

# Influencing the Beliefs of a Dialogue Partner

Mare Koit<sup>(✉)</sup>

University of Tartu, J. Liivi 2, 50409 Tartu, Estonia

[mare.koit@ut.ee](mailto:mare.koit@ut.ee)

<http://www.cl.ut.ee>

**Abstract.** A model of argumentation dialogue which includes reasoning is introduced in the paper. The communicative goal of the initiator is to convince the partner to do an action. The choice of an argument depends, on the one hand, on the needed resources and the beliefs about the positive and negative aspects of doing the action, and on the other hand, on the result of reasoning based on these beliefs. The initiator of dialogue is using a partner model – the hypothetical beliefs about the partner who at the same time operates with the actual beliefs. Both the participants' models are changing during a dialogue as influenced by the partners' arguments. Two implementations have been created. In one implementation, the computer initiates a dialogue and attempts to influence the user to make a decision about doing an action. In the other implementation, the roles of the computer and the user are reversed. Interaction is text-based, participants are using ready-made sentences in natural language which are classified semantically. The paper studies how the participants are updating their beliefs in dialogue. The study is based on the interactions with our dialogue systems.

**Keywords:** Beliefs · Updating · Reasoning model · Argumentation dialogue

## 1 Introduction

A dialogue system (DS), or conversational agent, is a computer system intended to interact with a human using text, speech, graphics, gestures and other modes for communication. A dialogue manager is a component of a DS which controls the conversation. The dialogue manager reads the input modalities, updates the current state of the dialogue, decides what to do next, and generates output [1].

Four kinds of dialogue management architectures are most common: plan-based, finite-state, frame-based, and information-state [2, chap. 24]. One of the earliest models of conversational agent is based on the use of artificial intelligence planning techniques. Using plans to generate and interpret sentences require the models of beliefs, desires, and intentions (BDI) [3, 4]. Plan-based approaches, though complex and difficult to embed in practical dialogue systems, are seen as more amenable to flexible dialogue behavior [5].

The simplest dialogue manager architecture, used in many practical implementations, is a finite-state manager. Frame-based dialogue managers ask the

user questions to fill slots in a frame until there is enough information to perform a data base query, and then return the result to the user. If the user answers more than one question at a time, the system has to fill in these slots and then remember not to ask the user the associated questions for the slots. In this way, the user can also guide the dialogue [6].

More advanced architecture for dialogue management which allows for sophisticated components is the information-state architecture [5, 7]. We use this approach in our implementations.

‘Dialogue state tracking’ refers to accurately estimating the user’s goal as a dialogue progresses. It is sometimes also called ‘belief tracking’ [8].

A conventional DS is a speech-based system that gives information to a user. In this paper, however, we will consider another kind of dialogues in natural language – negotiations that include argumentation. We suppose that there are two participants –  $A$  and  $B$  – and one of them – let it be  $A$  – initiates the dialogue making a proposal to his partner  $B$  to do (or not) an action  $D$ . If  $B$  refuses then  $A$  attempts to influence her in a dialogue, proposing several arguments for/against doing  $D$ .  $A$ ’s arguments are based on the partner model – his image about  $B$ ’s beliefs. At the same time,  $B$  can present counter arguments if her goal is opposite to  $A$ . The counter arguments indicate which beliefs of  $A$  about  $B$  are wrong and how the partner model has to be updated by  $A$ . One possible scenario for such an interaction is that  $A$  is a conversational agent (DS) and  $B$  is a human user. However, we do not exclude the other scenarios: (a)  $A$  is a user and  $B$  is a conversational agent, or (b) both  $A$  and  $B$  are artificial agents, or (c) both of them are humans. These scenarios give us an opportunity to study and to model behaviour of both participants in order to understand how their beliefs are changing during a dialogue as influenced by the partner’s arguments. We have created two different experimental dialogue systems – one is playing the role of  $A$  and the other the role of  $B$  in an argumentation dialogue with a user. At the moment, the interaction is text-based. A possible future practical application, as we see, could be to train the user’s argumentation skills in interaction with a DS.

The remainder of this paper is organised as follows. Section 2 describes our model of argumentation dialogue. Section 3 studies how the partner model (i.e. the agent’s beliefs about the partner) is updated in a conversation and how the actual beliefs of an agent are changing due to the partner’s arguments. The relationship between the partner model and the actual beliefs will be discussed in Sect. 4. Section 5 makes conclusions.

## 2 A Model of Argumentation Dialogue

Let us consider a dialogue in natural language between two participants  $A$  and  $B$  (human or artificial agents). Let  $A$  be the initiator of dialogue, and let his communicative goal be “ $B$  makes  $A$  decision to do an action  $D$ ” or, respectively, “ $B$  makes a decision not to do  $D$ ”.  $B$ ’s communicative goal can be either the same or opposite. In interaction,  $A$  is influencing  $B$  to make the decision which coincides with his communicative goal. The following cases can occur:

1.  $A$ 's goal is " $B$  decides to do  $D$ " but  $B$ 's goal is " $B$  will not do  $D$ "
2.  $A$ 's goal is " $B$  decides not to do  $D$ " but  $B$ 's goal is " $B$  will do  $D$ "
3.  $A$ 's goal is " $B$  decides to do  $D$ ", and  $B$ 's goal is " $B$  will do  $D$ "
4.  $A$ 's goal is " $B$  decides not to do  $D$ ", and  $B$ 's goal is " $B$  will not do  $D$ ".

$A$ 's and  $B$ 's communicative goals are opposite in the cases (1) and (2). When interacting, the initiator  $A$  presents arguments in order to influence  $B$  to adopt  $A$ 's goal and to abandon her own initial goal. At the same time,  $B$  can present counter arguments which should bring  $A$  to adopt  $B$ 's goal and to abandon his own initial goal.

$A$ 's and  $B$ 's communicative goals coincide in the cases (3) and (4). When interacting, they cooperatively look for arguments in support of doing (respectively, not doing)  $D$  and find out how to overcome possible obstacles before doing  $D$  or, respectively, to prevent possible undesirable consequences of not doing  $D$ .

Let us, for example, consider the case (1). The initiator  $A$  has a partner model – an image about  $B$ 's beliefs. The partner model gives him an opportunity to suppose that  $B$  will agree to accept his communicative goal (to do the action  $D$ ). When constructing his first turn,  $A$  chooses the dialogue acts (e.g. request, proposal, question, etc. depending on his image about  $B$ ) and determines their verbal form (utterances). The partner  $B$  analyses  $A$ 's turn and in order to make a decision – to do  $D$  or not – she triggers a reasoning procedure in her mind. In the reasoning process,  $B$  is weighing her resources for doing  $D$ , positive and negative aspects of doing  $D$  and finally, she makes a decision. Then  $B$  in her turn chooses the dialogue acts (agreement, refusal) and their verbal form in order to inform  $A$  about her decision. If  $B$  agrees to do  $D$  then the dialogue finishes ( $A$  has reached his communicative goal). If  $B$  refuses then  $A$  has to change his partner model (it did not correspond to the reality) and find out new arguments in order to convince  $B$  to make a positive decision, cf. [9].

$B$  can add arguments to her refusal. These (counter) arguments give information about the reasoning process that brought  $B$  to the (negative) decision.  $A$  uses the arguments given by  $B$  for updating his current partner model.

## 2.1 Reasoning Model

Our reasoning model is introduced in [9]. In general, it follows the ideas realised in the well-known BDI model.

The reasoning model consists of two parts: (1) a model of human motivational sphere; (2) reasoning procedures. In the motivational sphere three basic factors are differentiated that regulate reasoning of a subject concerning of doing an action  $D$ . First, a subject may *wish* to do  $D$  if the pleasant aspects of  $D$  for him/her overweight the unpleasant ones; secondly, a subject may find it reasonable to do  $D$  if  $D$  is *needed* to reach some higher goal, and the useful aspects of  $D$  overweight the harmful ones; and thirdly, a 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.

If the subject is reasoning about not doing  $D$  then the basic factors which trigger the reasoning are analogous: first, the subject *does not wish* to do  $D$  if the unpleasant aspects of  $D$  overweight the pleasant ones; secondly, doing  $D$  is *not needed* for him/her if the harmful aspects of  $D$  overweight the useful aspects; and thirdly, doing  $D$  is *not allowed (prohibited)* for him/her if it will cause some punishment.

We represent the model of motivational sphere of a reasoning subject by the following vector of 'weights' of beliefs (with numerical values of its components):

$\mathbf{wD} = (wD(\text{resources}), wD(\text{pleasant}), wD(\text{unpleasant}), wD(\text{useful}), wD(\text{harmful}), wD(\text{obligatory}), wD(\text{prohibited}), wD(\text{punishment-do}), wD(\text{punishment-not}))$ .

In the description,  $wD(\text{pleasant})$ , etc. mean the weight of pleasant, etc. aspects of  $D$ ;  $wD(\text{punishment-do})$  – the weight of punishment for doing  $D$  if it is prohibited, and  $wD(\text{punishment-not})$  – the weight of punishment for not doing  $D$  if it is obligatory. Further,  $wD(\text{resources}) = 1$  if subject has all the resources necessary to do  $D$  (otherwise 0);  $wD(\text{obligatory}) = 1$  if  $D$  is obligatory for the reasoning subject (otherwise 0);  $wD(\text{prohibited}) = 1$  if  $D$  is prohibited (otherwise 0). The values of other weights can be non-negative natural numbers.

The second part of the reasoning model consists of reasoning procedures that supposedly regulate human action-oriented reasoning. Every reasoning procedure represents steps that the subject goes through in his/her reasoning process; these consist in comparing the summarised weights of different aspects of  $D$ ; and the result is the decision: to do  $D$  or not.

We use two vectors of motivational sphere. The vector  $\mathbf{wDAB}$  represents  $A$ 's beliefs concerning  $B$ 's evaluations and it is used as a partner model. The vector  $\mathbf{wDB}$  represents  $B$ 's actual evaluations of  $D$ 's aspects (which exact values  $A$  does not know) and it is used as the model of  $B$  herself. In the paper, we will consider the needed changes that will be made and tracked by the participants due to arguments presented in a dialogue. In the following, we suppose that the action  $D$  is fixed and we do not indicate it in the vectors.

A reasoning procedure depends on the determinant which triggers it. As an example, let us present the reasoning procedure WISH as a step-form algorithm triggered by the *wish* of the reasoning subject to do  $D$ , that is,  $D$  is more pleasant than unpleasant for the subject, cf. [9].

Presumption:  $w(\text{pleasant}) \geq w(\text{unpleasant})$ .

1. Is  $w(\text{resources}) = 1$ ? If not then go to 11.
2. Is  $w(\text{pleasant}) > w(\text{unpleasant}) + w(\text{harmful})$ ? If not then go to 6.
3. Is  $w(\text{prohibited}) = 1$ ? If not then go to 10.
4. Is  $w(\text{pleasant}) > w(\text{unpleasant}) + w(\text{harmful}) + w(\text{punishment-do})$ ? If yes then go to 10.
5. Is  $w(\text{pleasant}) + w(\text{useful}) > w(\text{unpleasant}) + w(\text{harmful}) + w(\text{punishment-do})$ ? If yes then go to 10 else go to 11.
6. Is  $w(\text{pleasant}) + w(\text{useful}) \leq w(\text{unpleasant}) + w(\text{harmful})$ ? If not then go to 9.
7. Is  $w(\text{obligatory}) = 1$ ? If not then go to 11.
8. Is  $w(\text{pleasant}) + w(\text{useful}) + w(\text{punishment-not}) > w(\text{unpleasant}) + w(\text{harmful})$ ? If yes then go to 10 else go to 11.

9. Is  $w(\text{prohibited}) = 1$ ? If yes then go to 5 else go to 10.
10. Decide: do  $D$ . End.
11. Decide: do not do  $D$ .

The idea is quite simple: a reasoning subject is step-by-step weighing positive and negative aspects of doing  $D$ . If the positive aspects in sum weigh more then the decision will be “do  $D$ ” else “do not do  $D$ ”. The reasoning can also be considered as argumentation for/against doing  $D$ .

If  $D$  is more unpleasant than pleasant but more useful than harmful then the reasoning procedure NEEDED can be triggered by the reasoning subject. If  $D$  is obligatory and not doing  $D$  involves a punishment then the subject can use the reasoning procedure MUST, cf. [9].

A *communicative strategy* is an algorithm used by a participant for achieving his/her goal in the interaction. In order to achieve the goal in argumentation dialogue, a participant can present different arguments for/against  $D$  in a systematic way. For example, if the initiator  $A$  has the communicative goal “ $B$  will do  $D$ ” then he can over and over again stress the pleasant aspects of  $D$  (i.e. *entice* the partner  $B$  to do  $D$ ), or stress usefulness of  $D$  for  $B$  (i.e. *persuade*  $B$ ), or stress punishment for not doing  $D$  if it is obligatory (*threaten*  $B$ ), etc. We call *communicative tactics* these concrete ways of applying a communicative strategy [9]. The participant  $A$ , trying to direct  $B$ ’s reasoning to the desirable decision, proposes arguments for doing  $D$  (respectively, not doing  $D$ ) while  $B$ , when opposing, proposes counter arguments. While enticing (respectively, persuading or threatening) the partner for doing  $D$ ,  $A$  attempts to trigger the reasoning procedure WISH (respectively, NEEDED or MUST) in  $B$ ’s mind.

## 2.2 Argumentation-Based Dialogue

Let us make some more assumptions. First, let us consider the general scenario (b, Sect. 1) supposing that both  $A$  and  $B$  are conversational agents interacting in a natural language. (If one of them actually is a human user then some of the introduced models and formalisms are not needed.) Both  $A$  and  $B$  have access to a common set of reasoning procedures. Both  $A$  and  $B$  can use fixed sets of dialogue acts and the corresponding utterances in a natural language which are pre-classified semantically, e.g. the set  $P_{\text{increasingresources}}$  for indicating that there exist resources for doing a certain action  $D$  (e.g. *The company will cover all your expenses*),  $P_{\text{increasingpleasantness}}$  for stressing the pleasantness of  $D$  (e.g. *You can meet interesting people*),  $P_{\text{missingresources}}$  for indicating that some resources for doing  $D$  are missing (e.g. *I don’t have proper dresses*), etc. Therefore, no linguistic analysis or generation will be made during a dialogue (in our implementation). However, these restrictions will involve that the generated dialogues can be not quite coherent.

$A$ , starting an interaction, generates a partner model  $w_{AB}$  (using his knowledge) and determines the communicative tactics  $T$  which he will use (e.g. *entice-ment*), i.e. he accordingly fixes a reasoning procedure  $R$  which he will try to trigger in  $B$ ’s mind (e.g. WISH).  $B$  has her own model  $w_B$  (which exact values

A does not know). She in her turn determines a reasoning procedure RB which she will use in order to make a decision about doing  $D$  (which can be different from R fixed by A) and her communicative tactics TB.

In the following dialogue example (generated with our DS), A is the manager of a company and B is working for the company but is at the same time studying at a university. A presents arguments for doing  $D$  by B (to travel to N. in order to conclude a contract). He succeeds to avert B's counter arguments and convince B to accept his goal. In the example, we also annotate the arguments in **RC** (reason-claim) formalism [10] using additional notations for statements (given in parentheses).

- (1) A: *The company offers you a trip to N. in order to conclude a contract.* (trip)  
 $Y$ ou can meet interesting people. (people) **R**(people): **C**(trip)
- (2) B: *I don't have proper dresses.* (dresses)  
**R**(**R**(dresses) : **C**( $\neg$  trip)) : **C**( $\neg$ **R**(people) : **C**(trip)) [strong rebuttal]
- (3) A: *The company will pay your executive expenses.* (expenses) *The nature is very nice in N.* (nature)  
**R**(**R**(expenses) : **C**( $\neg$  dresses)) : **C**( $\neg$ **R**(dresses) : **C**( $\neg$  trip)) [strong premise attack]  
 & **R**(nature) : **C**(trip)
- (4) B: *I can have some problems at my university.* (university)  
**R**(**R**(university) :  $\neg$ **C**(trip)) : **C**( $\neg$ **R**(nature) : **C**(trip)) [weak rebuttal]
- (5) A: *It's all right – your examinations period will be extended.* (extension) *You can sunbathe in N. early in spring already.* (sunbathe)  
**R**(**R**(extension) : **C**( $\neg$  university)) : **C**( $\neg$ **R**(university) : **C**( $\neg$  trip)) [strong premise attack]  
 & **R**(sunbathe) : **C**(trip)
- (6) B: *OK, I'll do it.*
- (7) A: *I am glad.*

Let us point out that the participants are explicitly presenting only reasons of arguments (when speaking in terms of **RC** formalism); the claim, or conclusion (doing resp. not doing the action) is implicit.

### 3 Changes of Beliefs

#### 3.1 Incremental Update of the Partner Model

Let us consider the example dialogue in Sect. 2.2 in order to demonstrate in more details how the partner model is used in interaction. Let us suppose that a conversational agent is playing A's role. The communicative goal of A is to reach B's decision to do the action  $D =$  'to travel to N. in order to conclude a contract'. A will implement the tactics of enticement and generates a partner model, let it be  $w_{AB} = (w(\text{resources})=1, w(\text{pleasant})=3, w(\text{unpleasant})=2, w(\text{useful})=2, w(\text{harmful})=1, w(\text{obligatory})=0, w(\text{prohibited})=0, w(\text{punishment-do})=0,$

$-w(\text{punishment-not}) = 0$ ). The reasoning procedure WISH (Sect. 2.1) is applicable and yields a positive decision in this model. *A* tries to trigger the reasoning procedure WISH in *B*.

We suppose here that every statement (argument) presented in dialogue will increase (or respectively, decrease) the corresponding weight in a model of beliefs by one unit.

*A* starts the dialogue with a proposal. Using the tactics of enticement and attempting to trigger the reasoning procedure WISH in *B* he adds an argument to the proposal for increasing the pleasantness (turn 1). Therefore, he increases the initial value of the pleasantness in his partner model by 1. The current reasoning procedure WISH still gives a positive decision in the updated model. However, *B*'s counter argument (turn 2) demonstrates that *B* actually has resources missing (*I don't have proper dresses*) therefore, *A* has to change the value of  $w(\text{resources})$  in his partner model from 1 to 0. Now *A* has to find an argument indicating that the resources actually exist: he selects an utterance from the set  $P_{\text{increasingresources}}$  (*The company will pay your executive expenses*) and when following the tactics of enticement in turn 3 he adds an argument for increasing the pleasantness (*You can meet interesting people*). The value of  $w(\text{resources})$  will be 1 and the value of  $w(\text{pleasant})$  will be increased by 1 in the updated partner model. The reasoning in the updated model gives a positive decision. Nevertheless, *B* has a new counter argument indicating the harmfulness of the action: *I can have some problems at my university* (turn 4). This turn needs more comments. *B*'s statement for harmfulness increases the weight  $w(\text{harmful})$  in the partner model by 4 and not by 1. Why? Let us consider the step-form reasoning algorithm WISH (Sect. 2.1). There are four possible paths to achieve a negative decision (do not do *D*): coming through the steps 1–2–3–4–5 or 1–2–6–9–5 (if *D* is prohibited); 1–2–6–7–8 (if *D* is obligatory); or 1–2–6–7 (if *D* is neither obligatory nor prohibited). The last path is acceptable according the current model  $w_{AB}$  and the weight  $w(\text{harmful})$  indicated by *B* has to be increased so much that the condition checked in step 6 will be fulfilled – (at least) by 4.

Responding to *B*'s counter argument *A* decreases the value of  $w(\text{harmful})$  by 1 using the utterance *It's all right – your examinations period will be extended* and increases the value of  $w(\text{pleasant})$  once more using the utterance *You can sunbathe in *N.* early in spring already* (turn 5). The reasoning procedure WISH gives a positive decision in the updated partner model. Now it turns out that *B* has made this same decision (turn 6). *A* has achieved his communicative goal and finishes the dialogue (turn 7).

### 3.2 Changes of Own Beliefs

The last example (Sect. 3.1) demonstrates how *A* is updating the partner model  $w_{AB}$  in argumentation dialogue with *B*. As compared with the initial model, the values of two weights have been increased:  $w(\text{pleasantness})$  from 3 to 6 and  $w(\text{harmfulness})$  from 1 to 4. The changes have been caused by *A*'s arguments and *B*'s counter arguments.

Does the final model  $\mathbf{wAB}$  coincide with  $B$ 's actual model  $\mathbf{wB}$ , i.e. has  $A$  correctly guessed all the actual weights of  $B$ 's beliefs? The answer is 'not'. Let us discuss why. Let us again consider the example dialogue (Sect. 2.2). Let us suppose that  $B$  is a conversational agent (not a human user) and that  $B$ 's actual model is  $\mathbf{wB} = (0, 3, 2, 1, 5, 1, 1, 0, 0)$  at the beginning of the dialogue (different from  $\mathbf{wAB}$  as in Sect. 3.1). Thus,  $B$  considers  $D$  as an obligatory action therefore not doing  $D$  involves a punishment (differently from  $A$ 's picture about  $B$ ). In addition, let us suppose that  $B$ 's communicative goal coincides with  $A$ 's one:  $B$  has a wish to do  $D$  (it is more pleasant than unpleasant). She triggers a reasoning procedure WISH in her model of beliefs in order to check her resources and the other aspects of doing  $D$  and to make a decision.

In dialogue,  $A$  is updating his partner model  $\mathbf{wAB}$ . At the same time,  $B$  has to update the model  $\mathbf{wB}$  of herself as based on the arguments presented by  $A$ . Similarly with  $A$  who does not know the exact values of  $B$ 's beliefs in  $\mathbf{wB}$ , also  $B$  does not know the exact values of beliefs in the model  $\mathbf{wAB}$ . Both participants can make conclusions only based on arguments presented by the partner.

Let us suppose that  $A$  is acting as considered in Sect. 3.1.  $A$  makes a proposal to  $B$  and adds an argument which increases the weight  $w(\text{pleasantness})$  in the initial model of  $B$  herself by 1. The reasoning procedure WISH triggered by  $B$  in  $\mathbf{wB}$  gives a negative decision: resources are missing (*I don't have proper dresses*, turn 2).  $A$ 's next utterances (turn 3) increase the weights  $wB(\text{resources})$  and  $wB(\text{pleasant})$  by 1. Constructing her next turn (5)  $B$  again triggers the reasoning procedure WISH in the updated model  $\mathbf{wB}$  and over again comes to a negative decision. She chooses to indicate the harmfulness (*I can have some problems at my university*). When responding (turn 6)  $A$  presents an argument which decreases the harmfulness by 1 and when enticing he adds an argument which increases the pleasantness by 1 in the model  $\mathbf{wB}$ . Now  $B$ , after triggering the reasoning procedure WISH in her updated model, gets a positive decision (turn 6). Her final model will be  $\mathbf{wB} = (1, 6, 2, 1, 4, 1, 1, 0, 0)$ . The dialogue finishes, both  $A$  and  $B$  have achieved their common communicative goal.  $A$  has been able to convince  $B$  to make a decision to do  $D$  using the arguments by which  $B$  updated her initial model of beliefs in order to come to a positive decision. Although the models  $\mathbf{wAB}$  and  $\mathbf{wB}$  do not coincide at the end of the dialogue, the proportions of the weights of the positive (pleasantness, usefulness) and negative aspects of doing  $D$  (unpleasantness, harmfulness) are similar.

## 4 Discussion

When attempting to direct  $B$ 's reasoning to the desirable decision ("do  $D$ " in the considered example),  $A$  presents several arguments stressing the positive and downgrading the negative aspects of  $D$ . The choice of  $A$ 's argument is based, on one hand, on the (counter) argument presented by the partner and on the other hand, on the partner model. When choosing the next argument,  $A$  triggers a reasoning procedure in his partner model, in order to be sure that the reasoning will give a positive decision after presenting this argument.  $B$  herself can use



the same or a different reasoning procedure triggering it in her own model. (In the example, both participants are using the same reasoning procedure WISH.) After the updates made both by  $A$  and  $B$  in the two models during a dialogue, the models will approach each to another but, in general, do not equalise. Although, the results of reasoning in both models can be equal as demonstrated the example.

Therefore,  $A$  can convince  $B$  to do  $D$  even if not having a right picture of her. Our dialogue model considers only a limited kind of dialogues but although, it illustrates the situation where the dialogue participants are able to change their beliefs and bring them closer one to another by using arguments. The initiator  $A$  does not need to know whether the counter arguments of the partner  $B$  have been caused by  $B$ 's opposite goal or are there simply obstacles before their common goal and can be eliminated by arguments.  $A$ 's goal, on the contrary is not hidden from  $B$ . Secondly, as said in Sect. 2.1 the different communicative tactics used by  $A$  are aimed to trigger different reasoning procedures in  $B$ 's mind.  $A$  can fail to trigger the pursued communicative tactics but however, he can achieve his communicative goal when having a sufficient number of statements for supporting the communicative goal.

## 5 Conclusion and Further Work

We are considering the dialogues where two (human or artificial) agents  $A$  and  $B$  discuss about doing an action  $D$  by one of them ( $B$ ). Their initial communicative goals can conform or be opposite. They present arguments for and against of doing  $D$ , in order to achieve their goals.  $A$ 's arguments are based on his partner model whilst  $B$ 's arguments are based on her model of herself. Both models include the beliefs about the resources, positive and negative aspects of doing  $D$  which have numerical values (weights) in our implementation. Both models are changing during a dialogue. We study how the models are updated in a dialogue, and track the changes.

We have created two implementations – two experimental DSs which interact with a user using texts in a natural language. In one implementation, the computer is playing  $A$ 's role and in the other –  $B$ 's role. When attempting to direct  $B$ 's reasoning to the decision “do  $D$ ”,  $A$  presents several arguments (statements) stressing the positive and downgrading the negative aspects of  $D$ . The choice of statements is based on the partner model. Before bringing out an argument,  $A$  triggers a reasoning procedure in his partner model, in order to be sure that the reasoning will give a positive decision. When opposing,  $B$  can use the same or a different reasoning procedure triggering it in the model of herself. After the changes have been made by the participants during a dialogue, the two models of beliefs ( $A$ 's model of  $B$  and  $B$ 's model of herself) will approach each to other. The results of reasoning in both models can be (or not be) equal.

Our future work includes development of the implementations. When adding text and speech processing tools to a DS we can achieve more natural interaction of a user with the system.

**Acknowledgments.** This study was supported by the Estonian Ministry of Education and Research (IUT20-56), and by the European Union through the European Regional Development Fund (Centre of Excellence in Estonian Studies).

## References

1. Bos, J., Klein, E., Lemon, O., Oka, T.: DIPPER: description and formalisation of an information-state update dialogue system architecture. In: Proceedings of SIGDial Workshop on Discourse and Dialogue, Sapporo, pp. 115–124 (2003). [http://sigdial.org/workshops/workshop.4/proceedings/11\\_SHORT\\_bos\\_dipper.pdf](http://sigdial.org/workshops/workshop.4/proceedings/11_SHORT_bos_dipper.pdf)
2. Jurafsky, D., Martin, J.M.: Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall, Upper Saddle River (2009)
3. Cohen, P.R., Perrault, C.R.: Elements of a plan-based theory of speech acts. *Cogn. Sci.* **3**, 177–212 (1979)
4. Galitsky, B.: Exhaustive simulation of consecutive mental states of human agents. *Knowl.-Based Syst.* **43**, 1–20 (2012). <http://dx.doi.org/10.1016/j.knsys.2012.11.001>
5. Traum, D.R., Larsson, S.: The information state approach to dialogue management. In: van Kuppevelt, J., Smith, R.W. (eds.) *Current and New Directions in Discourse and Dialogue*. TSLT, vol. 22, pp. 325–353. Springer, Dordrecht (2003)
6. Nestorovič, T.: A frame-based dialogue management approach. In: 2nd International Conference on the Applications of Digital Information and Web Technologies, ICADIWT 2009, pp. 327–332 (2009). <http://dx.doi.org/10.1109/ICADIWT.2009.5273964>
7. Young, S., Schatzmann, J., Thomson, B., Weilhammer, K., Ye, H.: The hidden information state dialogue manager: a real-world POMDP-based system. In: Proceedings of NAACL HLT, Rochester pp. 27–28 (2007)
8. Williams, J.D.: A belief tracking challenge task for spoken dialog systems. In: Proceedings of NAACL HLT 2012 Workshop on Future Directions and Needs in the Spoken Dialog Community: Tools and Data, Association for Computational Linguistics, p. 2 (2012). <http://research.microsoft.com/pubs/163683/naaclhlt2012.pdf>
9. Koit, M., Õim, H.: A computational model of argumentation in agreement negotiation processes. *Argum. Comput.* **5**(2–3), 209–236 (2014). Taylor & Francis, <http://dx.doi.org/10.1080/19462166.2014.915233>
10. Amgoud, L., Besnard, P., Hunter, A.: Logical representation and analysis for RC-arguments. In: Proceedings of ICTAI, p. 8 (2015). [www.irit.fr/~Leila.Amgoud/ictai2015-1.pdf](http://www.irit.fr/~Leila.Amgoud/ictai2015-1.pdf)