Modelling Attitudes of a Conversational Agent

Mare Koit[a]
Institute of Computer Science, University of Tartu, J. Liivi 2, Tartu, Estonia

Keywords: Conversational Agent, Negotiation, Reasoning Model, Attitude, Knowledge Representation.

Abstract: The paper introduces a work in progress on modelling of attitudes of a conversational agent in negotiation. Two kinds of attitudes are under consideration: (1) related to different aspects of a negotiation object (in our case, doing an action) which direct reasoning about the action, and (2) related to a communication partner (dominance, collaboration, communicative distance, etc.) which are modelled by using the concept of multidimensional social space. Attitudes of participants have been annotated in a small sub-corpus of the Estonian dialogue corpus. An example from the sub-corpus is presented in order to illustrate how the models describe the change of attitudes of human participants. A limited version of the model of a conversational agent is implemented on the computer. Our further aim is to develop a dialogue system to train the user's negotiation skills by interacting with him in a natural language.

1 INTRODUCTION

Modelling of conversational agents and development of dialogue systems (DS) is aimed to make interaction of human users with the computer more convenient. Conversational agents communicate with users in natural language making travel arrangements, answering questions about weather or sports, routing telephone calls, acting as a general telephone assistant, or performing even more sophisticated tasks (Jurafsky and Martin, 2013).

We are studying the dialogues where one participant (*A*) requests her partner (*B*) to do an action *D* and proposes several arguments for doing *D*, trying to influence *B*. The paper introduces a model of conversational agent, concentrating on the attitudes of the agent when it is involved into negotiation with a user or with another conversational agent. Two kinds of attitudes are under consideration: (1) related to different aspects of a negotiation object (in our case, it is doing an action), and (2) related to a communication partner, or social attitudes.

Our aim is here to justify some aspects of the model on actual human-human dialogues in order to include these aspects into the DS that interacts with the user in the Estonian language. We currently limit us with verbal interaction and do not consider nonverbal means.

The paper is structured as follows. Section 2 describes the related work. Section 3 introduces a model of conversational agent, including representation of information states that the agent passes during negotiation as well as the participants' attitudes that are changing in negotiation. An authentic human-human negotiation is analyzed in Section 4 in order to demonstrate how well work the models of attitudes. Section 5 presents an implementation and Section 6 discusses how the model of conversational agent can be used when developing a DS. Section 7 draws conclusions.

2 RELATED WORK

Negotiation is simultaneously a linguistic and a reasoning problem, in which intent must be formulated and then verbally realized. Such dialogues require agents to understand, plan, and generate utterances to achieve their goals (Traum et al., 2008; Lewis et al., 2017). Automated negotiation agents capable of negotiating efficiently with people must rely on a good opponent modeling component to model their counterpart, adapt their behavior to their partner, influencing the partner's opinions and attitudes (Oshrat et al., 2009). There are several approaches to model change of a person's attitude, incl. the Elaboration

[a] https://orcid.org/0000-0002-7318-087X

Koit, M.
Modelling Attitudes of a Conversational Agent.
DOI: 10.5220/0008067102250232
In *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2019)*, pages 225-232
ISBN: 978-989-758-382-7
Copyright © 2019 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

Likelihood Model (ELM) that is a theory of thinking process, Social Judgment Theory that emphasizes the distance in opinions, and Social Impact Theory which emphasizes the number, strength and immediacy of the people trying to influence a person to change its mind. A naïve, intuitive model of functioning of the human mind reflects knowledge that human beings have and use about their partners in everyday communication (Õim, 1996).

For virtual agents, the expressions of attitudes in groups is a key element to improve the social believability of the virtual worlds that they populate as well as the user's experience, for example in entertainment or training applications (Ravenet et al., 2015; DeVaultand et al., 2015; Gratch et al., 2015; Callejas et al., 2014).

Concept of social attitude or interpersonal stance in interaction (being polite, distant, cold, warm, supportive, contemptuous, etc.) is considered in (Carofiglio, 2009). Ravenet et al. (2012) show the influence of dominance and liking in the nonverbal behavior depending on the gender of the speaker. These attitudes can be conveyed by words and voice features but also by nonverbal means – facial expression, body movement, and gestures (Knapp and Hall, 2009).

Computational approaches to dialogue fall into two categories of computational task: dialogue modelling and dialogue management. A dialogue system will have both dialogue modeling and dialogue management components (Traum, 2017). The functions of the dialogue manager can be formalized in terms of information state update. An information state includes beliefs, assumptions, expectations, goals, preferences and other attitudes of a dialogue participant's that may influence the participant's interpretation and generation of communicative behavior (Bunt, 2014).

An interesting and useful kind of dialogue systems rapidly developing in the last years are embodied conversational agents (Harthold et al., 2013; Ravenet et al., 2015; Dermouche, 2016; Jokinen, 2018). Such agents are interacting with human users in a natural language, and they are designed to play a certain social role in interaction.

3 CONVERSATIONAL AGENT AND INFORMATION STATE UPDATES

Let us consider interaction between a conversational agent A and its partner B (which can be whether another conversational agent or a human user). A initializes the interaction by requesting B to do an action D. The process is determined if the following is given (cf. Koit, 2018):

1) set G of communicative goals where both participants choose their own initial goals (G^A and G^B, respectively). In our case, G^A = "B makes a decision to do D"

2) set S of communicative strategies of the participants. A communicative strategy is an algorithm used by a participant for achieving his/her communicative goal. This algorithm determines the activity of a participant at each communicative step

3) set T of communicative tactics, i.e. methods of influencing the partner when applying a communicative strategy. For example, A can entice, persuade, or threaten B in order to achieve its goal G^A

4) set R of reasoning models which is used by participants when reasoning about doing the action D. A reasoning model is an algorithm the result of which is a positive or negative decision about the reasoning object (in our case, the action D)

5) set P of participant models, i.e. a participant's depiction of the attitudes himself/herself and his/her partner in relation to the reasoning object:

$P = \{P^A(A), P^A(B), P^B(A), P^B(B)\}$

6) set of world knowledge

7) set of linguistic knowledge.

A conversational agent passes several *information states* during interaction starting from initial state and going to following states by applying update rules. Information state represents cumulative additions from previous actions in the dialogue, motivating future actions. There are two parts of an information state of a conversational agent (Traum and Larsson, 2003) – private (information accessible only for the agent) and shared (accessible for both participants).

The private part of an information state of the conversational agent A consists of the following information: (a) current partner model and social attitudes in relation to the partner, (b) communicative tactics t_i^A which A has chosen for influencing B, (c) the reasoning model r_j which A is trying to trigger in B and bring it to the positive decision (it is determined by the chosen tactics, e.g. when enticing, A tries to increase B's wish to do D), (d) set of dialogue acts $DA=\{d_1^A, d_2^A, ..., d_n^A\}$ which A can use, (e) set of utterances for increasing or decreasing the values of B's attitudes in relation to D (arguments for/against of doing D) $U=\{u_{i1}^A, u_{i2}^A, ..., u_{iki}^A\}$.

The shared part of an information state contains (a) set of reasoning models $R=\{r_1,...,r_k\}$, (b) set of communicative tactics $T=\{t_1, t_2, ..., t_p\}$, and (c)

dialogue history $p_1{:}u_1[d_1]$, $p_2{:}u_2[d_2]$,..., $p_i{:}u_i[d_i]$ where $p_1=A$, p_2, etc. are A or B.

There are two categories of *update rules* that will be used by a conversational agent for moving from the current information state into the next one: (1) the rules used by the agent in order to interpret the partner's turns and (2) the rules used in order to generate its own turns.

3.1 Reasoning Model and Attitudes related to Conversation Object

The reasoning process of a subject about doing an action D consists of steps where the resources, positive and negative aspects of D will be weighed. A communication partner can take part in this process only implicitly by presenting arguments to stress the positive aspects of D and downgrade the negative ones.

Our used reasoning model includes two parts: (1) a model of (human) motivational sphere that represents the attitudes of a reasoning subject in relation to the aspects of the action under consideration, and (2) reasoning procedures. It is a kind of BDI (Belief-Desire-Intention) model (Bratman, 1999).

We represent *the model of motivational sphere* of a communication participant as a vector with numerical coordinates that express the attitudes of the participant in relation to different aspects of the action D (Koit and Õim, 2014):

w_D = (w(resources$_D$), w(pleasant$_D$), w(un¬ple¬asant$_D$), w(useful$_D$), w(harm¬ful$_D$), w(ob¬liga¬tory$_D$), w(punishme¬nt-not$_D$), w(pro¬hi¬bited$_D$), w(pu¬nish¬ment-do$_D$)).

Here w(pleasant$_D$), w(unpleasant$_D$), etc, indicate the pleasantness, unpleasantness, etc. of D or its consequences; w(punishment-do$_D$) is the punishment for doing a prohibited action and w(punishment-not$_D$) – the punishment for not doing an obligatory action. The value of w(resources$_D$) is 1 if the reasoning subject has all the resources needed for doing D (or 0 if some of the resources are missing), w(obligatory$_D$) is 1 if the action is obligatory for the subject (otherwise 0), w(prohibited$_D$) is 1 if the action is prohibited (otherwise 0). The values of the other coordinates can be numbers on the scale from 0 to 10. The model of motivational sphere is used by the subject when reasoning about doing D.

The reasoning itself depends on the determinant which triggers it. With respect to a naïve theory, there are three kinds of determinants that can cause humans to reason about an action D (Õim, 1996): his/her wish, need and obligation. Therefore, three different prototypical *reasoning procedures* can be described. Every procedure consists of steps passed by a reasoning subject and it finishes with a decision: do D or not. When reasoning in order to make a decision, B considers his resources as well as different positive and negative aspects of doing D. If the positive aspects (pleasantness, etc.) weigh more than negative (unpleasantness, etc.) then the decision will be "do D" otherwise "do not do D".

Presumption: w(pleasant) ≥ w(unpleasant).
1) Is w(resources) = 1? If not then go to 11.
2) Is w(pleasant) > w(unpleasant) + w(harmful)? If not then go to 6.
3) Is w(prohibited) = 1? If not then go to 10.
4) Is w(pleasant) > w(unpleasant) + w(harmful) + w(punishment-do)? If yes then go to 10.
5) Is w(pleasant) + w(useful) > w(unpleasant)+w(harmful) + w(punishment-do)? If yes then go to 10 else go to 11.
6) Is w(pleasant) + w(useful) ≤ w(unpleasant) + w(harmful)? If not then go to 9.
7) Is w(obligatory) = 1? If not then go to 11.
8) Is w(pleasant) + w(useful) + w(punishment-not) > w(unpleasant) + w(harmful)? If yes then go to 10 else go to 11.
9) Is w(prohibited) = 1? If yes then go to 5. else go to 10.
10) Decide: do D. End.
11) Decide: do not do D.

Figure 1: Reasoning procedure WISH.

Let us present two reasoning procedures (WISH and NEEDED) which will be used in the example in the following section. One procedure is triggered by the wish and the other by the need of the reasoning subject to do the action D. Both procedures are presented as step-form algorithms in Fig.1 and 2, respectively. (We do not indicate here the action D.)

Presumption: w(useful) ≥ w(harmful).
1) Is w(resources) = 1? If not then go to 8.
2) Is w(pleasant) > w(unpleasant)? If not then go to 5.
3) Is w(prohibited) = 1? If not then go to 7.
4) Is w(pleasant) + w(useful) > w(unpleasant) + w(harmful) + w(punishment-do)? If yes then go to 7 otherwise go to 8.
5) Is w(obligatory) = 1? If not then go to 8.
6) Is w(pleasant) + w(useful) + w(punishment-not) > w(unpleasant) + w(harmful)? if not then go to 8.
7) Decide: do D. End.
8) Decide: do not do D.

Figure 2: Reasoning procedure NEEDED.

We use two vectors ($w^B{}_D$ and $w^{AB}{}_D$) which capture the attitudes of communication participants in relation to the action D under consideration. Here $w^B{}_D$ is the model of motivational sphere of B who has to make a decision about doing D; the vector includes B's (actual) evaluations (attitudes) of D's aspects, it is used by B when reasoning about doing D. The other vector $w^{AB}{}_D$ is the partner model that includes A's

beliefs concerning *B*'s attitudes, it is used by *A* when planning the next turn in dialogue. We suppose that *A* has some preliminary information about *B* in order to compose the initial partner model before making the proposal to do *D*. Both the models will change as influenced by the arguments presented by the participants in negotiation. For example, every argument presented by *A* targeting the pleasantness of *D* should increase the corresponding values of $w^{B}_{D}(pleasant)$ as well as $w^{AB}_{D}(pleasant_{D})$.

3.2 Social Attitudes related to Communication Partner

In order to model the attitudes of a participant in relation to a communication partner we use the dimensions that characterize the relationships of participants in a communicative encounter. Communication can be collaborative or confrontational, personal or impersonal; it can also be characterized by the social distance of participants (near, far), etc. People have an intuitive, naïve theory of these dimensions; the values can be expressed by specific words. Still, we use numerical values as approximations to the words in our model (like in the model of human motivational sphere in the previous subsection). The following dimensions of such a 'social space' can be specified:

1. Dominance (on the scale from dominant to submissive)
2. Communicative distance to the partner (from near to far)
3. Cooperation (from collaborative to confrontational)
4. Politeness (from polite to impolite)
5. Personality (from personal to impersonal)
6. Modality (from friendly to hostile)
7. Intensity (from indolent to vehement).

We use the numbers 1, 0 and -1 for the values of the coordinates of social space. For example, the value of communicative distance is -1 if the person feels closeness in relation to his/her communication partner and 1 if he/she is far from the partner. The value 1 on the scale of modality means friendly and the value -1 means hostile interaction. On any scale, 0 is the neutral value. Still, a bigger number of values than three can be considered on every scale. It is also possible to use continuous scales instead of discrete values.

The attitudes of participants in relation to the partner can be represented by the (7-dimensional) vectors $\boldsymbol{s}^{AB}$ and $\boldsymbol{s}^{BA}$, respectively which determine the 'points' in social space. The participants can be located in different points. For example, customers sometimes angrily communicate with a service man who, as an official person, has to remain neutral or even friendly. Moreover, the participants can also 'move' from one point to another during communication. For instance, the participants who were on confrontational positions at the outset can reach the collaborative one at the end.

4 ANALYSIS OF HUMAN-HUMAN NEGOTIATION

In order to evaluate the model of conversational agent and especially, the models of attitudes, we are studying human-human dialogues.

Our current analysis is based on the Estonian dialogue corpus (Hennoste et al., 2008). The biggest part of the corpus – about 1000 spoken dialogues – is recorded in authentic situations and transcribed by using the transcription system of Conversation Analysis (Sidnell and Sivers, 2012). Each transcription is provided with a header that lists situational factors (meta-knowledge about the dialogue session), which affect language use – participants' names, social characteristics, relations between participants in the situation, specification of situation (private/public place, private/institutional conversation), etc.

For the current study, we have chosen five phone calls from the corpus where the acquaintances are negotiating about doing an action by one of them (cf. Koit, 2019). We are studying how the attitudes of the negotiation participants are changing and how well the models describe the changes. To do so, we annotated both the attitudes in relation to the negotiation object and in relation to the communication partner in every dialogue move using corresponding vectors.

In the next example (s. the following Table 1, left column) taken from our analyzed sub-corpus, mother *A* entices her son *B* to bake gingersnaps (action *D*).

She presents several arguments in order to increase the son's wish to do the proposed action by stressing the pleasantness of the action until he finally agrees. In the current study, the initial attitudes (the coordinates of the vectors $\boldsymbol{w}^{AB}$ and $\boldsymbol{w}^{B}$, s. Table 1) have been determined by an informal analysis of the whole dialogue text. Further, we suppose that every argument presented by *A* will change the targeted attitude (the pleasantness of *D*) by the value 1. (Therefore, we suppose that all the arguments have the weights equal to 1; still, this is a simplification.)

Table 1: Dialogue example (*A* – mother, *B* – son).

Utterances	*A*'s partner model w^{AB} (*A*'s beliefs about *B*'s attitudes in relation to *D*)	*B*'s model w^{B} (*B*'s actual attitudes in relation to *D*)	*A*'s attitudes s^{AB} in relation to *B*	*B*'s attitudes s^{BA} in relation to *A*
/---/ ***(1)*** *A*: ´küsimus. **A question.** (0.6) .hhhhh kas sulle pakuks ´pinget ´piparkookide ´küpsetamine. **Would you like to bake gingersnaps?** ***(2)*** *B*: ´praegu. **Just now?** ***(3)*** *A*: jah. **Yes.**	(1,6,2,1,1,1,0,0,0) WISH (Fig.1) gives the decision "do" (steps 1, 2, 3, 10)	(1,1,5,2,1,1,1,0,0) WISH is not applicable (presumption is not fulfilled: 1<5); NEEDED (Fig.2) gives the decision „do not do" (steps 1, 2, 5, 6, 8)	(1,1,1,0,1,1,0)	
(4) *B:* .hhhhhhh ma=i=´tea vist ´mitte. **I don't know, perhaps no.**	(1,6,2,**2**,1,1,1,0,0,0) WISH gives „do not do"			(-1,1,-1,0,1,0,0)
(5) *A*: ja=sis gla´suurimine=ja=´nii. **And then glazing and so on.** (0.6)	(1,2 **3**,2,1,1,1,1,0,0,0) WISH gives „do"	(1,**2**,5,2,1,1,1,0,0) NEEDED gives „do not do"	"	
(6) *B*: ´ei, ´ei, ´ei ei=´ei. **No, no, no, no, no.** (0.9)	(1,**3**,**2**,2,1,1,1,0,0,0) WISH gives „do not do"			(-1,1,-1,0,1,0,**1**) intensity – vehement
(7) *A*: me saaksime nad ´vanaema=jurde ´kaasa võtta. **We can take them with us when going to grandmother.**	(1,**2**,**3**,2,1,1,1,0,0,0) WISH gives „do"	(1,**3**,5,2,1,1,1,0,0) NEEDED gives „do not do"	"	
(8) *B*: ´präägu ei=´taha. **I just don't want.** (1.3)	(1,**2**,2,2,1,1,1,0,0,0) WISH gives „do not do"			(-1,1,-1,0,1,0,**0**) intensity – neutral
(9) *B*: aga (.) noh, kas sa mõtled nagu .hhh kui sa tuled ´koju=vä. **But what do you think – after you come home?** ***(10)*** *A*: .hhh ei **No.** ma mõtlen: kui mind kodus ei=´ole. **I think when I'm not home.** ***(11)*** *B*: aa. **Aha.** (0.5) .hhh et ´lähen ostan ´tainast=vä. **Then I'll go to buy paste, right?** ***(12)*** *A*: ja=niimodi=jah, **Yes, right.**	(1,**3**,2,2,1,1,1,0,0,0) WISH gives „do"	(1,**4**,5,2,1,1,1,0,0) NEEDED gives „do"	" "	(-1,1,**0**,0,1,0,0) cooperation – neutral (-1,1,**1**,0,1,0,0) cooperation – collaborative
(13) .hhh sinna:: ´Pereleiva ´kohvikusse võiksid minna @ ´võiksid seal endale ühe ´kohvi lubada=ja @ (2.7) teha ostmise ´mõnusaks=ja (0.8) ja=siis tulla ´koju=ja? (1.7) ´piparkooke teha=ja **And you could go to Pereleiva cafe and take a coffee in order to make buying pleasant for you, and then go home to bake gingersnaps.**	(1,**4**,2,1,1,1,0,0,0) WISH gives „do"		"	
(1.2) ***(14)*** *B*: okei? **OK.** /---/		(1,**5**,5,2,1,1,1,0,0) Both NEEDED and WISH are applicable, WISH gives „do"		"
(15) *A*: .hhhhhhhhhh (0.2) ja ´siis ma tahtsin sulle öelda=et ´külmkapis on: ´sulatatud või tähendab=ned ´külmutatud ja ´ülessulanud ´maasikad ja ´vaarika´mõmm.=hh **And I'd like to tell you that there are frozen strawberries and raspberries in the refrigerator.**	(1,**5**,2,1,1,1,0,0,0) WISH gives "do" (		"	
(16) *B*: jah **Yes.** ***(17)*** *A*: palun ´paku endale sealt. **Please help yourself.** /---/		(1,**6**,5,2,1,1,1,0,0) WISH gives „do"		"

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

A initiates the dialogue. Her communicative goal is to convince *B* to do the proposed action *D* (baking gingersnaps). *A* is using the partner model $\boldsymbol{w}^{AB}$ (corresponding to her image of *B*) by supposing that *B* has all the resources to do the action (the value of the first coordinate is 1), the pleasantness (6) is much greater than the unpleasantness (2), the usefulness and the harmfulness are equal (both 1), the action is obligatory (value 1) for *B* (because son is obliged to fulfil mother's requests). Still, mother will not punish (0) son if he does not fulfill the request. Further, the action being obligatory is not prohibited (0) and no punishment (0) will follow if it will be done. *A* applies the reasoning procedure WISH (Fig. 1) in the partner model and achieves the decision "do *D*" (s. the Table, 2nd column).

At the same time, *B*'s actual attitudes are different (model $\boldsymbol{w}^{B}$, s. the Table, 3rd column). He cannot apply the reasoning procedure WISH which presumption is not fulfilled (Fig.1). Instead, he applies the procedure NEEDED (Fig.2) which unfortunately gives the decion „do not do *D*" as indicated by *B*'s answer (utterance 4).

Now *A* has to introduce the changes into the partner model in order to get the same decision like *B* got and to present an argument for the pleasantness (utterance 5). Influenced by *A*'s argument, *B* increases his attitude in relation to the pleasantness (by 1 as we suppose) and applies the reasoning procedure NEEDED in his changed model.

The dialogue continues in the similar way until *A*'s argument presented by the utterance (10) makes it possible for *B* by applying the procedure NEEDED to get the decision "do" (11). It turns out that *A*'s next argument (13) will increase the pleasantness for *B* in such a way that he can apply the reasoning procedure WISH (prerequisite is fulfilled) therefore, at this moment he started to want to do the action. *A* does not stop but presents one more argument (15) which increases *B*'s wish once more. Therefore, both *A* and *B* finally achieved their communicative goal.

The initial *social attitudes* have been determined by using meta-knowledge about the dialogue session that are given in the header of the transcript and by the analysis of the first turns. Mother dominates over son (the value on the dominance scale 1 for mother and -1 for son), communicative distance is 'near' for the both participants (value 1), mother expresses cooperativity (value 1) but son – antagonism (value -1), the politeness is neutral (0), communication is personal (value 1), modality is 1 for mother and 0 for son, intensity is neutral for both. As seen in the Table (two last columns), mother keeps her social attitudes during the whole negotiation but son's utterance (6) – strong rejection – expresses the increased modality (value 1) which decreases to 0 in his next utterance (8). Son's utterance (9) demonstrates that antagonism has decreased – he started to doubt in his previous rejection and asks an adjusting question (value 0 on the cooperativity scale). By the utterance (11) son expresses cooperativity (value 1). Therefore, the final attitudes can be represented by the vectors $\boldsymbol{s}^{AB}$ = (1,1,1,0,1,1,0) and $\boldsymbol{s}^{BA}$ = (-1,1,1,0,1,1,0), respectively. Thus, the participants have approached one to another during the negotiation.

Here we do not consider the problem of how to determine automatically the initial attitudes of both kinds. This needs semantic analysis of dialogues that is currently not available for Estonian. In addition, recognition of social attitudes is hard without taking into account nonverbal means.

5 IMPLEMENTATION

A limited version of the conversational agent is implemented as a simple DS that interacts with a user in written Estonian. Information-state dialogue manager is used in the implementation. The programming language is Java.

The agent plays *A*'s role and the user *B*'s role. *A*'s communicative goal is "*B* will do *D*". The computer has ready-made sentences (assertions) for expressing of arguments, i.e., for stressing or downgrading the values of different aspects of the proposed action, which depend on its user model. The user (*B*) can choose one of two actions – traveling to a certain place or becoming a vegetarian. The user can optionally use another set of ready-made sentences or put in free texts. In the last case, keywords are used in order to analyze the texts. In the current implementation, the social attitudes of participants are not included.

Starting a dialogue, *A* determines a partner model $\boldsymbol{w}^{AB}{}_{D}$, fixes its communicative strategy and chooses the communicative tactics that it will follow, that is, the computer respectively determines a reasoning procedure that it will try to trigger in *B*'s mind. *A* applies the reasoning procedure in its partner model, in order to 'put itself' into *B*'s role and use suitable arguments when convincing *B* to decide to do *D*. The user is not obliged (but can) to follow neither certain communicative tactics nor reasoning procedures. He/she is also not obliged to fix his/her attitudes in relation to *D* by composing the vector $\boldsymbol{w}^{B}{}_{D}$. However, *A* does not 'know' *B*'s attitudes but it only can choose its arguments on the basis of *B*'s rejection and/or counterarguments. Respectively, *A* is making

changes in its partner model during a dialogue. The process runs in the similar way as described in the Table above.

6 DISCUSSION

Our model of conversational agent considers two kinds of attitudes of a dialogue participant: (1) the attitudes in relation of doing an action (which is the negotiation object), and (2) the attitudes related to a communication partner. Both kinds of attitudes are changing in dialogue as influenced by the arguments of the communication partner.

To model *the attitudes in relation to a reasoning object* (in our case, doing an action), we use the vector which coordinates correspond to the different aspects of the action (its pleasantness, usefulness, etc.). We evaluate these aspects by giving them discrete numerical values on a numerical scale. Still, people do not operate with numbers and what is more, with the exact values of the aspects of an action. They rather make 'fuzzy calculations', for example, they believe that doing an action is more useful than harmful and therefore it is needed to do it. In addition, if a reasoning object is different (not doing an action like in our case) then the attitudes of a reasoning subject can be characterized by a different set of aspects.

Social attitudes are modelled by using the concept of social space that dimensions correspond to the different relations of communication participants: dominance-subordination, social distance, politeness, etc. Currently, we use the values 1, 0 and -1 for the coordinates. Still, it is possible to operate with continuous scales instead of discrete values (cf. Mesiarová-Zemánková, 2016). It is also possible to use words of a natural language for the values. For example, the intensity of communication can be indolent, restrained, vehement, etc. However, annotation of the points of social space in written dialogues (transcripts of spoken dialogues) is difficult and subjective already with three different values (1, 0, -1). So far, we have not done that automatically.

Further empirical research is needed in order to determine the list of dimensions of communicative space, their relations and values on different scales (which can be different). Linguistic cues can be used for recognizing of values of some coordinates. For example, if a participant uses the 2nd person singular form of pronouns in Estonian then he/she is indicating a short communicative distance (-1) and a big value on the personality scale (+1). Feeling words can be used for recognizing the values of some coordinates, for example, *please* and *thank* indicate politeness. The comments of transcribers in transcripts of spoken dialogues help to determine the modality of communication (for example, the comment ((violently)) indicates the value 1 on the intensity scale). The dialogue act tags also contribute to recognizing of some coordinates. Opinion mining (Liu, 2015) can be used to automatically annotate communication points. Still, the small size of the Estonian dialogue corpus does not yet allow implement statistical or machine learning methods.

How to use the proposed models in human-computer systems? The conversational agents and especially, the conversational characters, which have recently become popular, take into account only the features of a limited field (e.g., a virtual guide of an art exhibition). At the same time, the agents can be created which could be 'tuned' to behave according to certain locations in social space depending on the user. For example, a travel agent gives information about a trip but it can also add various advices being neutral, advertising or even intrusive.

7 CONCLUSION

The paper describes a model of conversational agent that we are implementing in an experimental dialogue system. We are considering argument-based negotiations where the goal of the initiator (A) is to get the partner (B) to carry out a certain action D.

We consider two kinds of attitudes expressed by participants in negotiation about doing an action: (1) related to the action, and (2) related to a communication partner. We represent the first kind of attitudes as coordinates of a vector of motivational sphere of a participant who is reasoning whether to do an action or not. The second kind of attitudes is represented by using the concept of social space – a mental space where communication takes place.

We introduce information states of the conversational agent. Update rules allow move from one information state into another. An information state of the agent A includes a partner model that consists of evaluations of different aspects of the action under consideration. The attitudes included into the partner model as well as the actual attitudes of the partner are changing during the interaction, based on the arguments and counter-arguments presented. The conversational agent acts in social space that represents its attitudes related to the partner (social distance, politeness, etc.). The partner similarly behaves according his/her attitudes that are expressed in utterances. We consider here only verbal

communication, non-verbal means are not taken into account.

In order to evaluate our attitude models we analyze a small sub-corpus of the Estonian dialogue corpus formed by human-human phone calls. The analysis demonstrates how the attitudes of the negotiation participants are changing and how the different points in social space are visited during conversation.

We have implemented an experimental conversational agent, which argues for doing an action by interacting with the user in written Estonian. We believe that including the model of social space into the system will make the interaction more human-like. This needs more advanced processing of Estonian and remains for the further work.

ACKNOWLEDGEMENTS

This work was supported by institutional research funding IUT (20-56) of the Estonian Ministry of Education and Research, and by the European Union through the European Regional Development Fund (Centre of Excellence in Estonian Studies).

REFERENCES

Bratman, M.E. 1999. Intention, Plans, and Practical Reason, *CSLI Publications*.

Bunt, H. 2014. A context-change semantics for dialogue acts, *Computing Meaning*, 4, 177–201.

Callejas, Z., Ravenet, B., Ochs, M., Pelachaud, C. 2014. A Computational model of Social Attitudes for a Virtual Recruiter, *AAMAS,* 8 p.

Carofiglio, V., De Carolis, B., Mazzotta, I., Novielli, N., Pizzutilo, S. 2009. Towards a Socially Intelligent ECA, *IxD&A* 5-6, 99–106.

Dermouche, S. 2016. Computational Model for Interpersonal Attitude Expression, *ICMI*, 554–558.

DeVaultand, D., Mell, J., Gratch, J. 2015. Toward Natural Turn-Taking in a Virtual Human Negotiation Agent, *AAAI Spring Symposium on Turn-taking and Coordination in Human-Machine Interaction*, 9 p.

Gratch, J., Hill, S., Morency, L.-P., Pynadath, D., Traum, D. 2015. Exploring the Implications of Virtual Human Research for Human-Robot Teams, *VAMR*, 186–196.

Hartholt, A., Traum, D., Marsella, S.C., Shapiro, A., Stratou, G., Leuski, A., Morency, L.P., Gratch, J. 2013. All together now: Introducing the Virtual Human toolkit, *IVA,* 368–381.

Hennoste, T., Gerassimenko, O., Kasterpalu, R., Koit, M., Rääbis, A., and Strandson, K. 2008. From Human Communication to Intelligent User Interfaces: Corpora of Spoken Estonian, *LREC,* 2025–2032.

Jokinen, K. 2018. AI-based Dialogue Modelling for Social Robots, *FAIM/ISCA Workshop on Artificial Intelligence for Multimodal Human Robot Interaction*, 57-60.

Jurafsky, D., Martin, J. 2013. *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall; 2nd edition.

Knapp, M.L., Hall, J.A., 2009. *Nonverbal Communication in Human Interaction.* Wadsworth Publishing.

Koit, M. 2019. Changing of Participants' Attitudes in Argument-Based Negotiation, *ICAART*, 2, 770−777.

Koit, M. 2018. Reasoning and communicative strategies in a model of argument-based negotiation*, Journal of Information and Telecommunication TJIT*, 2, 14 p. Taylor & Francis Online

Koit, M., Õim, H. 2014. A computational model of argumentation in agreement negotiation processes, *Argument & Computation*, 5 (2-3), 209–236. Taylor & Francis Online.

Lewis, M., Yarats, D., Dauphin, Y.N., Parikh, D., Batra, D. 2017. Deal or No Deal? End-to-End Learning for Negotiation Dialogues, *EMNLP,* 2443–2453.

Liu, B. 2015. *Sentiment Analysis. Mining opinions, sentiments, and emotions*. Cambridge University Press.

Mensio, M., Rizzo, G., Morisio, M. 2018. The Rise of Emotion-aware Conversational Agents: Threats in Digital Emotions. *The 2018 Web Conference Companion,* Lyon, France, 4 p.

Mesiarová-Zemánková, A. 2016. Sensitivity analysis of fuzzy rule-based classification systems by means of the Lipschitz condition, *Soft Computing*, 103–113.

Oshrat, Y., Lin, R., Kraus, S. 2009. Facing the Challenge of Human-Agent Negotiations via Effective General Opponent Modeling, *AAMAS,* 377–384.

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

Sidnell, J., Stivers, T. (Eds.) 2012. *Handbook of Conversation Analysis*. Boston: Wiley-Blackwell.

Ravenet, B., Cafaro, A., Biancardi, B., Ochs, M., Pelachaud, C. 2015. Conversational Behavior Reflecting Interpersonal Attitudes in Small Group Interactions, *IVA*, 375–388.

Ravenet, B., Ochs, M., Pelachaud, C. 2012. A computational model of social attitude effects on the nonverbal behavior for a relational agent, *WACAI,* 9 p.

Rosenfeld, A., Zuckerman, I., Segal-Halevi, E., Drein, O., Kraus, S. 2014. NegoChat: A Chat-based Negotiation Agent, *AAMAS*, 525–532.

Traum, D. 2017. Computational Approaches to Dialogue, *The Routledge Handbook of Language and Dialogue*.

Traum, D., Larsson, S., 2003. The Information State Approach to Dialogue Management, *Current and New Directions in Discourse and Dialogue*, 325–353.

Traum, D., Marsella, S., Gratch, J., Lee, J., Hartholt, A. 2008. Multi-party, Multi-issue, Multi-strategy Negotiation for Multi-modal Virtual Agents, *IVA,* 117–130.