Biparty Decision Theory for Dialogical Argumentation

Emmanuel HADOUX , Anthony HUNTER [1], and Sylwia POLBERG
Department of Computer Science,
University College London, London, UK

Abstract. Proposals for strategies for dialogical argumentation often focus on situations where one of the agents wins the dialogue and the other agent loses. Yet in real-world argumentation, it is common for agents to not involve such zero-sum situations. Rather, the agents may enter into a dialogue with divergent but not necessarily opposing views on what is important in the outcomes from the argumentation. In order to model this kind of situation, we investigate a decision-theoretic approach that allows different participants to have different utility evaluations of a dialogue, and for the proponent to model the opponent's utility evaluation in order to optimize the choice of move in the dialogue.

Keywords. Dialogical argumentation; Argumentation strategies, Opponent modelling.

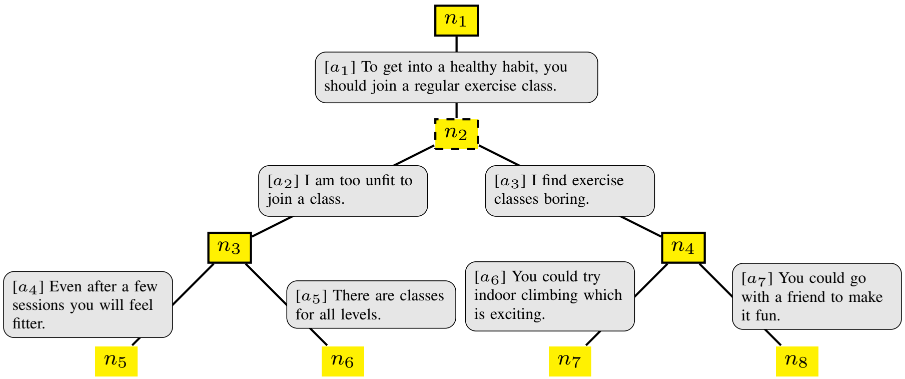
1. Introduction

In the literature on computational models of dialogical argument, there is an emphasis on situations where agents take opposing positions on the outcomes from the dialogues, and the aim of each agent is to win and thereby make the other agent lose. However, for some dialogues, this may be an over-simplification since it may overlook the fact that each participant may be seeking different but not necessarily conflicting outcomes.

For example, in a discussion on politics, participant P_1 might want to persuade participant P_2 to accept an argument A_1 about the need to raise funding in hospitals, and participant P_2 might want to persuade P_1 to accept an argument A_2 about the need to cut government expenditure. Depending on the other arguments and the relationships between them, it is possible for both agents to accept both A_1 and A_2, yet participant P_1 might have high utility in A_1 and low utility in A_2 being accepted, and participant P_2 might have the reverse.

In this paper, we investigate how by taking into account the utility function of the opponent, the proponent has a more accurate picture of how the opponent may behave in a dialogue. This information can be harnessed by the proponent to make better choices of move. For example, if the proponent wishes to persuade the opponent to accept a claim, the proponent may attempt to use an argument for that claim for which the opponent would assign a high utility value. Our approach is to use games in extensive form [11] and adapt them for argumentation.

[1] Corresponding Author: Anthony Hunter, UCL Department of Computer Science, UCL, London, UK; Email: anthony.hunter@ucl.ac.uk



n_1
[a_1] To get into a healthy habit, you should join a regular exercise class.
n_2
[a_2] I am too unfit to join a class.
[a_3] I find exercise classes boring.
n_3
n_4
[a_4] Even after a few sessions you will feel fitter.
[a_5] There are classes for all levels.
[a_6] You could try indoor climbing which is exciting.
[a_7] You could go with a friend to make it fun.
n_5
n_6
n_7
n_8

Figure 1. A decision tree for an argumentation dialogue. Each arc is labelled with an argument that is posited in a dialogue. Each branch denotes a dialogue involving exactly three arguments with the first (respectively second) being posited by the proponent (respectively opponent). Proponent (decision) nodes are solid boxes and opponent (chance) nodes are dashed boxes.

2. Decision trees for argumentation

We consider a subset of extensive form games that we refer to as decision trees. A decision tree represents all the possible combinations of decisions and outcomes of a sequential decision-making problem. In a two-agent problem, where the agents take turns, a path from the root to any leaf crosses alternately nodes associated with the proponent (called *decision nodes*) and nodes associated with the opponent (called *chance nodes*).

In the case of dialogical argumentation, a decision tree represents all possible dialogues. Each path is one possible permutation of the moves permitted by the dialogue protocol *i.e.*, one possible complete dialogue between the two agents. An edge between any two nodes n and n' in the tree is the decision (i.e. dialogue move) that has to be taken by the corresponding agent in order to transition from node n to node n'.

In this paper, we do not restrict ourselves to a specific dialogue protocol, and so there are various options for constructing a decision tree. For instance, we could assume that the only kind of move is the posit of an argument from an argument graph, or we could allow moves such as queries, concessions, and retractions. It could be non-exhaustive so as to reflect when one or both agents decide to stop participating in the dialogue. Alternatively, it could be exhaustive, for instance in order to directly reflect Dung's semantics [4], which can be undertaken by adopting the dialogue protocols of Vreeswijk and Prakken [18] or Caminada [1]. In the examples, we restrict consideration to each move being a posit of an argument and the participants take turns to present their argument. Each argument is a label on an arc in the decision tree (as illustrated in Figure 1).

Once the decision tree is built, we select, in each decision node, an action to perform (*e.g.*, an argument to posit in each state of the debate) from the point of view of the proponent. This association of a node with the action to perform in this node is called a *policy*. The aim is to compute an *optimal policy*. This is the policy that selects the best action to perform in each decision node. For this, we use a decision rule, composed of two parts: one taking account of the values of all children of a decision node and the other taking account of the values of all the children of a chance node. We consider some options for decision rules in the following sections.

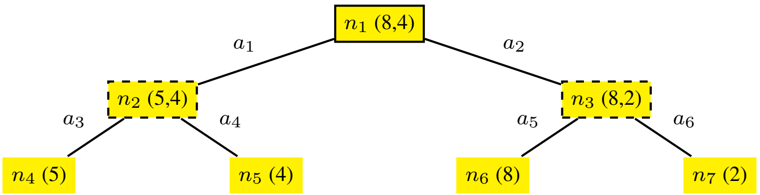


Figure 2. A decision tree to illustrate the maximax and maximin decision rules. The node n_1 is a decision node (box with solid line), and the nodes n_2 and n_3 are chance nodes (box with dashed line). Each arc is labelled with an argument. For the leaf nodes, the value in brackets is the utility value. For the non-leaf nodes, the first value in brackets is the value for the Q function for the maximax case, and the second value is the value for the maximin case. We assume $\delta = 1$. The policy for the maximax case is $\Pi(n_1) = a_2$ and the policy for the maximin case is $\Pi(n_1) = a_1$.

3. Decision rules for argumentation

We now review two simple decision rules and show how they can be used for our decision rules for argumentation. In the following definitions, we assume that we have a decision tree T, a labelling function L that assigns a label (which represents a move in the dialogue such as positing a particular argument) to each arc in T, a utility function U (which gives a utility value to each leaf), and a discount factor δ (which is used to discount the utility of longer branches). Also, for a utility function U, and a node n in the decision tree T, we define the AMax and AMin functions as follows, where $\mathsf{Children}(T, n)$ is the set of children for n in T. Essentially, AMax (respectively AMin) gives the children with maximum (respectively minimum) utility.

$$\mathsf{AMax}(T, U, n) = \{n' \in \mathsf{Children}(T, n) \mid \text{ for all } n'' \in \mathsf{Children}(T, n), U(n') \geq U(n'')\}$$
$$\mathsf{AMin}(T, U, n) = \{n' \in \mathsf{Children}(T, n) \mid \text{ for all } n'' \in \mathsf{Children}(T, n), U(n') \leq U(n'')\}$$

We start with the maximax rule which is specified in Definition 1 and illustrated in Figure 2. This rule is applied if the proponent wants to adopt an optimistic behaviour, *i.e.*, to consider that the opponent wants to maximize the outcome as well. Recall that the outcome is given from the point of view of the proponent only. The Q function captures the utility backed-up the tree (i.e. for each non-leaf node, the utility assigned to the node is based on the maximum utility of its children), and it is decreased by the discount factor. Note that the maximax policy is not necessarily unique; indeed, none of the proposals we consider in this paper necessarily give a unique policy.

Definition 1. *A* **maximax policy** *for* (T, L, U, δ) *is* $\Pi : \mathsf{Nodes}(T) \rightarrow \mathsf{Nodes}(G)$ *defined as follows using the calculation of the* $Q : \mathsf{Nodes}(T) \rightarrow \mathbb{R}$ *function.*

- *If n is a leaf node, then $Q(n) = U(n)$.*
- *If n is a chance node, and $n_i \in \mathsf{AMax}(T, Q, n)$, then $Q(n) = \delta \times Q(n_i)$.*
- *If n is a decision node, and $n_i \in \mathsf{AMax}(T, Q, n)$, then $Q(n) = \delta \times Q(n_i)$, and $\Pi(n) = L(n, n_i)$.*

We now consider the maximin rule which is specified in Definition 2 and illustrated in Figure 2 [19]. This is the pessimistic selection since the proponent assumes that the opponent will always try to minimize the outcome. It can be related to a two-player zero-sum game where a negative outcome for one player is positive for the other.

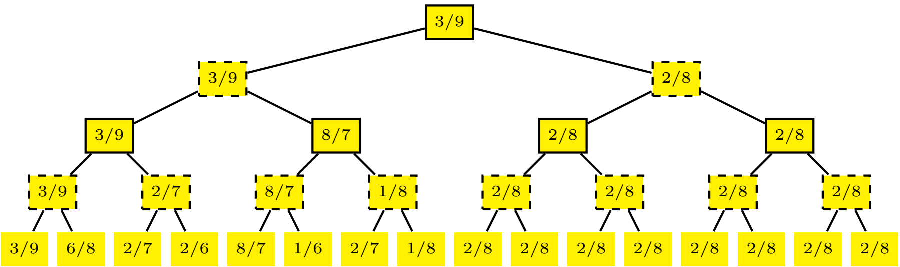


Figure 3. A decision tree with the bimaximax values passed back up the tree. Each node n in the decision tree has the biparty utility in the form x/y where x is the backed-up proponent utility $Q^p(n)$ and y is the backed-up opponent utility $Q^o(n)$. So at a proponent node, the child with x/y such that x is largest is chosen, whereas at an opponent node, the child with x/y such that y is largest is chosen.

Definition 2. *A* **maximin policy** *for* (T, L, U, δ) *is* $\Pi : \mathsf{Nodes}(T) \rightarrow \mathsf{Nodes}(G)$ *defined as follows using the calculation of the* $Q : \mathsf{Nodes}(T) \rightarrow \mathbb{R}$ *function.*

- *If* n *is a leaf node, then* $Q(n) = U(n)$.
- *If* n *is a chance node, and* $n_i \in \mathsf{AMin}(T, Q, n)$*, then* $Q(n) = \delta \times Q(n_i)$.
- *If* n *is a decision node, and* $n_i \in \mathsf{AMax}(T, Q, n)$*, then* $Q(n) = \delta \times Q(n_i)$*, and* $\Pi(n) = L(n, n_i)$.

In previous work, we have shown how dialogical argumentation can be optimized by the application of decision rules [9]. The outcome of each branch of a dialogue can be judged, the utility of that outcome evaluated, and the optimal policy determined. The outcome of a branch can be judged in various ways including in terms of belief in a persuasion goal at the leaf or the set of accepted arguments at the leaf. For the latter, an abstract argument graph can be constructed from the moves in the dialogue, and then Dung's dialectical semantics (as defined in [4]) can be applied to the resulting graph.

4. Biparty decision theory

In biparty decision theory, we assume that each agent i has a utility function U^i. The aim is for U^i to reflect what agent i would regard as the benefit of each possible outcome (i.e. each leaf of the decision tree). To use this, we consider the bimaximax decision rule in Definition 3 and illustrate it in Figure 3. This assumes that we have a utility function U^p for the proponent and a utility function U^o for the opponent. The bimaximax rule is a generalization of the maximax rule where the U^p is maximized at a decision node, and U^o is maximized at a chance node.

Definition 3. *A* **bimaximax policy** *for* (T, L, U^p, U^o, δ) *is* $\Pi : \mathsf{Nodes}(T) \rightarrow \mathsf{Nodes}(G)$ *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* $Q^o(n) = \delta \times Q^o(n_i)$ *and* $Q^p(n) = \delta \times Q^p(n_i)$.
- *If* n *is a decision node, and* $n_i \in \mathsf{AMax}(T, Q^p, n)$*, then* $Q^o(n) = \delta \times Q^o(n_i)$ *and* $Q^p(n) = \delta \times Q^p(n_i)$ *and* $\Pi(n) = L(n, n_i)$.

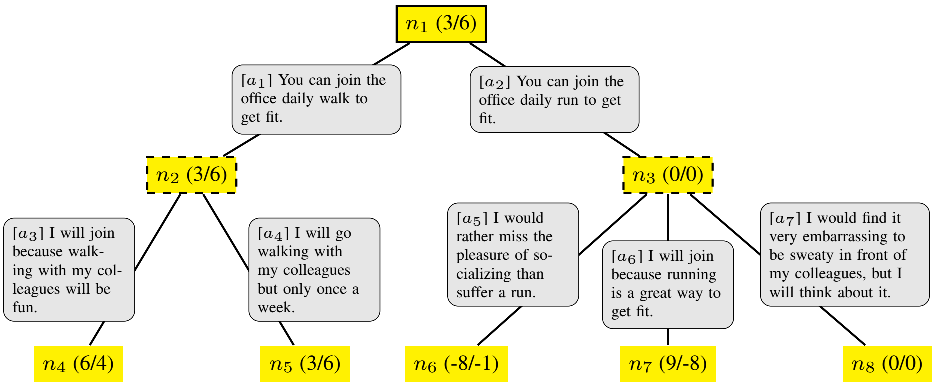

n_1 (3/6)
[a_1] You can join the office daily walk to get fit.
[a_2] You can join the office daily run to get fit.
n_2 (3/6)
n_3 (0/0)
[a_3] I will join because walking with my colleagues will be fun.
[a_4] I will go walking with my colleagues but only once a week.
[a_5] I would rather miss the pleasure of socializing than suffer a run.
[a_6] I will join because running is a great way to get fit.
[a_7] I would find it very embarrassing to be sweaty in front of my colleagues, but I will think about it.
n_4 (6/4)
n_5 (3/6)
n_6 (-8/-1)
n_7 (9/-8)
n_8 (0/0)

Figure 4. A decision tree for Example 1 with the bimaximax values passed back up the tree. Each node n in the decision tree has the biparty utility in the form x/y where x is the backed-up proponent utility $Q^p(n)$ and y is the backed-up opponent utility $Q^o(n)$. Each arc is labelled with an argument that is posited in a dialogue. Each branch denotes a dialogue involving exactly two arguments with the first and third (respectively second) being posited by the proponent (respectively opponent).

Example 1. *Consider the arguments (which are presented as enthymemes and so involve some implicit premises and/or claims) that appear in the decision tree in Figure 4. We suppose that each branch is a possible dialogue according to an assumed protocol. From the point of view of the proponent (who wants the opponent to increase exercise), the utility could be the following.*

- n_4 *Having the opponent agree to a daily walk is a good outcome* ($U^p(n_4) = 6$).
- n_5 *Having the opponent disagree to a daily run but agree to a weekly walk is positive but not as good as agreeing to a daily walk* ($U^p(n_5) = 3$).
- n_6 *Having the opponent disagree to a run and therefore disagreeing to do exercise is the worst outcome* ($U^p(n_6) = -8$).
- n_7 *Having the opponent agree to a daily run is the best outcome* ($U^p(n_7) = 9$).
- n_8 *Having the opponent reject a daily run, but agree to think about it, is a neutral outcome* ($U^p(n_8) = 0$).

From the point of view of the opponent (who is mildly interested in improving his/her health by exercise but does not want to do much exercise and in addition likes socializing), the utility could be the following.

- n_4 *Agreeing to the opportunity for socializing, despite doing a walk daily, is a reasonable good outcome* ($U^0(n_4) = 4$).
- n_5 *Agreeing to a weekly walk is better than to a daily walk* ($U^0(n_5) = 6$).
- n_6 *Disagreeing to a daily run means that there is no exercise is done and so the desire to do a little exercise is not met* ($U^0(n_6) = -1$).
- n_7 *Agreeing to a daily run would be a bad outcome* ($U^0(n_7) = -8$).
- n_8 *Disagreeing to a daily run but agreeing to think about it is a neutral outcome* ($U^0(n_8) = 0$).

So the participants agree that n_4 *and* n_5 *have positive utility (though there is some difference in the actual values assigned),* n_6 *has negative utility (though with substantial difference in the actual value assigned), and* n_8 *has zero utility. The participants completely*

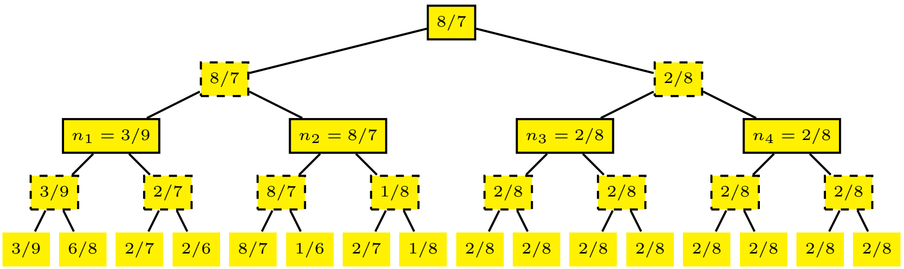


Figure 5. A decision tree with the bimaxilocal values passed back up the tree. Each node n in the decision tree has the biparty utility in the form x/y where x is the backed-up proponent utility $Q^p(n)$ and y is the backed-up opponent utility $Q^o(n)$. So at a proponent node, the child with x/y such that x is largest is chosen, whereas at an opponent node, the child with arc with label with highest utility (according to the opponent utility function) is chosen. For this example, suppose $U^o(n_2) > U^o(n_1)$ and $U^o(n_3) > U^o(n_4)$.

disagree on n_7. Given these considerations by the participants, we obtain the backed-up values given in the decision tree in Figure 4.

We now consider a variant of the bimaximax decision rule called the bimaxilocal decision rule in Definition 4 and illustrate it in Figure 5. At a chance node, the rule assumes that the opponent chooses the action that maximizes its utility at that step. So it assumes that the opponent makes a local choice and does not consider the utility of the whole branch. For argumentation, it captures an agent who just decides what to say without considering the wider ramifications.

Definition 4. *A* **bimaxilocal policy** *for* (T, L, U^p, U^o, δ) *is* $\Pi : \mathsf{Nodes}(T) \rightarrow \mathsf{Nodes}(G)$ *defined as follows using the calculation of the* $Q^p : \mathsf{Nodes}(T) \rightarrow \mathbb{R}$ *function.*

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

The bimaxilocal policy models an opponent who provides its highest utility counterargument to each argument by the proponent, but does not think in terms of choosing its counterarguments based on what further arguments the proponent may put forward. This may be useful for modelling an agent who is behaving intelligently within a conversation but not behaving strategically. As an example, consider a doctor (proponent) trying to persuade a patient (opponent) to give up smoking. The patient may choose the best arguments to posit at each stage of the dialogical but might not be aiming to maximize its utility from the overall dialogue.

5. Properties of biparty decision rules

We can show that the bimaximax decision rule subsumes the maximax decision rules by choosing the biparty utility appropriately. For instance, we can show this when the utility for the opponent and the utility for the proponent are the same for each outcome. Similarly, we can show that the bimaximax decision rule subsumes the maximin decision rules by choosing the biparty utility appropriately. For instance, we can show this when the utility for the opponent is the reciprocal of the utility for the proponent for each outcome. Note, all propositions and proofs are available in the proof appendix[2].

The maximax rule captures the bimaximax rule in the sense that we can find a utility function to use with the maximax rule that gives the same policy as a bimaximax policy. However, there is no simple way of obtaining this utility function. Indeed, it seems to call for applying the bimaximax rule to the decision tree in order to reverse engineer an appropriate utility function.

In the definitions of the bimaximax and bimaxilocal policies, we use the function Q^p to identify the backed-up utility for the proponent. So for a policy Π, the value for Q^p at the root is the utility that the proponent would expect to get by following the policy. We call this value the **gain** of the policy Π. For example, in Figure 3, the gain is 3, and in Figure 5, the gain is 8. In the following results, we consider gain.

Proposition 1. *Let Π_1 and Π_2 be bimaximax policies of (T, L, U^p, U^o, δ), and let γ_1 (respectively γ_2) be the gain of Π_1 (respectively Π_2). If U^p and U^o are injective functions (i.e. for each function, no two nodes have the same utility value). then $\gamma_1 = \gamma_2$, else it is not guaranteed that $\gamma_1 = \gamma_2$.*

Proposition 2. *If Π_1 is a bimaximax policy of (T, L, U^p, U^o, δ), and Π_2 is a bimaxilocal policy of (T, L, U^p, U^o, δ), and γ_1 (respectively γ_2) is the gain of Π_1 (respectively Π_2), then it is not necessarily the case that $\gamma_1 \leq \gamma_2$ and or that $\gamma_1 \geq \gamma_2$.*

Considering gain is important if we want to be sure of what we can get out of a dialogue. It is also important if we want to consider the appropriateness of a decision rule for taking account of a particular opponent. For instance, as the distance between U^p and U^o grows, the maximax policy (which just uses U^p) tends to be over-optimistic in gain with respect to the bimaximax decision rule.

6. Discussion

Most proposals for dialogical argumentation focus on protocols (*e.g.*, [12,13,5,2]), though there is increasing interest in strategies (see [17] for a review). Game theoretic analyses have been applied to argumentation as a one step process where all arguments are presented and then evaluated, rather than in a dialogue where arguments are presented over a number of steps [14]. Mechanism design has been applied to dialogical argumentation with a focus on issues such as lying [6]. There are proposals for using probability theory to select a move based on a probabilistic model of the opponent (e.g. [15,7,10,8]). In previous work [9], we have shown how classical decision rules such as maximax, maximin, Hurwicz, and Laplace, can be used for optimizing the choice of move in dialogical argumentation taking into account the degree to which the opponent is being adversarial and/or the degree to which the model of the opponent is uncertain, though it does not take the utility of the opponent into account. Modelling the utility of the opponent has been considered in [15] using an adaptation of the M^* algorithm by [3], and in [16] using the notion of a subgame perfect equilibrium. There are further options within decision theory and game theory for modelling the utility of the opponent in

[2] For proof appendix, see http://www0.cs.ucl.ac.uk/staff/a.hunter/papers/bipartyappendix.pdf

games in extensive form, and that potentially these could be harnessed in argumentation. Taking a general approach, as done in this paper, without committing to a specific set of moves or protocol, may allow us to identify important properties or behaviour for a wide variety of argumentation systems.

Acknowledgements

This research is funded by EPSRC grant EP/N008294/1 *Framework for Computational Persuasion*.

References

[1] M. Caminada. Argumentation semantics as formal discussion. *IfCoLog Journal of Logic and its Applications*, 4(8):2457–2492, 2017.

[2] M. Caminada and M. Podlaszewski. Grounded semantics as persuasion dialogue. In *Proc. of the International Conference on Computational Models of Argument (COMMA)*, pages 478–485. IOS Press, 2012.

[3] D. Carmel and S. Markovitch. Incorporating opponent models into adversary search. In *Proc. of the National Conference on Artificial Intelligence (AAAI)*, pages 120–125. AAAI Press, 1996.

[4] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[5] X. Fan and F. Toni. Assumption-based argumentation dialogues. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 198–203. AAAI Press, 2011.

[6] X. Fan and F. Toni. Mechanism design for argumentation-based persuasion. In *Proc. of the International Conference on Computational Models of Argument (COMMA)*, pages 322–333. IOS Press, 2012.

[7] C. Hadjinikolis, Y. Siantos, S. Modgil, E. Black, and P. McBurney. Opponent modelling in persuasion dialogues. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 164–170. AAAI Press, 2013.

[8] E. Hadoux, A. Beynier, N. Maudet, P. Weng, and A. Hunter. Optimization of probabilistic argumentation with Markov decision models. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2004–2010. AAAI Press, 2015.

[9] E. Hadoux and A. Hunter. Strategic sequences of arguments for persuasion using decision trees. In *Proc. of the AAAI Conference on Artificial Intelligence (AAAI)*, pages 1128–1134. AAAI Press, 2017.

[10] A. Hunter. Probabilistic strategies in dialogical argumentation. In *Proc. of the International Conference on Scalable Uncertainty Management (SUM)*, volume 8720 of *LNCS*, pages 190–202. Springer, 2014.

[11] M. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1999.

[12] H. Prakken. Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation*, 15(6):1009–1040, 2005.

[13] H. Prakken. Formal sytems for persuasion dialogue. *Knowledge Engneering Review*, 21(2):163–188, 2006.

[14] I. Rahwan and K. Larson. Pareto optimality in abstract argumentation. In *Proc. of the AAAI Conference on Artificial Intelligence (AAAI)*, pages 150–155. AAAI Press, 2008.

[15] T. Rienstra, M. Thimm, and N. Oren. Opponent models with uncertainty for strategic argumentation. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 332–338. AAAI Press, 2013.

[16] R. Riveret, H. Prakken, A. Rotolo, and G. Sartor. Heuristics in argumentation: A game-theoretical investigation. In *Proc. of the International Conference on Computational Models of Argument (COMMA)*, pages 324–335. IOS Press, 2008.

[17] M. Thimm. Strategic argumentation in multi-agent systems. *Kunstliche Intelligenz*, 28:159–168, 2014.

[18] G. Vreeswijk and H. Prakken. Credulous and sceptical argument games for preferred semantics. In *Proc. of the European Workshop on Logic for Artificial Intelligence (JELIA)*, volume 1919 of *LNCS*, pages 239–253. Springer, 2000.

[19] A. Wald. Basic ideas of a general theory of statistical decision rules. In *Proc. of the International congress of Mathematicians*, volume 1, pages 308–325, 1950.