

CHANGING PROBABILISTIC BELIEFS IN PERSUASION

Katarzyna Budzyńska¹, Magdalena Kacprzak²

¹Institute of Philosophy, Cardinal Stefan Wyszyński University in Warsaw, Poland

²Faculty of Computer Science, Białystok University of Technology, Białystok, Poland

Abstract: The aim of the paper is to extend our formal model of persuasion with an aspect of change of uncertainty interpreted probabilistically. The general goal of our research is to apply this model to design a logic and a software tool that allow for verification of persuasive multi-agent systems (MAS). To develop such a model, we analyze and then adopt the Probabilistic Dynamic Epistemic Logic introduced by B. Kooi. We show that the extensions proposed in this paper allow us to represent selected aspects of persuasion and apply the model in the resource re-allocation problem in multi-agent systems.

Keywords: persuasion, beliefs, probabilistic logic, formal verification

1. Introduction

Persuasion plays an important role in resolving different problems in multi-agent systems (MAS). It allows agents to cooperate and perform collaborative decisions and actions since it is a tool for resolution of conflicts amongst agents (see e.g. [10]).

The general goal of our research is to develop a robust model of persuasion that will allow us to describe different phenomena specific to persuasive multi-agent systems. We concentrate on application of persuasion to resolution of the resource re-allocation problem (RrAP). This is the problem of effectively reallocating the resources such that all the agents have the resources they need. The formal model that we elaborate is used to develop a formalism (Logic of Actions and Graded Beliefs \mathcal{AG}_n [2]) and a software tool (the Perseus system [4]). The majority of existing work on agent persuasion considers protocols, which dictate what the possible legal next moves in persuasion are (e.g. [10]). We focus on verification of the persuasive systems for which protocols are already specified. The logic enables us to deductively test validity of formulas specifying agents participating in persuasion, as well as the

properties of systems that can be expressed via our model. The software allows us to semantically verify satisfaction of the logic formulas, which describe properties under consideration in a given model, as well as to perform parametrical verification that enables search for answers to questions about such properties.

In this paper, we focus on enriching the formal model of persuasion with an account of changing agents' uncertainty, interpreted probabilistically (this interpretation was insightfully studied in e.g. [1,5]; in this paper, however, we do not focus on the issue of probabilistic beliefs, but on the change of such beliefs). This provides a key first step towards extension of \mathcal{AG}_n into Probabilistic Logic of Actions and Graded Beliefs \mathcal{PAG}_n and a further development of the Perseus system. As far as we are aware, there are no other formal or software tools that allow verification of formulas with modalities expressing updates of probabilistic beliefs induced by persuasion.

The aspect of the uncertainty change in persuasion is important when we want to examine not only the final outcome of a given persuasion, but also to track how the successive actions modify agents' uncertainty about exchanging resources at each stage of persuasion (after the first persuasive action, after the second, etc.) [3]. This allows us to check and evaluate agents' strategies and, as a result, to plan optimal ones. The \mathcal{AG}_n logic enables expression of the uncertainty change in persuasion. The operator: $M_i^{d_1, d_2} \alpha$, intuitively means that an agent i considers d_2 doxastic alternatives (i.e. possible scenarios of a current global state) and d_1 of them satisfy α . Further, the operator: $\diamond(j : P)M_i^{d_1, d_2} \alpha$, intuitively means that after executing actions P by agent j , agent i may believe α with degree $\frac{d_1}{d_2}$. The strength of the \mathcal{AG}_n uncertainty operator is that it gives detailed information about local properties of a model we examine. For example, $M_i^{1,2} \alpha$ provides information that i assumes that α holds in exactly one state, while for $M_i^{2,4} \alpha$ the agent i assumes that α holds in two states. On the other hand, in pure \mathcal{AG}_n it is difficult to explore the uncertainty in terms of a ratio. Suppose that we want to examine if i believes α with degree $\frac{1}{2}$. To this end, we have to verify the formulas $M_i^{1,2} \alpha$, $M_i^{2,4} \alpha$, $M_i^{3,6} \alpha$ etc., since all of them describe the uncertainty ratio of $\frac{1}{2}$. A possible solution to this problem is to add the uncertainty operator interpreted probabilistically, since the probability is a natural way of expressing ratios. However, we must select a model, which would allow to describe not only the uncertainty, but also its change induced by persuasion.

In this paper, we examine a well-known framework proposed by Kooi [8]: Probabilistic Dynamic Epistemic Logic (PDEL). There are other logics that represent the change of degrees of beliefs, however, they do not refer to the probability in a direct manner. One such proposal is van Ditmarsch's model of graded beliefs

within Dynamic Epistemic Logic for Belief Revision [12]. In this framework, degrees of beliefs are related to agent's preferences which in turn correspond to a set of accessibility relations assigned to this agent. The other formalism is proposed by Laverny and Lang [9]. They define a graded version of the doxastic logic KD45 as the basis for the definition of belief-based programs and study the way the agents belief state is maintained when executing such programs.

Since our aim is to represent the change of probabilistic beliefs in persuasive MAS, the PDEL framework seems to be very promising. However, it has some serious limitations when directly applied to describe persuasion. A key contribution of this paper is that we not only identify those limitations but we also propose modifications that allow to avoid them.

The paper is organized as follows. Section 2 gives an overview of two frameworks that we explore in this paper: RrAP and PDEL. In Section 3, we propose the modifications to PDEL which are necessary if we want to apply it to the model of persuasion. In Section 4, we show how expressible the extended model is with respect to persuasion used in RrAP.

2. Background

In this section, we give a brief overview of the frameworks that we adopt to extend our model of persuasion. Moreover, we introduce an example that we use to illustrate our analysis in the next sections.

2.1 Resource re-allocation problem (RrAP)

The resource re-allocation problem can be intuitively described as the process of redistributing a number of items (resources) amongst a number of agents. During the resource re-allocation process, agents may disagree in some respects. Persuasion can provide a solution to such problems, since it allows resolution of conflicts. As a result, persuasion enhances the exchange of resources. Observe that in RrAP scenarios, persuasion may be accompanied by negotiations (see e.g. [7] for a framework enriching RrAP with negotiations), since conflict of opinion and conflict of interests often coexists.¹ However, for the clarity of the paper we limit our considerations to persuasion.

Recall that the general aim of our research is to build a logic and a software tool which will allow to verify the persuasive MAS. In this manner, we will be able to

¹ See [13] for details of a specification for persuasion and negotiation.

examine agents' strategies for exchanging resources and evaluate the correctness and effectiveness of applied algorithms for persuasion.

Consider the simplified example of RrAP. Assume a system with two agents: John and Ann. Both agents know that in the world they exist there are five keys, two of which are needed to open a safe. Ann knows identifiers of the appropriate keys and knows that John owns them. Therefore she tries to exchange the keys persuading John that after the exchange he will have the appropriate keys. John does not know which keys open the safe. Does he consent to the exchange?

Suppose that keys are marked with identifiers: 1, 2, 3, 4, 5. At the beginning Ann has the keys with identifiers 1, 2, 4, while John has keys 3 and 5. The keys which open the safe are also 3 and 5. Ann offers to John an exchange of key 2 for key 3. She justifies an action's necessity with a statement, which is obviously false, that in order to open the safe one odd and one even key is necessary. The John's response is strongly determined by his attitude to Ann. If John trusts Ann and knows that she is a reliable source of information, then he will agree to the keys' exchange and will believe that the pair of odd/even keys opens the safe. If John does not trust Ann, then he can respond in different ways (again, for simplicity we assume only two possible responses). The one manner is that John agrees to the keys' exchange, but he doesn't reset his beliefs. The other way determines that John assumes that Ann is not a credible source of information. Therefore, John does not accept the exchange and begins to believe that the safe may be opened only with a pair of odd/odd or even/even keys. As a result in the next sections we examine three cases:

- C1 John trusts Ann,
- C2 John does not trust Ann and is indifferent to her,
- C3 John does not trust Ann and believes the opposite of what she says.

2.2 Probabilistic Dynamic Epistemic Logic (PDEL)

In this section we show the syntax and semantics of PDEL introduced by Kooi [8]. Let $Agt = \{1, \dots, n\}$ be a finite set of names of *agents* and V_0 be a countable set of *propositional variables*.

The set of all well-formed expressions of PDEL is given by the following Backus-Naur form (BNF):

$$\alpha ::= p \mid \neg\alpha \mid \alpha \wedge \alpha \mid \Box_i \alpha \mid [\alpha_1] \alpha_2 \mid q_1 \mathbf{P}_i(\alpha_1) + \dots + q_k \mathbf{P}_i(\alpha_k) \geq q,$$

where $\alpha_1, \dots, \alpha_k$ are formulas, $p \in V_0$, $i \in Agt$, and q_1, \dots, q_k and q are rationals. For $q_1 \mathbf{P}_i(\alpha_1) + \dots + q_k \mathbf{P}_i(\alpha_k) \geq q$, the abbreviation $\sum_{j=1}^k q_j \mathbf{P}_i(\alpha_j) \geq q$ is used. Formulas

$\mathbf{P}_i(\alpha) = q, \mathbf{P}_i(\alpha) < q, \mathbf{P}_i(\alpha) \leq q, \mathbf{P}_i(\alpha) > q$ are defined from $\mathbf{P}_i(\alpha) \geq q$ in the classical way, e.g. $(\mathbf{P}_i(\alpha) \leq q)$ for $(\neg \mathbf{P}_i(\alpha) \geq -q)$ or $\mathbf{P}_i(\alpha) = q$ for $(\mathbf{P}_i(\alpha) \geq q) \wedge (\mathbf{P}_i(\alpha) \leq q)$.

The non-graded belief formula, $\Box_i \alpha$, says that i believes that α . The probabilistic belief formula, $\mathbf{P}_i(\alpha) \geq q$, means that the probability, i assigns to α , is greater than or equal to q . A formula for updates, $[\alpha_1] \alpha_2$, says that α_2 is the case, after everyone simultaneously and commonly learns that α_1 is the case.

By a probabilistic epistemic model we mean a Kripke structure $\mathcal{M} = (S, R, v, P)$ where

- S is a non-empty set of states (possible worlds),
- $R : Agt \longrightarrow 2^{S \times S}$ assigns to each agent an accessibility relation,
- $v : V_0 \longrightarrow 2^S$ is a valuation function,
- P assigns a probability function to each agent at each state such that its domain is a non-empty subset of S
 $P : (Agt \times S) \longrightarrow (S \rightarrow [0, 1])$ such that
 $\forall i \in Agt \forall s \in S \sum_{s' \in dom(P(i,s))} P(i,s)(s') = 1$,
 where \rightarrow means that it is a partial function, i.e., some states may not be in the domain of the function.

The semantics of formulas of PDEL are defined by two interdependent definitions with respect to a Kripke structure \mathcal{M} . The first definition gives the semantics for the PDEL language, and the second, for updates.

Definition 1 For a given structure $\mathcal{M} = (S, R, v, P)$ and a given state $s \in S$ the Boolean value of the formula α is denoted by $\mathcal{M}, s \models \alpha$ and is defined inductively as follows:
 $\mathcal{M}, s \models p$ iff $s \in v(p)$, for $p \in V_0$,
 $\mathcal{M}, s \models \neg \alpha$ iff $\mathcal{M}, s \not\models \alpha$,
 $\mathcal{M}, s \models \alpha \wedge \beta$ iff $\mathcal{M}, s \models \alpha$ and $\mathcal{M}, s \models \beta$,
 $\mathcal{M}, s \models \Box_i \alpha$ iff $\mathcal{M}, s' \models \alpha$ for all s' such that $(s, s') \in R(i)$,
 $\mathcal{M}, s \models [\alpha] \beta$ iff $(\mathcal{M}_\alpha, s_\alpha) \models \beta$ (see Definition 2),
 $\mathcal{M}, s \models \sum_{j=1}^k q_j \mathbf{P}_i(\alpha_j) \geq q$ iff $\sum_{j=1}^k q_j P(i, s)(\alpha_j) \geq q$,
 where $P(i, s)(\alpha_j) = P(i, s)(\{s' \in dom(P(i, s)) \mid \mathcal{M}, s' \models \alpha_j\})$.

Definition 2 Let a model $\mathcal{M} = (S, R, v, P)$ and a state $s \in S$ be given. The updated model $\mathcal{M}_\alpha = (S_\alpha, R_\alpha, v_\alpha, P_\alpha)$ is defined as follows:

- $S_\alpha = S$,
- $R_\alpha(i) = \{(s, s') \mid (s, s') \in R(i) \text{ and } \mathcal{M}, s' \models \alpha\}$,
- $v_\alpha = v$,
- $dom(P_\alpha(i, s)) = dom(P(i, s))$ if $P(i, s)(\alpha) = 0$ and $dom(P_\alpha(i, s)) = \{s \in$

$dom(P(i, s)) : (\mathcal{M}, s) \models \alpha$ otherwise,
 - $P_\alpha(i, s)(s') = P(i, s)(s')$ if $P(i, s)(\alpha) = 0$ and $P_\alpha(i, s)(s') = \frac{P(i, s)(s')}{P(i, s)(\alpha)}$ otherwise.

The public announcement α updates the model by changing the accessibility relations and probability functions. The only states that remain accessible for each agent are the states where α holds. The probability functions work in a similar way. Their domains become limited to the states where α holds.

2.3 PDEL in RrAP-example

Let us analyze the initial step in the persuasion dialogue described in Section 2.1 with the use of PDEL. We need to define a probabilistic epistemic model \mathcal{M} . Assume that $Agt = \{John, Ann\}$ and $V_0 = \{p, even_odd\}$, where p means that John has the correct set of keys, which enables him to open the safe and $even_odd$ means that the combination of the pair of one even and one odd key opens the safe. The set of states is the set $S = \{(A, J, C) : A, J, C \subseteq \{1, 2, 3, 4, 5\}, |C| = 2, A \cap J = \emptyset, A \cup J = \{1, 2, 3, 4, 5\}\}$. Thus a state $s = (A, J, C) \in S$ consists of three sets. The first one is a set of Ann's keys. The second is the set of John's keys. The intersection of the sets A and J is the empty set because the resources can not be shared. The union of these sets equals $\{1, 2, 3, 4, 5\}$ because Ann and John own all accessible resources. The third set, C , is a set of the keys which open the safe. Cardinality of C equals 2 since there are exactly two correct keys.

In this model there are two propositions: p and $even_odd$. Proposition p is true in every state in which the set of keys opening the safe is a subset of the set of keys owned by John, i.e., $v(p) = \{s \in S : s = (A, J, C) \text{ and } C \subseteq J\}$. Proposition $even_odd$ is true in states in which even/odd combination of keys opens the safe, i.e., $v(even_odd) = \{s \in S : s = (A, J, C) \text{ and } C = \{1, 2\} \text{ or } C = \{1, 4\} \text{ or } C = \{2, 3\} \text{ or } C = \{2, 5\} \text{ or } C = \{3, 4\} \text{ or } C = \{4, 5\}\}$.

Moreover assume that Ann has all the information about the actual state, i.e., when she is at state s she knows that she is at this state. As a result, her accessibility relation is defined as follows: $R(Ann) = \{(s, s') \in S^2 : s = s'\}$. John knows the keys he has and knows that Ann has the other keys, so his accessibility relation is $R(John) = \{(s, s') \in S^2 : s = (A, J, C), s' = (A', J', C'), J' = J, A' = A\}$.

Furthermore, say that the probability function P is as follows: $P(i, s)(s') = \frac{1}{|dom(P(i, s))|}$ for every $i \in Agt, s, s' \in S$ where $dom(P(i, s)) = \{s'' \in S : (s, s'') \in R(i)\}$. Notice that we propose to define the probability function in a classical way, i.e. we assume that $P(i, s)(s')$ is the quotient of 1 (one state s') and the number of all elements belonging to the domain of probability. This means that every state s' accessible

from s has assigned the same probability. Furthermore, we assume that the domain of probability function is a set of all accessible states. In this case there are no reasons to separate these sets. For example, if at state s John considers 10 accessible states s' then $P(\text{John}, s)(s') = \frac{1}{10}$ for every state s' .

At the beginning Ann has the keys 1, 2, and 4, John has the keys 3 and 5 and the same keys open the safe. So the initial state is $s_0 = (\{1, 2, 4\}, \{3, 5\}, \{3, 5\})$. John's accessibility relation and probability function for s_0 are depicted in Fig. 1.

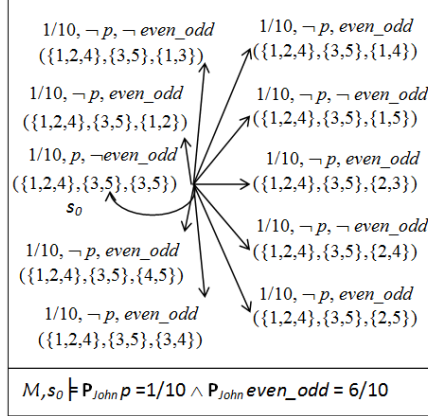


Fig. 1. John's accessibility relation and probability function before persuasion.

3. Adaptation of PDEL to persuasion model

In this section we analyze the limitations of PDEL with respect to representing the persuasion in RrAP. Moreover, we propose modifications that allow those limitations to be overcome.

In order to study a persuasive situation from the example we need to express the issues related to three stages of the persuasion:

- before the persuasion: John's attitude to the statement p – “John has the good (opening the safe) couple of keys”. In other words, how strongly does he believe it to be true? Formally we can ask: to what degree does John believe that p holds in the initial state s_0 ?

- John’s attitude to Ann’s persuasive actions (e.g. proposal of the keys’ exchange). In what manner might he react? What can he do and what will the result of John’s behavior be?
- after the persuasion: John’s attitude to the statement p . Does the degree of his belief change? If yes, how big a change is it?

The issues from the initial and the final stage can be successfully described in PDEL. Recall that John does not know the identifiers of keys which open the safe. Therefore at the beginning he considers all the possibilities (see Fig. 1). Since there are 10 possible situations and in only one of them p is true, in John’s opinion the probability that in actual state he has good keys is $\frac{1}{10}$. Formally:

$$\mathcal{M}, s_0 \models \mathbf{P}_{John}(p) = \frac{1}{10}.$$

Similarly, we could compute the probability which John assigns to p when the persuasion is finished. However, we must first know and represent what happened in the intermediate stage of the persuasion. This section discusses the problems that we encounter when we want to express the issues belonging to that stage.

3.1 Trusting the persuader

Say that the first action that Ann performs is a public announcement that p is true. According to the PDEL definition of the satisfiability relation, it holds that $\mathcal{M}, s \models \mathbf{P}_{John}(p) = 1$ for a state s reachable from s_0 after execution of Ann’s action. Observe that in PDEL semantics, the outcome of an action is not related to the performer of the action. That is, John’s reaction will be exactly the same regardless of whether Ann or someone else says that p . Similarly, it is impossible to make John’s behavior dependent on his attitude to Ann.

In agents’ interactions (such as persuasions or negotiations), the reputation of the agent who performs the action can influence possibility of the action’s success. In MAS, this problem is studied within the Reputation Management framework (see e.g. [11,14]). Say that an agent knows a persuader, since they exchanged resources before. If the agent evaluates those exchanges as beneficial and fair, then he will trust the persuader and be easily persuaded the next time they are going to exchange resources. On the other hand, if an announcement is executed by an agent which is unknown to other agents, then it may be disregarded. The first limitation of the PDEL expressivity is:

L1 The success of persuasion cannot be affected by reputation of persuader.

To avoid the limitation **L1**, we need to label every action with an agent who performs it. However this is not sufficient on its own. First of all we need to be able to express agents' attitudes to each other, i.e., which agent is perceived as credible by which agent. Therefore, we need to add a trust function T to the model, so now $\mathcal{M}^T = (S, R, v, P, T)$. A trust function assigns to every pair of agents one of the values $0, \frac{1}{2}, 1$, i.e.,

$$T : \text{Agt} \times \text{Agt} \rightarrow \{0, \frac{1}{2}, 1\}.$$

With respect to the three cases from the example, the interpretation of T can be as follows:

- C1** if $T(\text{John}, \text{Ann}) = 1$ then John trusts Ann and accepts everything she says,
- C2** if $T(\text{John}, \text{Ann}) = \frac{1}{2}$ then John is indifferent to Ann, and as a result Ann's announcement does not influence John's beliefs,
- C3** if $T(\text{John}, \text{Ann}) = 0$ then John does not trust Ann and what is more he is sure that she always tells lies.

In future work we plan to extend this approach and introduce more trust degrees and allow agents to adopt various attitudes to each other. In order to do this we intend to adopt the well-known solution from the Reputation Management framework proposed by Yu and Singh [14].

3.2 Public announcement as argument

Recall that in PDEL after Ann's announcement that p is true, it holds that $\mathcal{M}, s \models \mathbf{P}_{\text{John}}(p) = 1$. It means that after the action of announcing p , John *must* believe that p . Thereby we deprive John of deciding whether Ann is right or not. In persuasive scenarios it is a strong limitation, since it assumes that an audience will believe everything a persuader says. The only exception is when the audience believes the persuader's claim with probability 0. In other words it is impossible to express within this framework reactions of indifference (i.e. an audience is neutral with respect to a persuader) and other reactions of distrust (e.g., maximal distrust, i.e. when before a persuasion dialogue an audience believes p and – after the announcement that p – he begins to believe $\neg p$). So, the second limitation is:

- L2** Audience must believe everything that a persuader claims, unless the claim is believed by audience with probability 0.

In order to avoid the limitation **L2**, the syntax of formula $[\alpha]\beta$ can be exchanged with $[j : \alpha]\beta$ where $j \in \text{Agt}$. Then

$$\mathcal{M}^T, s \models [j : \alpha]\beta \text{ iff } \mathcal{M}_{j,\alpha}^T, s_{j,\alpha} \models \beta$$

where $\mathcal{M}_{j,\alpha}^T = (S_{j,\alpha}; R_{j,\alpha}; v_{j,\alpha}; P_{j,\alpha}; T_{j,\alpha})$ is an updated model such that:

- if $T(i, j) = 1$ then $\mathcal{M}_{j,\alpha}^T = \mathcal{M}_\alpha^T$,
- if $T(i, j) = \frac{1}{2}$ then $\mathcal{M}_{j,\alpha}^T = \mathcal{M}^T$,
- if $T(i, j) = 0$ then $\mathcal{M}_{j,\alpha}^T = \mathcal{M}_{-\alpha}^T$,

where \mathcal{M}_α^T is a model \mathcal{M}_α (see Definition 2) extended with the trust function T .

In the running example:

C1 if $T(\text{John}, \text{Ann}) = 1$ then

$$\mathcal{M}^T, s_0 \models [\text{Ann} : \text{even_odd}](\mathbf{P}_{\text{John}}(\text{even_odd}) = 1),$$

i.e., if John trusts Ann then he agrees with everything she says (see Case 1 in Fig. 2 for John's probability function),

C2 if $T(\text{John}, \text{Ann}) = \frac{1}{2}$ then

$$\mathcal{M}^T, s_0 \models ([\text{Ann} : \text{even_odd}]\mathbf{P}_{\text{John}}(\text{even_odd}) = \frac{6}{10})$$

i.e., if John is indifferent to Ann then he does not change his beliefs (his probability function is the same as before the announcement – see Fig 1),

C3 if $T(\text{John}, \text{Ann}) = 0$ then

$$\mathcal{M}^T, s_0 \models [\text{Ann} : \text{even_odd}](\mathbf{P}_{\text{John}}(\text{even_odd}) = 0)$$

and

$$\mathcal{M}^T, s_0 \models [\text{Ann} : \text{even_odd}](\mathbf{P}_{\text{John}}(\neg\text{even_odd}) = 1),$$

i.e., if John does not trust Ann then he adopts the opposite of what she says (see Case 3 in Fig. 2).

3.3 Unpersuadable audience

Observe that, according to the semantics adapted from PDEL, for the formula $[j : \alpha]\beta$ it holds:

- if $\mathcal{M}^T, s \models (\mathbf{P}_i\beta = 1)$ then $\mathcal{M}^T, s \models [j : \alpha](\mathbf{P}_i\beta = 1)$ and
- if $\mathcal{M}^T, s \models (\mathbf{P}_i\beta = 0)$ then $\mathcal{M}^T, s \models [j : \alpha](\mathbf{P}_i\beta = 0)$,

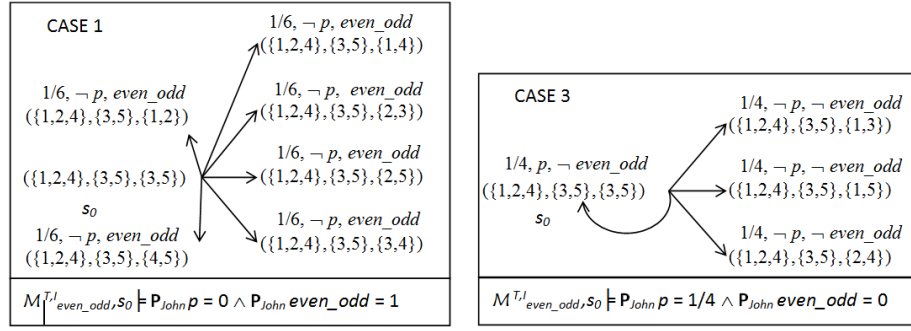


Fig. 2. John's probability function after Ann's public announcement that *even_odd* – cases 1 and 3.

for any formula α . Intuitively it means that if an agent i is sure that β is true then there is no way (no action which can be executed) to convince him that β is true with probability less than 1. A similar situation occurs when i is absolutely certain that β is false, i.e., if the probability of β is 0. In the context of persuasion scenarios, it is a serious limitation. For example, if an agent is sure that he needs some resources, then the other agent has no chance to persuade him to exchange it. So, the next problem with the PDEL expressivity is:

L3 A persuader has no chance to influence an audience about a claim in a case where it is absolutely sure that the claim is true or false.

The limitation **L3** is a consequence of an assumption that $\text{dom}(P_\alpha(i, s)) \subseteq \text{dom}(P(i, s))$. For instance, suppose that at state s_0 Ann says *even_odd*. Then $\mathcal{M}_{\text{even_odd}, s_0}^T \models \mathbf{P}_{\text{John}}(\neg \text{even_odd}) = 0$. Next Ann says $\neg \text{even_odd}$. Now, since John believes $\neg \text{even_odd}$ with probability 0, both probability function and its domain are not changed (see Definition 2). As a result, John's beliefs remain unchanged, i.e. $\mathcal{M}_{\neg \text{even_odd}, s_0}^T \models \mathbf{P}_{\text{John}}(\neg \text{even_odd}) = 0$.

In persuasion we must often deal with updates with information that has probability zero. The approach given in PDEL is simply to ignore the information. This is to ensure that one does not divide by zero. Moreover, the logic cannot deal well with updates with inconsistent information. Typically, the accessibility relation become empty after an inconsistent update. There is no philosophical reason for such a choice. However, this makes the system and its completeness proof relatively simple. There are some more advanced approaches in probability theory for updating sentences with probability 0 (see [6] for an overview). We propose to cope with this limitations in the way described below.

In order to resolve the problem **L3**, we can allow that after an update an agent may take into consideration a state which was not considered before. Formally, we assume that there exists a state s such that $s \in \text{dom}(P_\alpha(i, s))$ and $s \notin \text{dom}(P(i, s))$, i.e., $\text{dom}(P_\alpha(i, s)) \not\subseteq \text{dom}(P(i, s))$. Of course in the general case it may be a big challenge to establish which states with what probabilities can be added to a domain of function P . However, in some concrete applications it seems to be easy and natural. In our example it can work as follows. Let

$$\begin{aligned} \text{dom}(P(\text{John}, s_0)) &= \{(\{1, 2, 4\}, \{3, 5\}, C) \in S : \\ &C \subseteq \{1, 2, 3, 4, 5\} \text{ and } |C| = 2\} \end{aligned}$$

and in the updated model $\mathcal{M}_{\text{even_odd}}^T$

$$\begin{aligned} \text{dom}(P_{\text{even_odd}}(\text{John}, s_0)) &= \\ \{(\{1, 2, 4\}, \{3, 5\}, \{n_1, n_2\}) \in S : n_1 \text{ is an even and } n_2 \text{ is an odd number}\}. \end{aligned}$$

Hence $\mathcal{M}_{\text{even_odd}}^T, s_0 \models \mathbf{P}_{\text{John}}(\neg \text{even_odd}) = 0$. Next if Ann says $\neg \text{even_odd}$ then

$$\begin{aligned} \text{dom}(P_{\neg \text{even_odd}}(\text{John}, s_0)) &= \\ \{(\{1, 2, 4\}, \{3, 5\}, \{n_1, n_2\}) : n_1, n_2 \text{ are even or odd numbers}\} \end{aligned}$$

and John's beliefs are changed, i.e. $\mathcal{M}_{\neg \text{even_odd}}^T, s_0 \models \mathbf{P}_{\text{John}}(\neg \text{even_odd}) = 1$.

3.4 Nonverbal actions during the persuasion process

In persuasion, the proponent aims to change beliefs of the audience. The persuasion process begins with the first action of the proponent which has a given aim, and finishes with the last action with this aim. Yet, during persuasion agents can perform actions (with or without persuasive aims) which change not only beliefs of the audience, but also the environment of the agents. For example, during their persuasion dialogue, John and Ann can exchange the keys (i.e. before and after the action of exchange Ann performs some persuasive actions). Observe that this action can change the circumstances in which the persuasion process will continue. That is, the new circumstances can be favorable to Ann and her next persuasive action can make John believe her claim, while in the old circumstances such an effect may not be obtained. In other words, during the persuasion process some nonverbal actions influencing the environment can change a course (an effect) of persuasive actions performed after this nonverbal action.

The pure PDEL allows expression only of the actions of public announcement which do not influence the beliefs of an agent (a doxastic relation), no those which influence the environment (a state). In particular, it is not possible to describe such

situations in RrAP in which before a persuasion agents have some resources and after it they have other resources. Moreover, verbal actions (like public announcement) do not change values of propositions. This means that in this framework persuasive actions can not change the actual world (only the probabilistic beliefs may be modified). Thus the exchange of resources which is necessary to solve RrAP can not be described. Therefore the last limitation is:

L4 There is no possibility of expressing actions other than public announcements.

Since nonverbal actions are often applied in persuasion, we need to resolve the limitation **L4**. To this end, we propose to combine PDEL with our logic \mathcal{AG}_n . The strength of \mathcal{AG}_n is that it is already adjusted to express persuasion. Assume a set Π_{nv} of nonverbal actions and enrich the model \mathcal{M}^T with interpretation I of this actions where

$$I : \Pi_{nv} \rightarrow (Agt \rightarrow 2^{S \times S}).$$

Now we have a new model $\mathcal{M}^{T,I} = (S, R, v, P, T, I)$. After the execution of an action $a \in \Pi_{nv}$, a system reaches a state in which not only new accessibility relations can be assigned to agents but also new logical values may be assigned to propositions. Next, let formula $[j : a]\beta$ for $j \in Agt$ and $a \in \Pi_{nv}$ says that after the execution of action a by agent j the condition β holds. The semantics of this formula is as follows:

$$\begin{aligned} \mathcal{M}^{T,I}, s \models [j : a]\beta \text{ iff } \mathcal{M}^{T,I}, s' \models \beta \\ \text{for every state } s' \text{ such that } (s, s') \in I(a)(j). \end{aligned}$$

In our example Ann intends to exchange the key 2 with the key 3. Let ex stand for the action of the keys exchange. The interpretation of ex is given below. If $s = (A, J, C)$ is a state such that $2 \in A$ and $3 \in J$ then for every state $s' = (A', J', C')$ it holds

$$(s, s') \in I(ex)(Ann) \text{ iff } A' = A \setminus \{2\} \cup \{3\}, J' = J \setminus \{3\} \cup \{2\}, \text{ and } C' = C.$$

Otherwise $(s, s') \in I(ex)(Ann)$ iff $s = s'$.

4. Expressivity of extended model persuasion

Now we are ready to analyze the running example. At the beginning at state s_0 John assigns to the proposition p (the statement ‘‘John has the good keys’’) probability $\frac{1}{10}$ (see Fig. 1):

$$\mathcal{M}^{T,I}, s_0 \models (\mathbf{P}_{John} p = \frac{1}{10}).$$

Next Ann says that the combination of one even and one odd key opens the safe. Again consider the three cases:

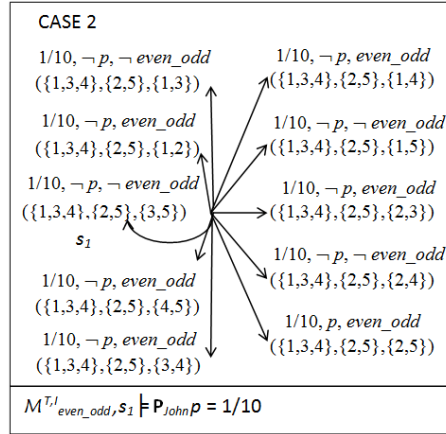


Fig. 3. John's probability function after the execution of the action ex – case 2.

C1 John trusts Ann and thinks that she is right. As a result, he removes from the domain of the probability function all states in which p is satisfied and thus assigns to p probability 0:

$$\mathcal{M}^{T,I}, s_0 \models [\text{Ann} : \text{even_odd}] (\mathbf{P}_{\text{John}} p = 0)$$

since (see Fig. 2): $\mathcal{M}_{\text{even_odd}}^{T,I}, s_0 \models (\mathbf{P}_{\text{John}} p = 0)$. Then Ann and John exchange keys 2 and 3. It is easy to compute that after this action the probability of p will be $\frac{1}{6}$:

$$\mathcal{M}^{T,I}, s_0 \models [\text{Ann} : \text{even_odd}] [\text{Ann} : ex] (\mathbf{P}_{\text{John}} p = \frac{1}{6})$$

since (see Fig. 4): $\mathcal{M}_{\text{even_odd}}^{T,I}, s_1 \models (\mathbf{P}_{\text{John}} p = \frac{1}{6})$ where s_1 is a state reachable after the execution of the action ex . Therefore the replacement causes the growth of probability which John assigns to the statement that he has right keys.

C2 John preserves a neutral position with respect to Ann. Hence, he assigns to p the probability as at the start of persuasion dialogue:

$$\mathcal{M}^{T,I}, s_0 \models [\text{Ann} : \text{even_odd}] (\mathbf{P}_{\text{John}} p = \frac{1}{10}).$$

Then Ann and John exchange keys 2 and 3. After the action the probability of p remains unchanged:

$$\mathcal{M}^{T,I}, s_0 \models [\text{Ann} : \text{even_odd}] [\text{Ann} : ex] (\mathbf{P}_{\text{John}} p = \frac{1}{10})$$

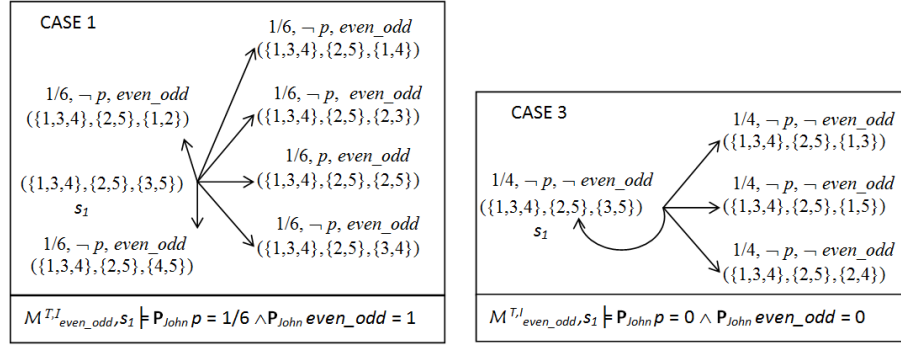


Fig. 4. John's probability function after the execution of the action *ex* – cases 1 and 3.

since (see Fig. 3): $\mathcal{M}_{\text{even_odd}, s_1}^{T,I} \models (\mathbf{P}_{\text{John}} p = \frac{1}{10})$. As a result, such an activity does not modify John's beliefs about whether he has right keys.

C3 John thinks that Ann lies. As a result he assigns to p degree $\frac{1}{4}$:

$$\mathcal{M}^{T,I}, s_0 \models [\text{Ann} : \text{even_odd}] (\mathbf{P}_{\text{John}} p = \frac{1}{4})$$

since (see Fig. 2) $\mathcal{M}_{\text{even_odd}, s_0}^{T,I} \models (\mathbf{P}_{\text{John}} p = \frac{1}{4})$.

Then Ann and John exchange keys 2 and 3. After the action the probability of p will be 0:

$$\mathcal{M}^{T,I}, s_0 \models [\text{Ann} : \text{even_odd}] [\text{Ann} : \text{ex}] (\mathbf{P}_{\text{John}} p = 0)$$

since (see Fig. 4): $\mathcal{M}_{\text{even_odd}, s_1}^{T,I} \models (\mathbf{P}_{\text{John}} p = 0)$. For that reason the exchange results in John believing that he has the right keys with degree 0.

5. Conclusions

PDEL is a powerful tool which can be used for reasoning about update of an agent's uncertainty. In this paper, we analyze the possibility of applying this framework to represent change of probabilistic beliefs induced by persuasion and executed in the resource re-allocation scenarios. First, we indicate some limitations of PDEL, when it is directly interpreted in a persuasive MAS. Next, we propose how to avoid those limitations such that the advantages of the PDEL tool could be fully exploited to represent the persuasion in RrAP.

We discuss four limitations: **L1** requires that the success of persuasion cannot be affected by the reputation of the persuader, **L2** limits the audience to believe

everything that the persuader says, with the exception when **L3** holds, i.e., when the audience is absolutely sure that the claim is true or false, in which case it is impossible to change the audience's mind, and **L4** does not allow expression of actions other than public announcements. In order to solve **L1**, we apply elements of the Reputation Management framework. For **L2** we propose changing the syntax and semantics of the PDEL formulas which describe public announcements. For **L3** we suggest changing the specification for the domain of the probability function. Finally, to resolve **L4** we propose using elements of the \mathcal{AG}_n logic.

The adaptation of PDEL to the persuasion model enriches the model expressivity with respect to change of probabilistic beliefs induced by persuasion. This provides a key first step towards creating \mathcal{PAG}_n , i.e., the Probabilistic Logic of Actions and Graded Beliefs, and extending the Perseus system designed to verify persuasive MAS. Moreover, in future work we are going to enrich the aspect of persuader's reputation e.g. by adding actions modifying trust. Such actions change neither accessibility relations nor values of propositions.

References

- [1] F. Bacchus, *Representing and Reasoning with Probabilistic Knowledge*, MIT Press, Cambridge, 1990.
- [2] K. Budzyńska and M. Kacprzak, 'A logic for reasoning about persuasion', *Fundamenta Informaticae*, **85**, 51–65, (2008).
- [3] K. Budzyńska, M. Kacprzak, and P. Rembelski, 'Modeling persuasiveness: change of uncertainty through agents' interactions', in *Frontiers in Artificial Intelligence and Applications*. IOS Press, (2008).
- [4] K. Budzyska, M. Kacprzak, and P. Rembelski, 'Perseus. software for analyzing persuasion process.', *Fundamenta Informaticae*, (91), (2009).
- [5] J. Y. Halpern, 'An analysis of first-order logics of probability', *Artificial Intelligence*, **46**, 311–350, (1990).
- [6] J.Y. Halpern, 'Lexicographic probability, conditional probability, and nonstandard probability', in *Proceedings of the Eighth Conference on Theoretical Aspects of Rationality and Knowledge*, 17–30, (2001).
- [7] A. Hussain and F. Toni, 'Bilateral agent negotiation with information-seeking', in *Proc. of EUMAS-2007*, (2007).
- [8] B. Kooi, 'Probabilistic dynamic epistemic logic', *Journal of Logic, Language and Information*, **12**, 381–408, (2003).
- [9] N. Laverny and J. Lang, 'From knowledge-based programs to graded belief-based programs part i: On-line reasoning', *Synthese*, (147), 277–321, (2005).

- [10] H. Prakken, 'Formal systems for persuasion dialogue', *The Knowledge Engineering Review*, **21**, 163–188, (2006).
- [11] S. D. Ramchurn, C. Mezzetti, A. Giovannucci, J. A. Rodriguez-Aguilar, R. K. Dash, and N.R. Jennings, 'Trust-based mechanisms for robust and efficient task allocation in the presence of execution uncertainty', *Journal of Artificial Intelligence Research*, (35), 119–159, (2009).
- [12] H.P. van Ditmarsch, 'Prolegomena to dynamic logic for belief revision', *Knowledge, Rationality & Action (Synthese)*, **147**, 229–275, (2005).
- [13] D. N. Walton and E. C. W. Krabbe, *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, State Univ. of N.Y. Press, 1995.
- [14] B. Yu and M.P. Singh, 'A social mechanism of reputation management in electronic communities', in *Proceedings of the Fourth International Workshop on Cooperative Information Agents*, (2000).

ZMIANA PROBABILISTYCZNYCH PRZEKONAŃ W PERSWAZJI

Streszczenie: Celem pracy jest rozszerzenie zaproponowanego przez nas formalnego modelu perswazji o aspekt zmiany niepewności przekonań agentów interpretowanych w teorii prawdopodobieństwa. Wzbogacony model jest podstawą do zdefiniowania logiki i zaprojektowania narzędzia, które umożliwi automatyczną weryfikację perswazyjnych systemów wieloagentowych. W celu realizacji tego zadania analizujemy i adaptujemy Probabilistyczną Dynamiczną Epistemiczną Logikę wprowadzoną przez B. Kooi. Zastosowanie zaproponowanego podejścia do analizowania wybranych aspektów perswazji omawiamy na przykładzie problemu alokacji zasobów w rozproszonych komputerowych systemach.

Słowa kluczowe: perswazja, przekonania, logika prawdopodobieństwa, formalna weryfikacja