ECAI 2024
U. Endriss et al. (Eds.)
© 2024 The Authors.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/FAIA240895

Compromises in Dialogical Argumentation: Aggregated Policies for Biparty Decision Theory

Ivan Donadello[a,*], Renan Lirio de Souza[b], Anthony Hunter[c] and Mauro Dragoni[b]

[a]Free University of Bozen-Bolzano
[b]Fondazione Bruno Kessler
[c]University College London
ORCID (Ivan Donadello): https://orcid.org/0000-0002-0701-5729, ORCID (Renan Lirio de Souza): https://orcid.org/0000-0001-5441-1858, ORCID (Anthony Hunter): https://orcid.org/0000-0001-5602-7446, ORCID (Mauro Dragoni): https://orcid.org/0000-0003-0380-6571

Abstract. Automated persuasion systems (APS) are conversational agents that exchange arguments and counterarguments with users during dialogues to persuade them to believe in something. Such systems use strategies (or policies) to carefully select a sequence of arguments that are tailored to the user's needs and will likely have a positive outcome, that is, changing the user's belief in a certain argument. Biparty Decision Theory (BDT) is a framework that uses game theory to formalize a dialogue between an APS and a user, that is, an exchange of (counter) arguments during each turn of the APS or the user. During the APS turn, the BDT policy selects the best argument to maximize only the utility for the APS and neglects the utility of that argument for the user. This is a reasonable choice in games, but in a persuasive dialogue, it can result in arguments that have a high utility for the APS but a modest utility for the user. There the user may be less likely to be persuaded. This is crucial in settings where there are no arguments with good utilities for both the APS and the user and a compromise has to be found. To this extent, we define a new family of policies for BDT, called *aggregated policies*, that consider, during the decisions of the APS, an aggregation of the APS and user's utilities. Such an aggregation considers both the APS and the user's needs leading toward a sequence of arguments representing the best trade-off of utilities. We evaluate the approach using both a new synthetic dataset and a published dataset of utilities for dialogical argumentation. The results show the aggregated policies find better compromise arguments w.r.t. the classical policy of BDT.

1 Introduction

Persuasion is an activity that involves one party trying to induce another party to believe something. It is an important and multifaceted human facility. It usually involves a dialogue in which arguments and counterarguments are exchanged between the persuader and the persuadee, and it is seen as increasingly important in domains such as healthcare [19, 30, 15, 14]. For example, a doctor can use arguments to persuade a patient to adopt a more healthy diet, and the patient can give counterarguments based on their understanding and preferences. Also, the doctor may provide counterarguments to attempt to overturn misconceptions held by the patient. An automated persuasion system (APS) is a software system that takes on the role of a persuader, and the user is the persuadee [28]. It aims to use convincing arguments in order to persuade the user towards achieving the goal of the system. Whether an argument is convincing depends on its own nature, on the context of the argumentation, and on the user's characteristics. An APS maintains a model for the user (the so-called *user model*) to choose the best arguments to pose in the dialogue [21]. For the APS turn, the APS presents an argument to the user, and then the user can answer with a counterargument selected from an input menu or with input in natural language form [28, 10].

As an example of an approach for dialogue argumentation, the bimaximax policy (presented in the context of Biparty Decision Theory (BDT) [23]) selects, during the persuader's turn, the argument that maximizes the utility for the persuader at the end of the dialogue by also considering the counterarguments that the user could pose. Indeed, a simple local choice (that ignores the user model) during the proponent's turn could choose arguments that will be overturned by the user's counterarguments. For example, in the healthcare domain, the APS suggestion of eating legumes instead of red meat has a higher utility for the APS than switching to white meat. However, the user model in the APS contains the information that the user will eat legumes just one meal per week, thus with a low impact on the diet. Instead, the user prefers eating white meat. Therefore, the APS will harness this information to propose eating white meat that has a high APS utility than eating legumes just once per week. However, the bimaximax policy presents an asymmetry as, during the persuader turn, it selects arguments that maximize only the final utility of the persuader neglecting the one of the persuadee. More generally, when the arguments are good for only one party, the bimaximax policy returns arguments that leave the other party unsatisfied. This can leave the opponent reluctant to accept those arguments. However, some of the available arguments can represent a good compromise between the proponent and opponent utilities, for example, the suggestion of eating white meat can have a higher utility for the opponent (with respect to the legumes argument) and be more likely to be accepted.

To address this need for compromise raised above, we propose a new family of policies for BDT, called *aggregated policies*, that consider, during the decisions of the proponent, an aggregation of both the proponent and opponent's utilities. Such an aggregation takes into account both the APS and the user's needs leading towards a

* Corresponding Author. Email: ivan.donadello@unibz.it

sequence of arguments representing the best compromise between the proponent and the opponent, that is, the best compromise/trade-off of the utilities. From this scenario, we formulate the following **research questions**, **RQ1**: "Are aggregated policies able to choose arguments for the APS to present that are the best compromise for an APS and a user in a dialogue?"; **RQ2**: "Could aggregated policies be more persuasive than non-aggregated policies?".

We will clarify what we mean by RQ1 in Section 4. The motivation for RQ2 is to consider how compromises in argumentation can contribute to more persuasive argumentation. Within the psychology literature, there are numerous studies that draw out connections between utility of a message, and its persuasiveness. In particular, within the domain of healthcare communications, there are numerous studies that show framing healthcare messages in terms of benefits to the individual or their family, are more persuasive [42, 18, 7, 31]. It is beyond the scope of this paper to undertake studies with subjects to determine whether our aggregated policies are more persuasive. So in order to investigate RQ2, we will investigate empirically how on average playing arguments that have higher utility (i.e. have greater benefits) for the opponent, are more acceptable to the opponent, and therefore more persuasive to the opponent.

The paper is structured as follows. Section 2 surveys the main strategies in the literature for dialogical argumentation and compromises in decision-making. Section 3 provides an overview of the Biparty Decision Theory, while Section 4 discusses the motivation behind our proposal. The new policies we developed are presented in Section 5 and empirically validated in Section 6. Section 7 concludes the paper.

2 Related Work

Within a persuasive dialogue, an APS must adopt a specific strategy to choose which argument to present to the user in order to move the dialogue on. The state-of-the-art contains several types of approaches targeting the selection of the best possible arguments to adopt. The contributions presented in [36, 37, 17] proposed a one-step process to manage dialogues (instead of a multi-step one) based on **Mechanism design**. Another family of techniques is based on **planning** that have been adopted with the aims of optimizing the choice of arguments based on belief in premises [5, 6] or minimizing the number of moves made [3]. Strategies based on **machine learning** have been designed as well, e.g. reinforcement learning [32, 41, 35, 29, 39, 2] and transfer learning [40]. Other methods are based on **probabilistic strategies** where moves are selected based on: (i) what an agent believes the other is aware of [38]; (ii) to approximately predict the argument an opponent might put forward based on data about the moves made by the opponent in previous dialogues [20]; (iii) using a decision- theoretic lottery [27]; and, (iv) using POMDPs when there is uncertainty about the internal state of the opponent [22]; Finally, both **Local strategies** and **Global strategies** have been proposed. The former is based on the concerns (i.e., issues raised or addressed by an argument) of one of the parties (e.g., the persuadee ones) [21, 10]. Instead, the latter is based on taking concerns and beliefs into account using, for example, Monte Carlo tree search [24].

Our motivation for aggregated policies is to allow for a compromise in the decision-making by the proponent. Our approach is to use games in extensive form [33], and adapt them for argumentation. The notion of compromise has long since been studied in game theory [4]. Diverse developments include a minimax procedure for compromise in negotiating treaties when a guarantee of non-manipulability is desirable [8]; investigation of the role of compromise in two-player games involving players with private information about their own strength [9]; and investigation of the role of compromise in the determining the existence of a cooperative dual to Nash equilibria for n-person games in strategic form [12]. Also aggregation of utilities has been considered for group decision-making [25]. However, compromises for extensive games have not been considered in the context of computational argumentation in multi-agent dialogues.

3 Background: Biparty Decision Theory

Decision trees (DTs) are one of the main formalisms in decision theory [34]. They have been harnessed to optimize the choice of dialogue move when trying to persuade an agent in believing something [26]. In this case, a DT represents all possible dialogues between two agents with the paths from the root to leaves that alternate *decision nodes* with *chance nodes*. The former are associated with the persuader/proponent and represent the arguments that will be posed by the APS. The latter are associated with the persuadee/opponent and represent the arguments posed by the user. The arc between two DT nodes n and m is labeled with the argument posed by the corresponding agent at node n. Figure 1 contains an example of DT for the persuasion goal of reducing red meat consumption, adapted from [16].

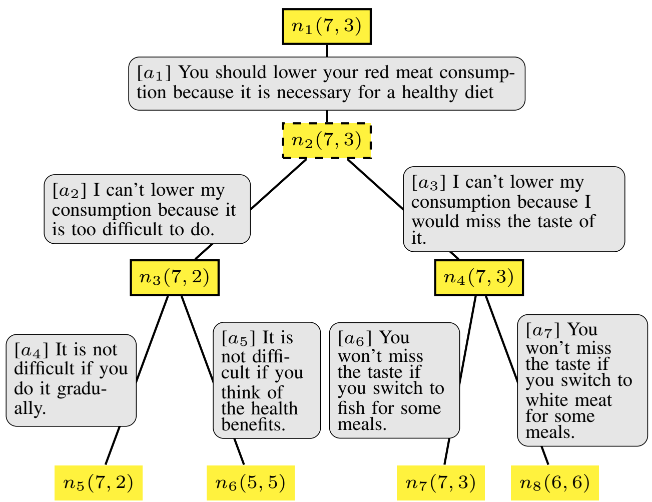

Figure 1: A decision tree for a persuasive dialogue about reducing red meat consumption. Solid boxes represent decision/proponent nodes whereas dashed boxes are for chance/opponent nodes.

A *policy* provides decision nodes with the best arguments to pose by the APS by considering the points of view of both the proponent and the opponent. Biparty Decision Theory (BDT) [23] models these viewpoints with the *utility functions* U^p and U^o. These associate each leaf with a utility value that represents the benefit the proponent p (respectively the opponent o) gets if the dialogue takes place according to the moves in the branch to that leaf. The utility is based on the change in behavior that is brought about by the arguments in the dialogue. In this situation, the utility of a leaf might be different for the proponent and the opponent. In our example, the APS might prefer that the user has a very low-cholesterol diet for health reasons, whereas the user might prefer a diet that has some meat even if it has more cholesterol. The **bimaximax policy** maximizes U^p at a decision node and U^o at a chance node. Let T be a DT, L be a labeling function that assigns an argument to each arc (n, m) between two nodes of T and $\mathsf{Children}(T, n)$ be the set of children of n. Let

$\mathsf{Nodes}(T)$ be the set of nodes in T. The $\mathsf{AMax}(T, U, n)$ function returns the children of n with highest utility U: $\mathsf{AMax}(T, U, n) = \{n' \in \mathsf{Children}(T, n) \mid \forall n'' \in \mathsf{Children}(T, n), U(n') \geq U(n'')\}$. Therefore:

Definition 1. *A* **bimaximax policy** *for* (T, L, U^p, U^o, δ) *is* $\Pi_{bim} : \mathsf{Nodes}(T) \rightarrow \mathsf{Nodes}(T)$ *defined as follows using the calculation of the* $Q^p : \mathsf{Nodes}(T) \rightarrow \mathbb{R}$ *and* $Q^o : \mathsf{Nodes}(T) \rightarrow \mathbb{R}$ *functions.*

- *If n is a leaf node, then $Q^p(n) = U^p(n)$ and $Q^o(n) = U^o(n)$.*
- *If n is a chance node, and $n_i \in \mathsf{AMax}(T, Q^o, n)$, then $Q^p(n) = \delta \times Q^p(n_i)$ and $Q^o(n) = \delta \times Q^o(n_i)$.*
- *If n is a decision node, and $n_i \in \mathsf{AMax}(T, Q^p, n)$, then $Q^p(n) = \delta \times Q^p(n_i)$, $Q^o(n) = \delta \times Q^o(n_i)$ and $\Pi_{bim}(n) = \langle n_i, L(n, n_i) \rangle$.*

Here, $\delta \in \mathbb{R}$ is a discount factor that decreases the utility of longer branches. This mechanism favors shorter dialogues instead of longer ones that require more attention from the user. Following the running example in Figure 1, the APS leaf utility values are $U^p(n_5) = 7$, $U^p(n_6) = 5$, $U^p(n_7) = 7$, $U^p(n_8) = 6$ whereas the opponent utilities are: $U^o(n_5) = 2$, $U^o(n_6) = 5$, $U^o(n_7) = 3$, $U^o(n_8) = 6$. The utility values are then propagated to the higher levels of the tree according to Definition 1. The policy Π_{bim} is therefore (with $\delta = 1$) $\Pi_{bim}(n_1) = \langle n_2, L(n_1, n_2) \rangle$, $\Pi_{bim}(n_2) = \langle n_4, L(n_2, n_4) \rangle$, $\Pi_{bim}(n_3) = \langle n_5, L(n_3, n_5) \rangle$ and $\Pi_{bim}(n_4) = \langle n_7, L(n_4, n_7) \rangle$.

A policy defines a rule for choosing the label/argument to pose in a decision node. Given a chance node instead, the next decision node is selected by the user via, for example, a menu. When real users are not available, which occurs when the proponent is using the DT to optimize its choices of move prior to a dialogue, the next decision node has to be computed according to some *simulated opponent policy* Π^o. Here, we adopt the solution in [16] where $\Pi^o(n_i) = \langle n_j, L(n_i, n_j) \rangle$ with $n_j \in \mathsf{AMax}(T, Q^o, n_i)$. A *simulated dialogue procedure* $\mathrm{SimDialogue}(\Pi, \Pi^o, T, L, \boldsymbol{u}^p, \boldsymbol{u}^o, \delta)$, is an algorithm that takes as input a policy Π (e.g., bimaximax), a simulated policy Π^o, a DT T, a labelling function L, the vectors $\boldsymbol{u}^p = \langle U^p(n_1), U^p(n_2), \ldots \rangle$ and $\boldsymbol{u}^o = \langle U^o(n_1), U^o(n_2), \ldots \rangle$ containing the proponent and opponent utility values, respectively, at the leaves and the discount factor δ and computes the best path $\boldsymbol{p}$ of nodes according to Π. In our example, $\boldsymbol{p} = \langle n_1, n_2, n_4, n_7 \rangle$. The values of $\boldsymbol{u}^p$ can be defined by domain experts from healthcare literature, e.g. [1]. Instead, the setting of $\boldsymbol{u}^o$ is an open challenge. In [16], for example, $\boldsymbol{u}^o$ is predicted using machine learning methods that leverage a training set of users' utilities.

4 Motivation

During the APS turn, the bimaximax policy selects the arguments to pose that maximize only the final utility for the APS by also considering the counter-arguments from the user. We can define this feature as a *selfishness property* for the APS as bimaximax neglects the opponent's utility during the APS turn. This could lead to arguments with low utility for the proponent and, therefore, hard to be accepted. This is shown in Figure 1, where argument a_6 is selected as it maximizes the APS final utility, i.e., 7. However, this argument has a low utility for the user (just 3, the user could not like fish) and it could be discarded with no impact on the diet.

We now clarify our intuition behind **RQ1**. As seen in the previous section, with the bimaximax method, each argument is paired with a pair of utilities, so we define the *argument space*, i.e., the Euclidean space having on the X-axis the utility values for the proponent and on the Y-axis the utility values for the opponent. This space can be partitioned into *regions of interest* according to the utility values as shown in Figure 2 (the utilities range from 1 to 10 just for the sake of simplicity):

The optimal region contains arguments (in green) that have high utilities for both the proponent and the opponent.

The unselfish region contains arguments (in orange) reflecting the priorities of the opponent, i.e., with high utility for the user and low for the APS. These arguments may be less effective with respect to the goal of an APS.

The selfish region contains arguments (in red) reflecting the priorities of the proponent, i.e., with high utility for the APS and low for the user. These arguments have the risk of not being accepted by the persuadee. In our example, the arguments a_4 and a_6 fall into this region.

The compromise region contains arguments (in blue) representing a compromise/trade-off from the dialogue perspective since their utilities are not maximal but are equally good for both the proponent and the opponent.

The best compromise region is a subset of the frontier of the compromise region with arguments (in black) that have the highest utilities for both the proponent and the opponent. The arguments represent the best compromise/trade-off between the proponent and the opponent. The argument a_5 is in the compromise region, whereas a_7 is in the best compromise region.

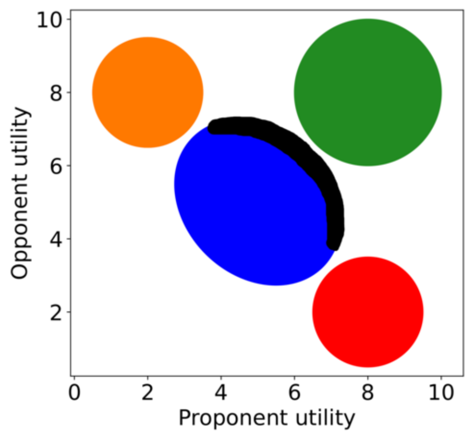

Figure 2: The argument space represents how arguments fit into the Euclidean space according to their utility values. Regions of interest indicate the quality of persuasiveness for an argument.

The optimal and the compromise regions contain arguments that are balanced, i.e., their utilities are close to each other, whereas the selfish and unselfish regions contain unbalanced arguments as their utilities are distant from each other.

We motivate our proposal for aggregated policies when the harvested arguments (along with their utilities) do not fall into the optimal region for a user. This is a reasonable assumption as, in some cases, the user and the APS can have opposite utilities for the available arguments. In such a setting, Π_{bim} returns arguments in the selfish region as shown in the running example. The aim of the **RQs** is to show whether i) the aggregated policies return, instead, more balanced arguments that fall in the best compromise region; ii) these arguments are more persuasive than arguments in other regions.

5 Method: Aggregated Policies

A possible solution to the *selfishness* of the bimaximax policy is the introduction of a new term that considers both the proponent and opponent utilities in the policy computation at decision nodes. This term can only be added in the decision nodes as the chance nodes totally depend on other parties (e.g., the user) that cannot be controlled. We define this term as an *aggregation function* $\mathcal{A} : \mathbb{R}^{|\mathcal{R}|} \to \mathbb{R}$ that takes a vector of utilities (one for each role in $\mathcal{R}$) and returns an aggregation value in $\mathbb{R}$. In BDT, $\mathcal{R} = \{p, o\}$, but, in principle, $\mathcal{R}$ can contain many roles thus allowing multiparty decision theory. We formally define the *aggregated policy* for BDT as:

Definition 2. *An* **aggregated policy** *for* $(T, L, U^p, U^o, \mathcal{A}, \delta)$ *is* $\Pi_{\mathcal{A}} : \mathsf{Nodes}(T) \to \mathsf{Nodes}(T)$ *defined as follows using the calculation of the* $Q^p : \mathsf{Nodes}(T) \to \mathbb{R}$, $Q^o : \mathsf{Nodes}(T) \to \mathbb{R}$ *and* $Q^a : \mathsf{Nodes}(T) \to \mathbb{R}$ *functions.*

- *If* n *is a leaf node, then* $Q^p(n) = U^p(n)$, $Q^o(n) = U^o(n)$ *and* $Q^a(n) = \mathcal{A}(U^p, U^o)$.
- *If* n *is a chance node, and* $n_i \in \mathsf{AMax}(T, Q^o, n)$, *then* $Q^p(n) = \delta \times Q^p(n_i)$, $Q^o(n) = \delta \times Q^o(n_i)$ *and* $Q^a(n) = \delta \times Q^a(n_i)$.
- *If* n *is a decision node, and* $n_i \in \mathsf{AMax}(T, Q^a, n)$, *then* $Q^p(n) = \delta \times Q^p(n_i)$, $Q^o(n) = \delta \times Q^o(n_i)$, $Q^a(n) = \delta \times \mathcal{A}(Q^p(n_i), Q^o(n_i))$ *and* $\Pi_{\mathcal{A}}(n) = \langle n_i, L(n, n_i)\rangle$.

Differently from the standard bimaximax policy Π_{bim} in Definition 1, the aggregated policy $\Pi_{\mathcal{A}}$ selects the node n_i with the highest aggregated value Q^a according to the aggregate function $\mathcal{A}$. In this manner, nodes with the highest overall utility value Q^a would be preferred instead of unbalanced nodes, i.e., nodes with a high utility only for one party. This would avoid the selfishness of the bimaximax policy that maximizes, in turn, the utility of a single party. We, therefore, revise the running example of before. In Figure 3, we consider the arithmetic mean for the aggregated function $\mathcal{A}$, that is, $Q^a = (Q^p + Q^o)/2$. Here, the nodes of the DT contain the values $\langle Q^p, Q^o, Q^a\rangle$. Differently from Figure 1, in the first layer of decision nodes (the one with n_3 and n_4), the policy $\Pi_{\mathcal{A}}$ selects the nodes n_6 and n_8 as they have a better-aggregated value than n_5 and n_7 ($Q^a = 5$ vs $Q^a = 4.5$ and $Q^a = 6$ vs $Q^a = 5$), respectively. The node n_2, instead, is a chance node and, therefore, propagates up the utility of node n_4 as it has the highest U^o value. As a consequence, in a possible dialogue between a real user and the APS, the node n_8 will be returned and it will be more likely accepted by the user with respect to n_7 (returned by Π_{bim} instead). From the APS perspective, n_8 has a lower utility than n_7 ($U^p = 6$ vs $U^p = 7$) but this is counterbalanced by a higher opponent utility ($U^o = 6$ vs $U^o = 3$) that it will make more suitable for the user. The policy $\Pi_{\mathcal{A}}$ is therefore (with $\delta = 1$) $\Pi_{\mathcal{A}}(n_1) = \langle n_2, L(n_1, n_2)\rangle$, $\Pi_{\mathcal{A}}(n_2) = \langle n_4, L(n_2, n_4)\rangle$, $\Pi_{\mathcal{A}}(n_3) = \langle n_6, L(n_3, n_6)\rangle$ and $\Pi_{\mathcal{A}}(n_4) = \langle n_8, L(n_4, n_8)\rangle$.

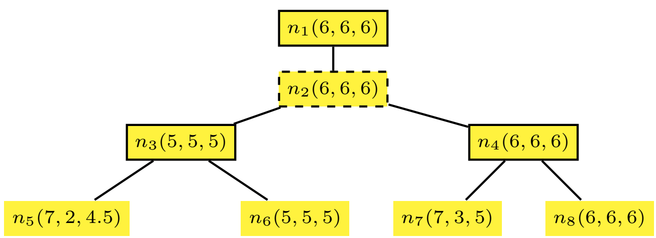

Figure 3: The running example with the aggregated policy and the arithmetic mean as $\mathcal{A}$. The vector in the nodes contains the values $\langle U^p, U^o, U^a\rangle$.

According to Definition 2, an aggregated policy requires a suitable aggregation function $\mathcal{A}$.[1] We propose and test three solutions. The first one is the **generalized mean** M_k:

$$\mathcal{A}_{M_k}(U^p, U^o) = \left(\frac{1}{|\mathcal{R}|}\sum_{i\in\mathcal{R}} U_i^k\right)^{\frac{1}{k}} \tag{1}$$

with $\mathcal{R} = \{p, o\}$ and k is a positive real number and U_i^k denoting the power of k of the utility value $U_{p/o}$. By changing k, different means can be obtained, such as $k = -1$ gives the harmonic mean, $k = 0$ the geometric mean, and $k = 1$ the arithmetic mean. Here, the set $\mathcal{R} = \{p, o\}$ contains only the roles of the proponent and the opponent but it can contain multiple roles making this aggregated policy valid for multiparty decision theory. The value for k has to be properly chosen, indeed, in our example of Figure 3 (where $k = 1$), the nodes n_6 and n_7 have the same value for Q^a even if they have different values for U^p and U^o and n_6 has balanced utilities whereas n_7 has unbalanced values. A possible solution is changing the value for k or the introduction of an aggregation function that penalizes nodes with unbalanced utility values. We call this second type of aggregation function **sum minus difference** (SMD_λ) of the U^p and U^o utilities and it is defined as:

$$\mathcal{A}_{SMD_\lambda}(U^p, U^o) = (1-\lambda)\cdot(U^p + U^o) - \lambda\cdot|U^p - U^o| \tag{2}$$

with $\lambda \in [0, 1]$ is a coefficient that weights the importance of having a high/low sum of the utilities with respect to a low/high difference of the utilities. In our example, the node n_6 will have $\mathcal{A}_{SMD_{0.5}}(5, 5) = 5$ that is higher of the one of n_7, that is, $\mathcal{A}_{SMD_{0.5}}(7, 3) = 3$. However, this policy cannot be generalized to multiparty decision theory. We, therefore, define a new policy as the mean of the U^p and U^o utilities over their standard deviation and call it **Mean over the Standard Deviation** ($MoSTD_\alpha$):

$$\mathcal{A}_{MoSTD_\alpha}(U^p, U^o) = \frac{\mu}{(\sigma + \alpha)} \tag{3}$$

where $\mu = \sum_{i\in\mathcal{R}} U^i/|\mathcal{R}|$ is the mean of the utilities, $\sigma = \sqrt{\sum_{i\in\mathcal{R}}(U^i - \mu)^2/|\mathcal{R}|}$ is the standard deviation and $\alpha > 0$ is a parameter for avoiding a zero denominator when all utilities have the same value. The set $\mathcal{R}$ can contain multiple roles and, therefore, the $MoSTD_\alpha$ policy can be used in a multiparty setting.

Table 1 compares the aggregation values Q^a of the leaves computed by the defined aggregated policy with the Q^p values considered by the bimaximax policy for defining Π in the running example. We can see that the most balanced (and with higher utilities) node, n_8, if selected if an aggregated policy is used. Bimaximax, on the contrary, will select the unbalanced nodes n_5 and n_7.

Table 1: The aggregated values Q^a of the leaves in the example compared with the Q^p values of the bimaximax policy.

Nodes \ Policy	M_{-1}	M_0	M_1	$SMD_{0.5}$	$MoSTD_1$	bim
$n_5(7,2)$	3.1	3.7	4.5	2.0	1.0	**7**
$n_6(5,5)$	5.0	5.0	5.0	5.0	5.0	5
$n_7(7,3)$	4.2	4.6	5.0	3.0	1.3	**7**
$n_8(6,6)$	**6.0**	**6.0**	**6.0**	**6.0**	**6.0**	6

[1] The function $\mathcal{A}$ is defined over both the U^p/U^o and the Q^p/Q^o, see first and third bullet of Definition 2. Hereafter, our proposals for $\mathcal{A}$ will use U^p/U^o just for simplicity of notation.

6 Empirical Evaluation

The goal of the evaluation is twofold: i) to test the ability of $\Pi_{\mathcal{A}}$ to provide the arguments in the best compromise region compared to Π_{bim} (**RQ1**); ii) to test whether $\Pi_{\mathcal{A}}$ results in more persuasive dialogues than Π_{bim}. That is, whether the arguments returned by $\Pi_{\mathcal{A}}$ are more likely to be accepted by users suggesting that $\Pi_{\mathcal{A}}$ is more effective in influencing their decisions (**RQ2**). However, since datasets of APS dialogues with real users whose arguments (with associated utilities) are not readily available, we evaluated the different policies on synthetically generated datasets. We tested on: i) an ad-hoc dataset whose arguments fit the (un) selfish and (best) compromise regions of Figure 2; ii) a recently published dataset [16]. Using synthetic data is an important intermediate step before attempting to undertake an empirical study with participants using an argument-based chatbot (such as undertaken in [21, 10, 11, 24]). It allows us to explore the space of possible combinations of proponent and opponent utilities, as well as other parameters, in order to get a better understanding of the behaviour of our aggregation methods. Such a preliminary understanding of the theory with synthetic data may reduce potential risks for participants in experiments. The source code, datasets, and appendices are online at https://github.com/ivanDonadello/Aggregated-Policies-Biparty-Decision-Theory.

6.1 Datasets

We compared $\Pi_{\mathcal{A}}$ with Π_{bim} on different abstract DTs equipped with the utilities $\boldsymbol{u}^p$ and $\boldsymbol{u}^o$. The set $\mathbb{T} = \{T_1 \dots T_{10}\}$ contains 10 abstract DTs taken from the dataset in [16]. Table 3 shows some figures, further details are in [16]. Each tree $T \in \mathbb{T}$ is paired with a set of utility vectors $\mathbb{U}_T = \{\langle \boldsymbol{u}^p, \boldsymbol{u}^o \rangle\}_{i=1}^N$ whose values are associated to the leaves in the tree T. The cardinality N of $\mathbb{U}_T$ represents the number of users in the dataset. Each user has its own utility values $\boldsymbol{u}^o$ for T whereas the utilities in $\boldsymbol{u}^p$ can be shared among users in the same subpopulation as they derive from domain knowledge. For the evaluation, we i) synthetically generate $\mathbb{U}_T$; ii) use the provided values in [16].

In the first case, we synthetically generate $\mathbb{U}_T$ to cover the selfish, unselfish, and (best) compromise regions of the argument space by implementing this procedure:

- we defined a rectangular area $\mathcal{D}$ of uniform probability distribution that ranges from $U^p_{min} = 5$ to $U^p_{max} = 7$ and from $U^o_{min} = 0$ to $U^o_{max} = 12$.
- We rotated $\mathcal{D}$ of $\pi/4$ counterclockwise and shift it to avoid negative values.
- Given $T \in \mathbb{T}$, we created vectors $\langle \boldsymbol{u}^p, \boldsymbol{u}^o \rangle$ by sampling from $\mathcal{D}$ a number of points equal to the number of leaves in T. The values are converted into integers. We repeated this $N = 100$ times.

We called this dataset **don2022NoOPT** and used it to compare the aggregated policies with bimaximax where arguments in the optimal region are lacking. We empirically investigate **RQ1** by inspecting whether Π_{bim} returns arguments that fall into the selfish region whereas $\Pi_{\mathcal{A}}$ is able to cover the best compromises region of arguments. In addition, we also use the set $\mathbb{U}_T$ provided in [16] (whose utility values range from 1 to 10) to study the behavior of the aggregated policies on a normal setting of arguments, that is, arguments that cover also the optimal region. We refer to this dataset as **Don2022** that is is more general (and challenging) than the previous one as it presents the optimal region and it is good to test whether $\Pi_{\mathcal{A}}$ returns arguments with higher utilities than the ones returned by Π_{bim} even in presence of optimal arguments.

6.2 Simulation Design and Settings

To compare $\Pi_{\mathcal{A}}$ with Π_{bim} on the datasets, we run the $\mathrm{SimDialogue}(\Pi, \Pi^o, T, L, \boldsymbol{u}^p, \boldsymbol{u}^o, \delta)$ procedure for each $\langle \boldsymbol{u}^p, \boldsymbol{u}^o \rangle \in \mathbb{U}_T$, for each $T \in \mathbb{T}$ and for each policy $\Pi \in \{\Pi_{bim}\} \cup \{\Pi_{\mathcal{A}}\}$ where $\mathcal{A} \in \{M_k, SMD_\lambda, MoSTD_\alpha\}$ and $\delta = 1$. As the DTs in $\mathbb{T}$ are abstract with no particular meaning, the labeling function L is not defined. For $\mathcal{A} = M_k$ we set $k \in \{-2, -1, 0, 1, 2\}$ whereas the weight λ ranges from 0 to 1 with a step of 0.1.[2] For $MoSTD_\alpha$ we set $\alpha = 1$. Each run of the SimDialogue returns a sequence $\boldsymbol{p}$ of n visited nodes in T where the last one $p_n = \langle U^p, U^o, U^a \rangle$ contains the last argument returned. Such last nodes are then collected in the set $\mathcal{P} = \{\langle U^p, U^o, U^a \rangle\}_{i=1}^N$ and evaluated.

We empirically investigate **RQ2** by testing the persuasiveness of the policies. We define for each user i an *acceptability threshold* th^a_i that indicates the lowest level of utility for a user to accept any argument. For example, a demanding user that would reject many arguments will have a high acceptability threshold as they would only accept arguments having a utility above this threshold. To perform this evaluation during our simulations, we define three clusters of users having low, medium or high th^a. Given a user i, their acceptability threshold is sampled from a normal distribution with $\mu \in \{3, 5, 7\}$ and $\sigma = 1$. Therefore, users in cluster with $\mu = 3$ will show the most receptive behavior with high argument acceptability. Users in cluster with $\mu = 5$ will be less selective, users in cluster with $\mu = 7$ will have a pickier behavior accepting few arguments.

6.3 Metrics

The aim of the evaluation is to compare the behavior of all policies and to test the ability of the aggregation policies to i) provide better compromise arguments with respect to the bimaximax policy; ii) select the arguments with the highest utilities for both the proponent and the opponent among all the compromise arguments; iii) provide arguments that are more likely to be accepted by users. In the mentioned example, both $n_6 = \langle 5, 5, 5 \rangle$ and $n_8 = \langle 6, 6, 6 \rangle$ are compromise arguments, but n_8 has to be chosen as it has higher utilities. The first point is evaluated with the metric **Average of Absolute Distance (AAD)** of the utilities:

$$AAD = \sum_{i \in \mathcal{P}} |U^p_i - U^o_i| / |\mathcal{P}|$$

AAD is an aggregated measure of how distant the utility values of a returned node are from each other for a given policy. Such a distance shows some characteristics and behaviors of different policies: a low AAD suggests that the policy tends to select nodes whose arguments have utility values that are relatively balanced and stay in the compromise region. In contrast, a policy with higher AAD returns unbalanced arguments. The second point is evaluated with the metrics **Average of the Proponent Utility (APU)** and **Average of the Opponent Utility (AOU)**:

$$APU = \sum_{i \in \mathcal{P}} U^p_i / |\mathcal{P}| \qquad AOU = \sum_{i \in \mathcal{P}} U^o_i / |\mathcal{P}|$$

These aggregated values suggest whether a policy returns arguments in the best compromise (high values for both APU and AOU) area with respect to arguments in the compromise area. The third point is

[2] We show the results only for a subset of these values. Full results are in Appendix A.

Table 2: Evaluation results for **don2022NoOPT**. AAD on the left, APU in the center, AOU on the right. The $\Pi_{\mathcal{A}}$ policies return arguments in the best compromise region with respect to Π_{bim}, see the lower AAD and balanced values for APU and AOU.

Policy \ Height	4	5	6
bim	4.085	2.93	2.85
M_{-2}	1.662	2.393	2.223
M_{-1}	1.948	2.783	2.487
M_0	2.425	3.347	3.063
M_1	4.155	5.51	5.433
M_2	6.348	7.037	7.247
$SMD_{0.1}$	3.158	4.317	4.14
$SMD_{0.5}$	1.510	1.933	1.753
$SMD_{0.9}$	1.248	**1.577**	1.327
$MoSTD_1$	**1.232**	1.627	**1.290**

Policy \ Height	4	5	6
bim	7.940	7.317	7.490
M_{-2}	5.635	5.117	5.257
M_{-1}	5.558	4.990	5.190
M_0	5.385	4.710	4.917
M_1	4.888	3.660	3.770
M_2	3.995	2.807	2.707
$SMD_{0.1}$	5.292	4.293	4.463
$SMD_{0.5}$	5.605	5.263	5.420
$SMD_{0.9}$	5.555	5.367	5.483
$MoSTD_1$	5.590	5.337	5.533

Policy \ Height	4	5	6
bim	4.075	4.907	4.747
M_{-2}	6.782	7.437	7.413
M_{-1}	6.945	7.680	7.623
M_0	7.240	8.023	7.933
M_1	7.822	9.103	9.170
M_2	8.432	9.677	9.953
$SMD_{0.1}$	7.455	8.517	8.490
$SMD_{0.5}$	6.575	7.103	7.060
$SMD_{0.9}$	6.352	6.783	6.723
$MoSTD_1$	6.358	6.823	6.697

Table 3: Figures for the decision trees in $\mathbb{T}$.

Tree id	0	1	2	3	4	5	6	7	8	9
Height	4	4	4	6	6	4	5	5	5	6
# nodes	46	19	24	171	168	35	99	60	52	105
# leaves	30	11	15	108	102	22	60	37	39	62

evaluated with the metric **Acceptance Rate (AR)**, which measures the persuasive power of a policy over the set of users $\mathcal{P}$:

$$AR = \frac{1}{|\mathcal{P}|} \sum_{i \in \mathcal{P}} \mathbb{I}[Q_i^o \geq th_i^a]$$

where $\mathbb{I}[\cdot]$ is the indicator function taking the value 1 when its argument is true and 0 when false. **AR** values range from 0 and 1 (the higher the better) and quantify the effectiveness of a policy: values close to 0 indicate a policy with limited persuasive ability, AR values close to 1 indicate a higher probability of user acceptance of an argument, indicating a more effective policy.

6.4 Results

Table 2 shows the results on the **don2022NoOPT** dataset according to the height of the tree. Such an analysis gives us insights on how results change according to the length of the dialogue between the APS and the user. The policies with the lowest AAD are Π_{MoSTD_1} and Π_{SMD_λ} with high values for λ. This is in line with the design of this policy as, for high values of λ, the arguments with utilities that are closer to each other will be preferred to arguments with unbalanced utilities, see Equation (2). In general, most of the aggregated policies have lower AAD values than Π_{bim} meaning that they are better able to return arguments in the compromise region than bimaximax. On the other hand, some aggregated policies have worse results for AAD than bimaximax. Indeed, as long as k in Π_{M_k} increases, it gives more importance to the outliers of the utilities in the node. The same holds for Π_{SMD_λ} with low values for λ. The Π_{bim} policy has a decreasing trend for AAD as long as the height of the tree increases. This is a positive property as longer dialogues can be less persuasive as they require bigger attention from the user. The $\Pi_{\mathcal{A}}$ policies tend to have an increasing trend of AAD but for $MoSTD_1$ and SMD_λ, with high values of λ, this is not so significant and can be considered as a stable trend. These results allow us to state that Π_{bim} returns arguments in the selfish region whereas the arguments returned by $\Pi_{\mathcal{A}}$ are spread more in the compromise area. However, we need to analyze the AAD in conjunction with APU and AOU metrics to understand whether $\Pi_{\mathcal{A}}$ finds arguments in the best compromise region. The APU and AOU values for Π_{bim} are quite distant with higher values for APU. This means that Π_{bim} returns arguments in the selfish area that have a high utility only for the APS, therefore, they will be more difficult to be accepted by the user. The $\Pi_{\mathcal{A}}$ policies, instead, have closer values for APU and AOU with higher values for the opponent. This is a positive property as such policies support more the user instead of giving arguments with high utility only for the APS. In general, the policies with low AAD, that is, $SMD_{0.9}$ and $MoSTD_1$, have values for APU and AOU that place the returned arguments in the best compromise region.

To further assess the relative performance of the policies, we applied the Nemenyi test (adapted from [13]) as a means for statistical comparison of the policies over the results in $\mathcal{P}$. We compared the opposite of the absolute distances (AD) $-|U_i^p - U_i^o|$ obtained by each policy for a given sample $i \in \mathcal{P}$. We considered the opposite of the AD as this test performs a ranking of the policies based on a metric score where the policy with the highest score will have the highest position in such a ranking. The resulting critical difference diagram (Figure 4), obtained using a 0.05 significance level, confirms that

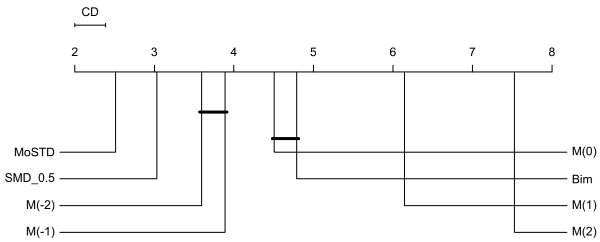

Figure 4: Comparison of the policies with the Nemenyi test according to the best AD. Groups of policies that are not significantly different (p-value < 0.05) are connected.

most of the aggregated policies, such as $MoSTD_1$, SMD_λ (with $\lambda \geq 0.5$), M_{-2} and M_{-1} are on average the ones that return arguments from the compromise area with respect to the bimaximax policy. The above analysis on the APU and AOU metrics can be statistically confirmed by the Nemenyi test, see Appendix B.

Table 4 shows the results for the **Don2022** dataset. All the aggregated policies have a lower value for AAD than Π_{bim} and, differently from **don2022NoOPT**, as the height of the DT increases, the AAD decreases. This suggests the ability of $\Pi_{\mathcal{A}}$ of returning arguments with balanced utilities as the dialogue increases. Such arguments are spread in the optimal region as we can see in the results for APU and AOU. On the contrary, Π_{bim} returns arguments in the border between the optimal region and the selfish region as the APU is around 10 and the AOU ranges from around 7 to around 8.

These results support a positive (and statistically robust) answer to **RQ1** with/without presence of arguments in the optimal region.

Table 5 shows the evaluation of the persuasive power for the **Don2022** dataset according to the users' clusters and the heights of the trees. Most aggregated policies have a higher AR value compared

Table 4: Evaluation results for **Don2022**. AAD on the left, APU in the center, AOU on the right. The $\Pi_{\mathcal{A}}$ policies return arguments in the best compromise region with respect to Π_{bim} even in presence of arguments in the optimal region. Wee the lower AAD and balanced values for APU and AOU.

Policy \ Height	4	5	6
bim	3.543	2.012	2.274
M_{-2}	1.947	1.524	1.415
M_{-1}	2.061	1.545	1.411
M_0	2.123	1.574	1.434
M_1	2.318	1.694	1.545
M_2	2.759	1.86	1.644
$SMD_{0.1}$	2.166	1.574	1.445
$SMD_{0.5}$	1.657	1.350	1.165
$SMD_{0.9}$	**0.799**	**0.677**	**0.433**
$MoSTD_1$	1.025	0.843	0.509

Policy \ Height	4	5	6
bim	10.171	9.861	10.606
M_{-2}	9.403	9.301	10.038
M_{-1}	9.452	9.331	10.069
M_0	9.408	9.308	10.086
M_1	9.298	9.288	10.114
M_2	8.963	9.219	10.182
$SMD_{0.1}$	9.342	9.304	10.073
$SMD_{0.5}$	9.341	9.178	9.894
$SMD_{0.9}$	7.573	8.599	8.872
$MoSTD_1$	8.450	8.630	8.963

Policy \ Height	4	5	6
bim	7.111	8.235	8.362
M_{-2}	8.336	8.781	9.059
M_{-1}	8.289	8.778	9.054
M_0	8.271	8.782	9.052
M_1	8.206	8.753	9.021
M_2	8.194	8.717	8.994
$SMD_{0.1}$	8.246	8.786	9.045
$SMD_{0.5}$	8.349	8.775	9.072
$SMD_{0.9}$	7.605	8.434	8.713
$MoSTD_1$	8.086	8.534	8.770

Table 5: Acceptance rates for **Don2022**: $\mu = 3$ left, $\mu = 5$ in the center, $\mu = 7$ on the right. The aggregated policies have higher persuasive power than Π_{bim}.

Policy \ Height	4	5	6
bim	0.996	1.0	1.0
M_{-2}	0.999	0.999	1.0
M_{-1}	0.999	1.0	1.0
M_0	0.999	0.999	1.0
M_1	0.998	0.999	1.0
M_2	0.999	0.999	1.0
$SMD_{0.1}$	0.998	1.0	1.0
$SMD_{0.5}$	0.999	1.0	1.0
$SMD_{0.9}$	0.993	1.0	1.0
$MoSTD_1$	0.998	0.999	1.0

Policy \ Height	4	5	6
bim	0.888	0.982	0.986
M_{-2}	0.963	0.991	0.997
M_{-1}	0.954	0.99	0.997
M_0	0.946	0.991	0.996
M_1	0.941	0.99	0.994
M_2	0.936	0.99	0.995
$SMD_{0.1}$	0.941	0.992	0.996
$SMD_{0.5}$	0.963	0.993	0.999
$SMD_{0.9}$	0.913	0.99	0.996
$MoSTD_1$	0.933	0.989	0.993

Policy \ Height	4	5	6
bim	0.524	0.848	0.76
M_{-2}	0.694	0.898	0.912
M_{-1}	0.678	0.894	0.902
M_0	0.676	0.89	0.9
M_1	0.657	0.884	0.888
M_2	0.669	0.886	0.872
$SMD_{0.1}$	0.672	0.888	0.897
$SMD_{0.5}$	0.698	0.906	0.924
$SMD_{0.9}$	0.589	0.863	0.861
$MoSTD_1$	0.628	0.873	0.867

Table 6: Acceptance rates for **Don2022NoOPT**: $\mu = 3$ left, $\mu = 5$ in the center, $\mu = 7$ on the right. The aggregated policies have higher persuasive power than Π_{bim}.

Policy \ Height	4	5	6
bim	0.703	0.857	0.907
M_{-2}	0.984	0.989	1.0
M_{-1}	0.992	0.989	1.0
M_0	1.0	1.0	1.0
M_1	0.984	1.0	1.0
M_2	0.875	1.0	0.99
$SMD_{0.1}$	0.992	1.0	1.0
$SMD_{0.5}$	0.992	0.989	1.0
$SMD_{0.9}$	0.984	0.989	1.0
$MoSTD_1$	0.992	0.989	1.0

Policy \ Height	4	5	6
bim	0.275	0.443	0.4
M_{-2}	0.906	0.938	0.968
M_{-1}	0.891	0.948	0.979
M_0	0.891	0.979	1.0
M_1	0.891	0.979	0.979
M_2	0.841	0.979	1.0
$SMD_{0.1}$	0.899	0.959	1.0
$SMD_{0.5}$	0.877	0.918	0.947
$SMD_{0.9}$	0.862	0.897	0.937
$MoSTD_1$	0.87	0.918	0.947

Policy \ Height	4	5	6
bim	0.007	0.054	0.083
M_{-2}	0.373	0.58	0.611
M_{-1}	0.455	0.625	0.722
M_0	0.537	0.741	0.806
M_1	0.709	0.875	0.954
M_2	0.716	0.973	1.0
$SMD_{0.1}$	0.575	0.821	0.843
$SMD_{0.5}$	0.276	0.509	0.546
$SMD_{0.9}$	0.246	0.366	0.407
$MoSTD_1$	0.224	0.393	0.454

to the bimaximax policy. For users in clusters with $\mu = 3$ and $\mu = 5$, all policies show high AR values, which is expected as these clusters include the most receptive users, that are more willing to accept arguments. On the other hand, users from cluster with $\mu = 7$ have lower AR values, due to their "picky" nature. Comparing results of this cluster with the AOU value from Table 4, we notice that the AR values are strictly linked to the AOU values; higher values of AOU mean a higher chance for the argument to be accepted by the user. Table 6 presents the results for **Don2022NoOPT** dataset. All the aggregated policies have a higher AR value compared to Π_{bim}, while the main difference, compared with Don2022 dataset, is that the Π_{bim} begins with significantly lower AR values for users in cluster with $\mu = 3$ compared with near-perfect scores achieved by aggregated policies. The AR also increases with the height of the DT, following the same trend as the AOU (see Tables 2 and 4). These results show higher persuasiveness of the aggregated policies giving evidence that compromises can potentially be more persuasive.

7 Conclusions

Biparty Decision Theory is a promising approach in dialogical argumentation. However, for the bimaximax policy, the maximization of the APS utility, irrespective of the user's utility, can be ineffective since it will make the APS choose arguments with a low utility for the user that they may be reluctant to accept. To address this, we have developed a new family of policies, the aggregated policies, that maximizes the utility for both the APS and the user during the APS turn. When tested on a new synthetic and on a published dataset, the aggregated policies return better arguments for both the APS and the user compared with the bimaximax policy. Such arguments have good utility values for both the APS and the user.

In future work, we will extend the theoretical basis of the new policies, especially to handle user utilities changing over time. Also, we will test the policies with real users and compare them with concern-based methods for dialogical argumentation [21]. We will also study aggregated policies in a multiparty setting where the APS has to persuade multiple opponents by trying to present arguments that best satisfies the majority without leaving some unsatisfied.

References

[1] I. Abete, D. Romaguera, A. R. Vieira, A. L. de Munain, and T. Norat. Association between total, processed, red and white meat consumption and all-cause, cvd and ihd mortality: a meta-analysis of cohort studies. *British Journal of Nutrition*, 112(5):762–775, 2014.

[2] S. Alahmari. *Reinforcement Learning for Argumentation*. PhD thesis, University of York, York, UK, 2020.

[3] K. Atkinson, P. Bench-Capon, and T. J. M. Bench-Capon. Efficiency in persuasion dialogues. In *ICAART (2)*, pages 23–32. SciTePress, 2012.

[4] R. Axelrod and W. D. Hamilton. The evolution of cooperation. *science*, 211(4489):1390–1396, 1981.

[5] E. Black, A. J. Coles, and S. Bernardini. Automated planning of simple persuasion dialogues. In *CLIMA*, volume 8624 of *LNCS*, pages 87–104, 2014.

[6] E. Black, A. J. Coles, and C. Hampson. Planning for persuasion. In *AAMAS*, pages 933–942. ACM, 2017.

[7] P. Borah, J. Hwang, and Y. C. L. Hsu. Covid-19 vaccination attitudes and intention: Message framing and the moderating role of perceived vaccine benefits. *Journal of Health Communication*, 26:523–533, 2021.

[8] S. Brams, D. Kilgour, and M. Sanver. *A Minimax Procedure for Negotiating Multilateral Treaties*, pages 265–282. Springer, 2007.

[9] J. Carrillo and T. Palfrey. The compromise game: Two-sided adverse selection in the laboratory. *American Economic Journal: Microeconomics*, 1:151—181, 2009.

[10] L. Chalaguine and A. Hunter. A persuasive chatbot using a crowd-sourced argument graph and concerns. In *COMMA*, volume 326 of *FAIA*, pages 9–20. IOS Press, 2020.

[11] L. Chalaguine and A. Hunter. Addressing popular concerns regarding COVID-19 vaccination with natural language argumentation dialogues. In *ECSQARU*, volume 12897 of *LNCS*, pages 59–73. Springer, 2021.

[12] H. Corley. A mixed cooperative dual to the nash equilibrium. *Game Theory*, 2015. Article ID 647246.

[13] J. Demsar. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.*, 7:1–30, 2006.

[14] I. Donadello and M. Dragoni. AI-enabled persuasive personal health assistant. *Soc. Netw. Anal. Min.*, 12(1):106, 2022.

[15] I. Donadello, M. Dragoni, and C. Eccher. Explaining reasoning algorithms with persuasiveness: a case study for a behavioural change system. In *SAC*, pages 646–653. ACM, 2020.

[16] I. Donadello, A. Hunter, S. Teso, and M. Dragoni. Machine learning for utility prediction in argument-based computational persuasion. In *AAAI*, pages 5592–5599. AAAI Press, 2022. URL https://ojs.aaai.org/index.php/AAAI/article/view/20499.

[17] X. Fan and F. Toni. Mechanism design for argumentation-based persuasion. In *COMMA*, volume 245 of *Frontiers in Artificial Intelligence and Applications*, pages 322–333. IOS Press, 2012.

[18] K. Gallagher and J. Updegraff. Health message framing effects on attitudes, intentions, and behavior: A meta-analytic review. *Annals of Behavioral Medicine*, 43(1):101–116, 2012.

[19] F. Grasso, A. Cawsey, and R. Jones. Dialectical argumentation to solve conflicts in advice giving: a case study in the promotion of healthy nutrition. *International Journal of Human-Computer Studies*, 53(6):1077–1115, 2000.

[20] C. Hadjinikolis, Y. Siantos, S. Modgil, E. Black, and P. McBurney. Opponent modelling in persuasion dialogues. In *IJCAI*, pages 164–170. IJCAI/AAAI, 2013.

[21] E. Hadoux and A. Hunter. Comfort or safety? gathering and using the concerns of a participant for better persuasion. *Argument & Computation*, 10:113–147, 2019.

[22] E. Hadoux, A. Beynier, N. Maudet, P. Weng, and A. Hunter. Optimization of probabilistic argumentation with markov decision models. In *IJCAI*, pages 2004–2010, 2015.

[23] E. Hadoux, A. Hunter, and S. Polberg. Biparty decision theory for dialogical argumentation. In *COMMA*, volume 305 of *Frontiers in Artificial Intelligence and Applications*, pages 233–240. IOS Press, 2018.

[24] E. Hadoux, A. Hunter, and S. Polberg. Strategic argumentation dialogues for persuasion: Framework and experiments based on modelling the beliefs and concerns of the persuadee. *Argument and Computation*, 14(2):109–161, 2023.

[25] Y. Huang, W. Chang, W. Li, and Z. Lin. Aggregation of utility-based individual preferences for group decision-making. *European Journal of Operational Research*, 229(2):462–469, 2013.

[26] A. Hunter. Towards a framework for computational persuasion with applications in behaviour change. *Argument & Computation*, 9(1):15–40, 2018.

[27] A. Hunter and M. Thimm. Optimization of dialectical outcomes in dialogical argumentation. *International Journal of Approximate Reasoning*, 78:73–102, 2016.

[28] A. Hunter, L. Chalaguine, T. Czernuszenko, E. Hadoux, and S. Polberg. Towards computational persuasion via natural language argumentation dialogues. In *KI*, volume 11793 of *LNCS*, pages 18–33. Springer, 2019.

[29] H. Katsumi, T. Hiraoka, K. Yoshino, K. Yamamoto, S. Motoura, K. Sadamasa, and S. Nakamura. Optimization of information-seeking dialogue strategy for argumentation-based dialogue system. In *Proceedings of DEEP-DIAL@AAAI'19*, volume abs/1811.10728. ArXiv, 2018.

[30] K. Kilic, S. Weck, T. Kampik, and H. Lindgren. Argument-based human-ai collaboration for supporting behavior change to improve health. *Frontiers in Artificial Intelligence*, 6, 2023.

[31] S. Lawes-Wickwar, D. Ghio, M. Tang, C. Keyworth, S. Stanescu, J. Westbrook, E. Jenkinson, A. Kassianos, D. Scanlan, N. Garnett, and *et al.* A rapid systematic review of public responses to health messages encouraging vaccination against infectious diseases in a pandemic or epidemic. *Vaccines*, 9(2), 2021.

[32] A. Monteserin and A. Amandi. A reinforcement learning approach to improve the argument selection effectiveness in argumentation-based negotiation. *Expert Systems with Applications*, 40:2182–2188, 2013.

[33] M. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[34] M. Peterson. *An Introduction to Decision Theory*. Cambridge University Press, 2009.

[35] N. Rach, W. Minker, and S. Ultes. Markov games for persuasive dialogue. In *COMMA*, volume 305 of *Frontiers in Artificial Intelligence and Applications*, pages 213–220, 2018.

[36] I. Rahwan and K. Larson. Pareto optimality in abstract argumentation. In *AAAI*, pages 150–155. AAAI Press, 2008.

[37] I. Rahwan, P. Pasquier, L. Sonenberg, and F. Dignum. A formal analysis of interest-based negotiation. *Annuals of Mathematics and Artificial Intelligence*, 55:253–276, 2009.

[38] T. Rienstra, M. Thimm, and N. Oren. Opponent models with uncertainty for strategic argumentation. In *IJCAI*, pages 332–338. IJCAI/AAAI, 2013.

[39] R. Riveret, Y. Gao, G. Governatori, A. Rotolo, J. Pitt, and G. Sartor. A probabilistic argumentation framework for reinforcement learning agents - towards a mentalistic approach to agent profiles. *Autonomous Agents and Multi-Agent Systems*, 33(1-2):216–274, 2019.

[40] A. Rosenfeld and S. Kraus. Providing arguments in discussions on the basis of the prediction of human argumentative behavior. *ACM Transactions on Interactive Intelligent Systems*, 6(4):30:1–30:33, 2016.

[41] A. Rosenfeld and S. Kraus. Strategical argumentative agent for human persuasion. In *ECAI*, volume 285 of *Frontiers in Artificial Intelligence and Applications*, pages 320–328. IOS Press, 2016.

[42] A. Rothman and P. Salovey. Shaping perceptions to motivate healthy behavior: The role of message. *Psychological Bulletin*, 121(1):3–19, 1997.