From Study of Human-human Dialogues to Reasoning Model
Conversational Agent in Argumentation Dialogue

Mare Koit[1] and Haldur Õim[2]
[1]*Institute of Computer Science, University of Tartu, 2 J. Liivi St., Tartu, Estonia*
[2]*Institute of Estonian and General Linguistics, University of Tartu, 2 J. Liivi St., Tartu, Estonia*

Keywords: Reasoning, Conversational Agent.

Abstract: We study human-human dialogues where one of the participants tries to influence the reasoning process of the dialogue partner in order to force the partner to make a decision to perform an action. Our further aim is to implement a dialogue system which would interact with a user in natural language. A model of the motivational sphere of a reasoning subject will be presented as a vector which consists of evaluations of different aspects of the action under consideration. Three reasoning procedures will be introduced, each of which is triggered by a so-called input factor. We examine the communicative strategies and communicative tactics that dialogue participants use to achieve their communicative goals. The models are implemented as a computer program.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

Several general approaches to the theory of pragmatics of natural communication have been used when developing models of dialogue (D´Andrade, 1987); (Davies and Stone, 1995); (Lester et al., 2004); (Jokinen, 2009); (Ginzburg and Fernández, 2010). One of the main problems for pragmatics is that of modelling the mechanisms people use to reach their communicative goals. This is done by manipulating the appropriate reasoning processes of other participants. Therefore, a pragmatic model of dialogue should include a commonsense, naïve model of reasoning and present the means to influence the reasoning processes that are regularly used by people in communication. Most of our naïve theories are and will remain implicit. It is the task of a real science, such as psychology or linguistics, to explicate this knowledge (Õim, 1996).

We study communications where one of the participants is trying to achieve the partner's decision to perform an action and have worked out our versions of the concepts of communicative strategies and tactics. Below we focus on these concepts and the concept of reasoning model, describing the current state of their implementation in our dialogue system. In general, our model of reasoning follows the ideas realized in the well-known BDI (belief-desire-intention) model (Rao and Georgeff, 1991).

The paper has the following structure. In section 2 we introduce our reasoning model. Section 3 examines how the reasoning can be influenced in interaction. We introduce a communicative strategy and communicative tactics as algorithms used by participants in order to achieve their communicative goals. Section 4 discusses how the dialogue model is and can be implemented. In section 5 we draw conclusions.

2 REASONING MODEL

2.1 Human Reasoning

Let us start with considering a dialogue example (Ex. 1) taken from the Estonian dialogue corpus. A client (participant *A*) is calling a travel agent (*B*) and is looking for a ski trip to Austria (action *D*). In the following examples the transcription of conversation analysis is used (Hutchby and Wooffitt, 1998).

(1)
A : .hh e sooviksin teada: natuke ´Austria suusareisi kohta. | REQUEST | **I'd like to know a bit about the ski trip to Austria**
B : jaa? | CONTINUER |

Koit M. and Õim H..
From Study of Human-human Dialogues to Reasoning Model - Conversational Agent in Argumentation Dialogue.
DOI: 10.5220/0004195902100216
In *Proceedings of the 5th International Conference on Agents and Artificial Intelligence* (ICAART-2013), pages 210-216
ISBN: 978-989-8565-39-6
Copyright © 2013 SCITEPRESS (Science and Technology Publications, Lda.)

yes
A : .hh et=mm (.) kui ´kaua se ültse ´kestab. | WH-QUESTION |
how long will it last
/---/
A : .hh ja (.) mis: ´hinnaklass see tuleb kui näiteks apartemendi vari´ant võtta.= | WH-QUESTION |
and what price category would it be for the option with accommodation in an apartment
/---/
A : .hh ja siis selle ´suusatamisega seal on: (.) ´mägedes suusatamine. | QUESTION OFFERING ANSWER |
and about skiing will the skiing be in mountains
B : jah. | YES |
yes
A : ja: kas need´suusad ja ´varustus on nagu ´lisatasu eest või see on: [(.) ´hinna sees.] | ALTERNATIVE QUESTION|
and the skies and the equipment are there for extra cost or is it included in the price

A is collecting information that he needs to make a decision about the action (here: ski trip). We do not know exactly how the reasoning process proceeds in *A*'s head. Only *A*'s utterances provide the indirect signals for the travel agent to help her draw conclusions about how several aspects of the action are weighed by *A*.

2.2 Model of Reasoning Subject

We assume that the reasoning process of a certain type is triggered by a so-called input factor. We distinguish between three types of input factors: (1) the reasoning subject may wish (would like) to perform action *D* (*wish*-factor), (2) the subject may depart from the assumption that doing *D* is *useful* for him (*needed*-factor), or (3) that doing *D* is *obligatory* (*must*-factor). In short, these factors constitute what can be called the (macro) model of human motivational system underlying his/her reasoning whether to take an action or not. When the reasoning process has started, the subject considers (weighs) the positive and negative aspects of *D*: how pleasant or unpleasant, useful or harmful it is, what punishment will follow if he does not do *D*, etc. If the positive aspects weigh more, the subject will make the decision to perform *D*, otherwise the decision will be not to do D.

When constructing our *model of reasoning* we assume that the reasoning subject is somehow able to evaluate the positive and negative aspects of the object of reasoning (in our case, action *D*). Here we assume that the relevant aspects of *D* can be characterized by scales that take certain numerical values, the so-called *weights*. Although, in reality people do not operate with numbers, the existence of certain scales also in human everyday reasoning is apparent. What would be a more adequate form for these scales, is a problem of future research.

We use the following notation: w(pleasant), w(unpleasant), w(useful), w(harmful) – weight of the pleasant, unpleasant, useful and harmful aspects of *D*, respectively; w(obligatory) – the value shows whether *D* is obligatory (=1) or not (=0), w(prohibited) – whether *D* is prohibited (=1) or not (=0), w(punishment-do) – weight of punishment for performing a prohibited action, w(punishment-do-not) – weight of punishment for not performing an obligatory action, w(resources) – the value indicates whether the subject has resources necessary for performing *D* (=1) or not (=0).

According to our present model, the motivational sphere of a reasoning subject can be represented by the following vector of weights:

w = (w(resources), w(pleasant), w(unpleasant), w(useful), w(harmful), w(obligatory), w(prohibited), w(punishment-do), w(punishment-do-not)).

2.3 The Three Reasoning Procedures

The model of the motivational sphere is used by a reasoning subject when weighing different aspects of the action under consideration. Reasoning is triggered by an input factor. As an example, we present the reasoning procedure WISH that originates in the wish of a subject to do *D* (Fig. 1). The prerequisite for triggering this procedure is w(pleasant) > w(unpleasant) based on the following assumption: if a person wishes to do something, then he assumes that the pleasant aspects of *D* (including its consequences) overweigh its unpleasant aspects.

Different aspects of *D* are subsequently evaluated and the final decision depends on the result of the comparison of different values as outcomes of the evaluation process. For instance, if the subject does not have enough resources for *D*, then, independently of his wish, he will decide not to do *D* (step 1 in the procedure). If the subject has the necessary resources and the weight of the pleasant aspects exceeds the sum of the unpleasant and the harmful ones and, in addition, *D* is not prohibited, then the subject will decide to perform *D* (steps 1, 2 and 3), etc. The same kinds of procedures, NEEDED

and MUST, are constructed for the reasoning processes triggered by the *needed* and *must* factors (Koit and Õim, 2004).

```
Prerequisite: w(pleasant) > w(unpleasant)
 1) Are there enough resources for doing
D? If not then do not do D.
 2) Is w(pleasant) > w(unpleasant) +
w(harmful)? If not then go to step 6.
 3) Is D prohibited? If not then do D.
 4) Is w(pleasant) > w(unpleasant) +
w(harmful) + w(punishment-do)? If yes
then do D.
 5) Is w(pleasant) + w(useful) >
w(unpleasant) + w(harmful) +
w(punishment-do)? If yes then to do D
else do not do D.
 6) Is w(pleasant) + w(useful) >
w(unpleasant) + w(harmful)? If yes then
go to step 9.
 7) Is D obligatory? If not then do not
do D.
 8) Is w(pleasant) + w(useful) +
w(punishment-do-not) > w(unpleasant) +
w(harmful)? If yes then do D else do not
do D.
 9) Is D prohibited? If not then do D.
10) Is w(pleasant) + w(useful) >
w(unpleasant) + w(harmful) +
w(punishment-do)? If yes then do D else
do not do D.
```

Figure 1: Reasoning procedure WISH.

Returning to Ex. 1, we can see how *A*'s utterances keep a "trace" of his reasoning process: *A* is weighing the duration of the trip, its price, the skiing location, etc. Every utterance refers to a specific aspect of *D*, the action under consideration.

2.4 Discussion

We do not claim that the above three reasoning procedures – WISH, NEEDED and MUST – exhaustively cover all the varieties of reasoning on which human action is based that can be encountered in the "real life". There are numerous kinds of situations not accounted for by our model so far, although we have dealt with them theoretically. For instance, there can be situations where a person has several simultaneously activated and competing input factors, e.g. two competing wishes for both of which he holds w(pleasant) > w(unpleasant). Such situations can be partly accounted for by certain general principles (e.g. "From two possible pleasant situations people prefer the more pleasant one"). There are no major problems with incorporating such principles into our model.

On the other hand, these motivational factors are not independent of each other. Thus, a useful outcome of an action is in some sense also pleasant for the subject, punishment is unpleasant (and can be harmful) for the punished person, etc., but we will not go into these details in our present model.

3 REASONING IN INTERACTION

As said above, our general goal is to model reasoning in communication. Our further aim is to build a conversational agent – a computer program that can participate in interaction with a human user.

3.1 Human-Human Communication

Both participants, *A* and *B*, have their own communicative goals. In Ex.1, the client's (*A*) communicative goal is to make a decision about a ski trip. *B* (a travel agent) is definitely interested in *A*'s positive decision. Let us consider how the communication between *A* and *B* continues (Ex. 2).

(2)

```
B:  [neil] on ´väga=ea (.) olemas
näiteks e jaanuarikuuks ´väga=ead
pakkumised Rootsi suusakuurortitesse.
| SPECIFICATION |
there are very good offers to Swedish
holiday resorts in January
A:  mhmh    | CONTINUER |
hem
B:  kus on noh ütleme ´hinnad on innad
on niuksed ´tõeliselt (.) ütleme teevad
Soomele ´ka (.) $ silmad ´ette. $
| ACCOUNT |
the prices are really let's say much
better than in Finland
A:  ah=nii.=    | CHANGE OF STATE |
I see
B:  =ja ´majutus on väga ´korralik    |
SPECIFICATION |
and accommodation is very descent
/---/
et sis kui te saate nagu selle ´tunde
kätte et siis juba ´siis juba minna
´Austriasse.    | OPINION |
and after you've got the feel of it
then you can go to Austria
A:  mhmh    | LIMITED ACCEPT |
I see
```

B has understood that a high price is a problem for *A*. Now she takes the initiative and offers another, cheaper trip to Sweden, trying thus to influence the client to make a positive decision about another action. What she can do is to stress the pleasant

aspects of *D* (i.e. to *entice*), or the usefulness of *D* (i.e. to *persuade*), or to *threaten B* with a punishment, if he does not do *D* (threatening is excluded in the current situation). We will say that *B* applies a communicative strategy which can be realized by using different *communicative tactics* (enticement, persuasion and threatening, accordingly). The dialogue will continue until *B* reaches her goal (i.e. the decision of *A* to perform the action) or gives up.

3.2 Communicative Goal and Communicative Strategy

In our approach the *communicative strategy* is formalized as an algorithm that a participant applies to achieve his/her communicative goal.

Two kinds of strategies are important in our case: attack and defence. In the first case, a participant tries to press his/her communicative goal onto the partner. In the second case, s/he averts taking over the partner's goal. In the situation under consideration, the communicative strategy used by *B* (attack) can be presented as the algorithm in Fig. 2.

```
1) Choose the communicative tactics.
2) Implement the tactics to generate a
turn (inform the partner of the
communicative goal – to do D).
3) Did the partner agree to do D? If yes
then finish (the communicative goal has
been reached).
4) Give up? If yes then finish (the
communicative goal has not been reached).
5) Change the communicative tactics? If
yes then choose the new tactics.
6) Implement the tactics to generate a
turn. Go to step 3.
```

Figure 2: Communicative strategy.

3.3 Communicative Tactics

The conversational agent we are modelling performs the role of *B*. In our model there are three different communicative tactics that *B* can use as part of its communicative strategy: *enticement, persuasion* and *threatening*. Each of the communicative tactics constitutes a procedure for compiling a turn in the ongoing dialogue: the tactic of enticement consists in increasing *A*'s *wish* to do *D*; persuasion consists in increasing *A*'s belief of the *usefulness* of *D* for him, and threatening consists in increasing *A*'s understanding that he *must* do *D*.

Communicative tactics are directly related to the reasoning processes of partner *A*. For instance, if *B* is applying the tactic of enticement it should be able to imagine the reasoning process of *A* that is triggered by the input factor *wish*. When *A* at a certain stage refuses to do *D*, then *B* should be able to guess at which point the reasoning of *A* went into the "negative" branch ("do not do *D*"), in order to adequately construct its reactive turn. Similarly, the tactic of persuasion is related to the reasoning process triggered by the *needed*-factor, and the threatening is related to the reasoning process triggered by the *must*-factor. Therefore, in order to model various communicative tactics, the reasoning model is used.

3.4 Model of Enticement

When implementing a communicative strategy the agent *B* uses a model of the motivational sphere of partner *A* – a vector $\boldsymbol{w}^{BA}$ – which includes its idea about weights of the aspects of action *D*. The more *B* knows about *A* the more similar the vector $\boldsymbol{w}^{BA}$ is with the actual vector $\boldsymbol{w}^{A}$ of the motivational sphere of partner *A*. Here we assume that *B* has several sets of statements for increasing/decreasing the weights of the different aspects of *D* for partner *A*. All the statements have their (numerical) weights as well (Koit, 2011).

As an illustration, we shortly describe the tactic of *enticement* that is based on the reasoning procedure WISH (Fig. 1). The general idea underlying this tactic is that *B* presents *A* with statements for pleasantness of *D* trying to keep the weight of pleasantness for *A* high enough and the values of negative aspects brought out by *A* low enough so that weighing positive and negative aspects would lead *A* to the decision to perform *D*. We suppose that *A* has a set of statements for indicating the aspect which weight caused his rejection. Here we assume that *B* when enticing uses each statement only once.

3.5 Discussion

The communicative tactics used by the participants are not limited to the three that were mentioned above. Firstly, while influencing the reasoning process of the partner a participant may repeatedly use the same argument in order to change a specific weight in the partner model. Secondly, communicative tactics need to be specified for *A* too. Two different scenarios are possible: (1) *A* and *B* have opposite communicative goals (*A* does not plan to do *D* but *B*'s goal is to get him do it); (2) both *A* and *B* have the same goal and cooperate with each other when looking for arguments that support

achieving it. However, the issue of specifying the communicative tactics of *A* will be left for the future work.

4 CONVERSATIONAL AGENT

We have implemented the described models as a conversational agent – a DS which interacts with a user in Estonian and tries to achieve the user's decision to perform an action by influencing his reasoning about the action.

4.1 Architecture

The DS we have been developing consists of the following functional blocks: natural language understanding and generation modules, a planner, a dialogue manager, and a problem solver. The problem solver enables the system to "tune in" to a specific problem domain. The other blocks form a basic interaction system (Jurafsky and Martin, 2008).

The modules use a knowledge base where different knowledge is kept: linguistic knowledge, also knowledge about the world (in our case – frames of actions), communication (communicative strategy, communicative tactics, dialogue acts), and users (partners' models, reasoning procedures).

The natural language understanding module analyzes the utterances of an input turn, and outputs their representations as the corresponding recognized dialogue acts (question, answer, request, etc.). The task of the planner is to construct a turn of the DS, either as a response to the user's turn or as a turn initiated by the DS itself. In this process, the planner contacts the problem solver (if a domain problem has to be solved) and the dialogue manager to determine the communicative structure of the turn (dialogue acts). The natural language generation module will compose the semantic structure underlying the planned output and transform it into a linguistic expression.

It is the task of the dialogue manager to determine how to proceed if the communicative goal of the DS (to achieve the user's decision to perform an action) has not been attained by the preceding turns. The reasoning model plays a crucial role in this: the dialogue manager has to decide where the reasoning of the partner went into a "negative" branch and try to find new material which is expected to lead to a positive outcome.

When comparing our reasoning model with BDI model, then beliefs are represented by knowledge of the conversational agent with reliability less than 1; desires are generated by the vector of weights $\boldsymbol{w}$; and intentions correspond to goals in goal base. In addition to desires, from the weights vector we also can derive some parameters of the motivational sphere that are not explicitly conveyed by the basic BDI model: needs, obligations and prohibitions.

4.2 Implementation

Presently, we have implemented a program which can play the role of *B* in a simple communication situation where the goal of *B* is that *A* (user) decides to perform action *D*. At the moment, the computer operates only with the semantic representations of the linguistic input/output. In the current implementation, ready-made Estonian sentences (texts) are used both by the computer and the user. The sentences are classified according to their function, e.g. for increasing the weight of pleasantness, for decreasing the weight of harmfulness, for expressing that pleasantness is too low, etc. Every sentence has a numerical weight 1. The work on a linguistic processor is in progress.

At the beginning of the dialogue the computer expresses the communicative goal (this is its first turn r^{B1}). If the user refuses to do *D* (after implementing normal human reasoning which we are trying to model here), based on the response (r^{A1}) the computer determines the aspect of *D* the weight of which does not match the reality and changes this weight so that the new model will give a negative result as before but it is an extreme case: if we increased this weight by one unit (in case of positive aspects of *D*) or decreased it (in case of negative ones) we should get a positive decision. The computer chooses its response r^{B2} from the set of sentences for increasing/decreasing this weight and at the same time it increases/decreases this weight in the partner model by the value of the chosen sentence. A reasoning procedure based on the new model will yield a positive decision. Now the user must choose his response and the process can continue in a similar way.

4.3 Discussion

A dialogue is generated jointly by the computer and a user. The computer uses its communicative strategy and tactics. Let us suppose that after the computer's proposal to perform an action *D*, the user will create a model of himself, i.e. he will attribute values to all aspects of *D* and will reason on the basis of this model. Of course, creating this model is

implicit, e.g. the user assesses that doing *D* would be more unpleasant than pleasant. By implementing its communicative strategy and tactics, the computer has to try to influence the partner model in the way that would cause the partner to make a positive decision based on the changed model. The problem is that the computer does not "know" the real weights attributed to different aspects of *D* by the user. It can only guess these values based on the user's negative responses.

At the beginning of a dialogue the computer randomly generates a user model. At the moment we have set only one restriction: we require that the initial model should satisfy the assumption(s) that underlie the corresponding reasoning procedure. Thus, for enticing w(pleasant) > w(unpleasant), for persuading w(useful) > w(harmful) and for threatening w(obligatory) = 1. When an initial model is generated the computer uses it as a partner model and informs the user about its communicative goal. It chooses a sentence (r^{BI}) from a special file of computer sentences. A user can choose his sentences r^{Ai} (i=1,2,...) from a special file of user sentences, i.e. he can "play a role" but cannot use unrestricted texts. If a user has chosen a sentence of refusal, the computer decides that the user model is inexact and needs amending. The corresponding class of user sentences of refusal will be recognized and the aspect of *D* determined the weight of which in the user model was either too small or too great, which brought about the false decision by the computer. Based on a valid reasoning procedure (and tactics) a new value will be computed for this weight, which is congruent with the negative decision (as explicated by the user expression).

Our research has a practical aim: to implement a communication trainer, a computer program that would allow the user to exercise his abilities to reach certain communicative goals: (a) getting the partner to decide to perform an action, or (b) on the contrary, opposing the partner (Koit, 2012).

5 FUTURE WORK

We have examined here only a very restricted type of dialogues where the user must play a particular rigid role. In the future we plan to model such situations where the computer will take the participant *A*'s role. In order to do that, *A*'s strategies and tactics need to be modeled.

One of our priorities will be to investigate the possibilities of adding contextual aspects to the reasoning model. One option is to include the personal background of the participants, e.g. by elaborating the notion of *communicative space* (Brown and Levinson, 1999). In our case, the communicative space is determined by a number of coordinates, such as social distance between the partners (far between adversaries, close between friends), intensity of communication (peaceful, vehement), etc. Without taking this information into account, formal reasoning about some action can easily run into problems such as inconsistency, due to considering the knowledge in a wrong context, inefficiency, when irrelevant knowledge is being considered, or incompleteness, when the relevant inferences are not made.

ACKNOWLEDGEMENTS

This work is supported by the European Regional Development Fund through the Estonian Centre of Excellence in Computer Science (EXCS), the Estonian Research Council (grant ETF9124), and the Estonian Ministry of Education and Research (grant SF0180078s08).

REFERENCES

Brown, P. and Levinson, S. C., 1999. Politeness: Some universals in language usage. In *The discourse reader.* London: Routledge, 321–335.

D´Andrade, R., 1987. A Folk Model of the Mind. In *Cultural Models of Language and Though.* London: Cambridge University Press, 112–148.

Davies, M. and Stone, T., 1995. *Folk psychology: the theory of mind debate.* Oxford, Cambridge, Massachusetts: Blackwell.

Ginzburg, J. and Fernández, R., 2010. Computational Models of Dialogue. In *The Handbook of Computational Linguistics and Natural Language Processing*, 429–481. Wiley Blackwell Publishers.

Hutchby, I. and Wooffitt, R., 1998. *Conversation Analysis. Principles, Practices and Applications*. Cambridge, UK: Polity Press.

Jokinen, K., 2009. *Constructive Dialogue Modelling: Speech Interaction and Rational Agents.* John Wiley & Sons Ltd.

Jurafsky, D. and Martin, J. H., 2008. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.* Prentice Hall.

Koit, M., 2012. Developing Software for Training Argumentation Skills. In *Proc. of CMNA 2012, Workshop of ECAI 2012,* 11–15. Montpellier, France. http://www.cmna.info/

Koit, M., 2011. Conversational Agent in Argumentation:

Updating of Information States. In *Proc. of KEOD 2011*, ,375–378. Paris: SciTEC Publications Ltd.

Koit, M. and Õim, H., 2004. Argumentation in the Agreement Negotiation Process: A Model that Involves Natural Reasoning. In *Proc. of the Workshop CMNA. 16th European Conference on Artificial Intelligence*, 53–56. Valencia, Spain. http://www.cmna.info/

Lester, J. Branting, K. and Mott, B. 2004. Conversational. Agents. In *The Practical Handbook of Internet Computing.* Chapman & Hall.

Õim, H. 1996. Naïve Theories and Communicative Competence: Reasoning in Communication. In *Estonian in the Changing World*, 211–231. University of Tartu.

Rao, A. S. and Georgeff, M. P. 1991. Modeling Rational Agents within a BDI-Architecture. In *Proc. of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, 473–484. Morgan Kaufmann.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS