of rejection then *A* can only stress the usefulness of *D* (when persuading).

The corpus analysis shows that our argumentation model is a coarse approximation to real human argumentation. For example, it does not allow deviations from the main line of argumentation by asking questions about possible new arguments. Nevertheless, we believe that the model can be useful for training argumentation in such a way that various arguments are classified and a strong

discipline is set to the order of using different arguments. For example, when persuading, stress the usefulness of *D* until the set of arguments becomes exhausted. At the moment, a simplified version of our model is implemented as an experimental DS which can optionally play the role of *A* or *B* using classified sets of pre-defined Estonian sentences and can be used as a "communication trainer".

Similar approaches are described in some other papers, e.g., presenting arguments in (Elhadad, 1995), car selling agent in (Piwek and van Deemter, 2007). Comparison of the approaches remains for the further work.

We are continuing our work in the following directions: (1) analysis of human-human dialogues in the EDiC in order to verify and to refine the model, (2) specifying resistance strategies, (3) developing linguistic knowledge of the DS.

REFERENCES

Allen, J., 1994. *Natural Language Understanding*. Second Edition, The Benjamins/Cummings Publ. Co.

Elhadad, M., 1995. Generating coherent argumentative paragraphs. In: *Journal of Pragmatics*. North Holland, Elsevier, 24, 189–220.

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

Koit, M., Roosmaa, T., Õim, H., 2009. Knowledge Representation for Human-machine Interaction. In: *Proc. of International Conference on Knowledge Engineering and Ontology Development.* Portugal, INSTICC, 396–399.

McTear, M.F., 2004. *Spoken Dialogue Technology. Toward the Conversational User Interface*, Springer.

Õim, H., 1996. Naïve theories and communicative competence. Reasoning in communication. In: *Estonian in the changing world.* Tartu University, 211–231.

Piwek, P., van Deemter, K., 2007. Generating under Global Constraints: the Case of Scripted Dialogue. In: *Journal of Research on Language and Computation*, 5(2), 237–263.