

# Strategic Dialogical Argumentation using Multi-Criteria Decision Making with Application to Epistemic and Emotional Aspects of Arguments

Emmanuel Hadoux<sup>1</sup> (✉), Anthony Hunter<sup>1</sup>, and Jean-Baptiste Corr ge<sup>2</sup>

<sup>1</sup> University College London, Department of Computer Science, London, UK  
{e.hadoux, anthony.hunter}@ucl.ac.uk

<sup>2</sup> LIMSI, CNRS, Universit  Paris-Saclay, F-91405 Orsay  
jean-baptiste.correge@limsi.fr

**Abstract.** Participants in dialogical argumentation often make strategic choices of move, for example to maximize the probability that they will persuade the other opponents. Multiple dimensions of information about the other agents (*e.g.*, the belief and likely emotional response that the other agents might have in the arguments) might be used to make this strategic choice. To support this, we present a framework with implementation for multi-criteria decision making for strategic argumentation. We provide methods to improve the computational viability of the framework, and analyze these methods theoretically and empirically. We finally present decision rules supported by the psychology literature and evidence using human experiments.

## 1 Introduction

In dialogical argumentation, a proponent can use strategic choices of argument when arguing with an opponent. In particular, when the number of moves is limited, the proponent needs to choose arguments that are more likely to be effective rather than exhaustively presenting all arguments. Consider for example, a doctor trying to persuade a patient to give up smoking. The doctor cannot expect the patient to have a discussion with hundreds of arguments and counterarguments being presented. Rather, the doctor has to think of what arguments and counterargument the patient believes, what arguments may have an emotional impact on the patient, etc., and then the doctor has to choose a line of argumentation that may be concluded with a relatively small number of arguments being presented by either side.

Most proposals for dialogical argumentation focus on protocols (*e.g.*, [1–4]) with strategies being under-developed. See [5] for a review of strategies in multi-agent argumentation. Strategies in argumentation have been analyzed using game theory (*e.g.*, [6, 7]), but these are more concerned with issues of manipulation rather than persuasion. There are also proposals for using probability theory to, for instance, select a move based on what an agent believes the other

is aware of [8], or, approximately predict the argument an opponent might put forward based on an history [9]. The problem can also be viewed as a probabilistic finite state machine, and generalized to POMDPs when there is uncertainty on the internal state of the opponent [10]. In [11], a planning system is used by the persuader to optimize choice of arguments based on belief in premises. But, none of these developments have systematically harnessed established notions in decision theory for maximizing the outcome of a dialogue.

To address this, Hadoux and Hunter [12] propose a general framework for representing persuasion dialogues as a decision tree, and for using decision rules such as MaxiMin for selecting moves. This is useful if we can focus on one dimension for modelling a user, such as her belief in the arguments, but it does not allow us to take into account multiple dimensions, and yet, multiple dimensions are often desirable. For instance, a user may want to maximize both the belief and appeal in an argument. In this paper, we will present a general framework for representing persuasion dialogues as a decision tree, and use multi-criteria decision making for computing an optimal policy. We will illustrate the use of these multiple dimensions by considering belief in arguments, and the emotional response evoked by arguments. For the latter, we draw on an established results from psychology for modelling emotional response in terms of valence (*i.e.*, polarity), arousal (*i.e.*, intensity) and dominance (*i.e.*, degree of feeling in control) invoked by an argument.

We proceed as follows: we review basic definitions for dialogical argumentation in Section 2; then in Section 3 we take emotional effect of arguments into account. We review decision trees and multi-criteria decision making for modelling argumentation dialogues in Section 4. We instantiate the multi-criteria decision making theory on multi-dimensional problems with emotions in Section 5. In Section 6 we specify size-reduction features to improve space efficiency and present some theoretical results concerning them. Sections 7 and 8 present both computational empirical results concerning the implementation and results on human experiments to validate our choice of decision rules. Finally, we discuss our contributions in Section 9.

## 2 Preliminaries

A *persuader* (the proponent) has a dialogue with a *persuadee* (the opponent) to make her believe (or disbelieve) some combination of arguments as a goal (*e.g.*, to do more exercise or to eat healthier food). We see that normally getting a persuadee to believe arguments is a prerequisite in the persuasion process.

For the sake of simplicity, in this paper, we deal with two agents and a singleton persuasion goal (*i.e.*, an individual argument as goal). However, our work can be extended to more agents and goals as long as only one persuader is involved. Building upon Dung’s [13] abstract argumentation, a dialogue concerns an argument graph  $G$  without self-attacks where  $\mathbf{Args}(G)$  is the set of arguments in  $G$ , and  $\mathbf{Attacks}(G)$  is the set of attack relations in  $G$ . We do not consider self-attacks in this work as we are concerned with applications where we assume

participants do not present self-contradictory statements. For real applications, this restriction is often not important.

More formally, a *persuasion dialogue* is a sequence of moves  $D = [m_1, \dots, m_h]$ . In this work, a move consists in positing an argument  $A \in \text{Args}(G)$ . The attacks to and from this argument in relation to the arguments already posited come from the original graph. The parameter  $h$  is the *horizon* of the debate, *i.e.*, the maximum number of moves that can be played. It is justified by the need to keep the persuadee engaged. A shorter debate (*i.e.*, a smaller value for  $h$ ) gives more chance to keep the persuadee in the debate until the end. However, it also lowers the number of ways to make a valid point.

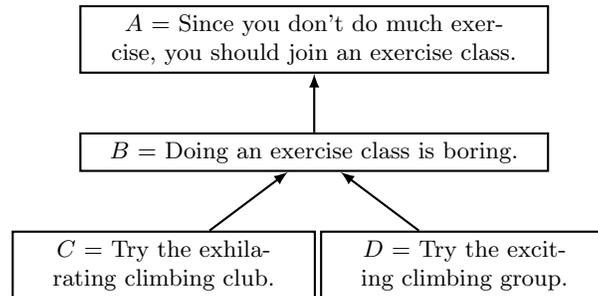
Each odd (resp. even) move in the dialogue is a persuader (resp. persuadee) move. However, the persuadee moves are played with respect to the arguments she believes, in reaction to the persuader positing an argument. Therefore, an efficient strategy needs to take into account the possible subsets of arguments the persuadee believes. Indeed, an agent is unlikely to posit arguments she does not have faith in. To that end, the persuader keeps and updates a *belief model* of the persuadee and uses it in her decision process. We use the epistemic approach to probabilistic argumentation [14–17], defining a model as an assignment in the unit interval to each argument where for an argument  $A$ ,  $B(A) > 0.5$  represents  $A$  is believed to some degree,  $B(A) = 0.5$  represents  $A$  is neither believed nor disbelieved, and  $B(A) < 0.5$  represents  $A$  is disbelieved to some degree.

### 3 Taking Emotion into Account

In addition to belief, the emotions invoked by arguments are important to take into account since they affect the way the argument is perceived by the persuadee. Emotions are the result of how an individual appraises a stimulus [18]. According to [19], appraisal is a cognitive process composed of a number of checks aimed at categorizing a stimulus: is it relevant, what does it imply, do I have the potential to cope and is it socially significant?

This process and the various patterns of checks generate different cognitive responses and coping strategies. These strategies in turn affect the way information is processed [20]. For example, guilt leads to the use of active strategies focused on repairing the wrong made, whereas shame leads to the use of more passive strategies focused on the self. Combined with gain-loss framing [21], the emotion conveyed by an argument can be used to increase the persuasiveness of this argument. Duhachek et al. [20] developed a study in which they tested different argument configurations, varying the emotional tone of the argument (guilt v. shame) and the framing of the argument (gain v. loss). The results showed that a positively-framed message associated with guilt (“What you have to gain by drinking responsibly”) is processed more fluently than a negatively-framed message associated with guilt (“What you have to lose by not drinking responsibly”).

While Ekman [22] considered only 6 basic emotions (anger, disgust, fear, happiness, sadness and surprise), the definition and characterization of emotions



**Fig. 1.** Example of an argument graph.

has been widely discussed in psychology. Emotions in argumentation have also been investigated recently using logic and sets of discrete emotions (see, *e.g.*, [23–25]).

In this paper, we propose to focus on affective norm as used in the database built by Bradley and Lang [26]. These capture the emotional response to specific words in three dimensions: arousal (ranging from excited to calm), valence (pleasant to unpleasant), and dominance (from being in control to being dominated). For example, for valence scores, *leukemia* and *murder* are low and *sunshine* and *lovable* are high; for arousal scores, *grain* and *dull* are low and *lover* and *terrorism* are high; and for dominance scores, *dementia* and *earthquake* are low, and *smile* and *completion* are high. There are a number of databases for affective norms, and they have been used in diverse studies of emotion, behaviour, and language processing.

In this work, we draw on this psychological research into affective norms to evaluate the emotion invoked by particular choices of word in an argument. We determine an emotional scoring of arguments taken from a recent affective norm database that has nearly 14 thousand words [27]. Each word in the database has been scored by around 20 participants using crowdsourcing. Scores are also given according to gender, age group, and educational background.

By determining the emotion likely to be invoked by different candidates for a posit, strategic choices can be made by the persuader. To illustrate, consider the argument graph given in Figure 1. Suppose the persuader wants to persuade the persuadee with argument *A*. If the other agent does not believe *A* and believes *B*, the persuader can posit either *C* or *D* as counterarguments to *B*. By taking into account the score for affective norms for words in the arguments, the persuader can determine a three-dimensional score for each of *C* and *D* according to the gender, age group, and educational background of the persuadee, or on the whole population, and thereby make a choice of which of *C* and *D* to present.

The literature has focused essentially on valence and arousal, and demonstrated that these two dimensions interact with each other [28–30]. Eder and Rothermund [28], in particular, showed that words with a positive valence were treated faster when their arousal level was low than when it was high. Con-

versely, words with a negative valence were treated faster when their arousal level was high. This means that negative valence combined with high arousal allows for a more efficient cognitive processing than negative valence with low arousal, and conversely for positive valence. In this work, we are dealing with superficial interactions limited to the exchange of several arguments. They do not necessarily mean that the persuadee puts a lot of effort into processing the arguments. So following the psychology literature, we aim at using arguments that will be treated faster and thereby have more impact on the persuadee.

Psychology classically considers arousal as an interval reflecting the intensity of the stimulus and ranging from null to extreme, *i.e.*, from 0 to a positive value. In this case, an argument with a low arousal level should be treated more fluently. However, the affective norm considers arousal as a bipolar scale ranging from calmness to excitation, *i.e.*, from -1 to 1. In this case, calmness does not correspond to a null intensity of arousal but rather to a positive value on a parallel scale. In other words, in classical psychology, there is a scale from 0 to  $n$  for degree of calmness with  $n$  being the maximum calmness, and there is a scale from 0 to  $n$  for degree of excitation with  $n$  being the maximum excitation. Therefore, the middle point in the bipolar scale (*i.e.*, 0) corresponds to a state of unarousal (*i.e.*, it corresponds to 0 on each of the classical scales). As we use the affective norm databases in this paper [27, 26], understanding this correspondence is important for the decision rules we present. So instead of minimizing the arousal, we are looking to neutralize it and thus select arguments with an arousal value as close to the midpoint of the bipolar scale as possible. Note, that the values in the affective norm databases are strictly positive but are meant to be rescaled into the  $[-1,1]$  interval.

## 4 Decision Making for Dialogues

In order to get the persuadee to accept the persuasion goal, the persuader has to posit the right sequence of arguments with respect to the persuadee.

### 4.1 Decision Trees

Recently, Hadoux and Hunter [12] proposed a framework able to compute an optimal policy by taking into account every possible sequence of arguments using a *decision tree*. A decision tree represents all the possible combinations of decisions and outcomes of a sequential decision-making problem. In a problem with two agents, a path from the root to any leaf crosses alternatively nodes associated with the proponent (called *decision nodes* in this work) and nodes associated with the opponent (called *chance nodes* or *nature nodes*). In the case of a dialogue represented as a decision tree, a path is one possible permutation of the argument set, *i.e.*, one possible complete dialogue between the two agents. If horizon  $h$  is smaller than the number of arguments, every execution (and thus path) is at most of length  $h$ . In this case, it is a permutation of a subset of the argument set. An edge between any two nodes  $n$  and  $n'$  in the tree is the

decision that has to be taken by the corresponding agent in order to transition from node  $n$  to node  $n'$ . Solving a problem modelled as a decision tree amounts to computing a *policy*, an action to perform (*i.e.*, an argument to posit) in each possible state of the dialogue.

However, Hadoux and Hunter [12] evaluate branches by averaging the values and using mono-criterion decision rules. In order, to handle multiple dimensions such as belief and emotion, we evaluate branches using multi-criteria decision making (see next subsection).

Note, for this paper, we assume that the opponent is not using a model of the proponent. Rather, she is selecting what she deems as the best arguments according to what she believes and regards as a strong emotional effect. This means we assume that the opponent is not behaving stochastically nor strategically, in particular not adversarially.

This is a reasonable assumption for some applications such as where the opponent is being co-operative or at least not being competitive. Consider for example, a proponent being a doctor trying to persuade a patient to eat more healthily. The doctor is thinking strategically because she wants to persuade the patient, whereas the patient is not trying to persuade the doctor nor trying to resist the doctor. Rather she is just playing counter-arguments that she for example believes. In future work, we will investigate how both the proponent and opponent can have a model of each other, and use these to play strategically against each other.

## 4.2 Multi-criteria Decision Making

In order to take into account several dimensions as a solution of a decision problem, we apply multi-criteria decision making. In this work, we use the traditional notion of *Pareto optimality* to compare multi-dimension solutions.

**Definition 1.** (*Pareto optimality*) A solution (*i.e.*, a multi-dimension value of a leaf node)  $x = (x_1, \dots, x_n)$  **dominates** a solution  $y = (y_1, \dots, y_n)$  iff  $\forall i \in \{1, \dots, n\}, x_i \succeq y_i$  and  $\exists j \in \{1, \dots, n\}$  s.t.  $x_j \succ y_j$  where  $\succ$  denotes the preference operator.

A solution  $x$  is **Pareto optimal** if no solution  $y$  dominates it. The **Pareto front** is the set of all Pareto optimal solutions.

Unfortunately, computing the Pareto front is a very costly operation because each solution has to be compared, in the worst case scenario, to all the other ones. One might use aggregation functions in order to reduce the problem to a mono-criterion decision making problem on the aggregated criterion. However, doing so reduces the number of Pareto optimal solutions that can be considered. Using for instance the Weighted Average function implies that only the solutions located on the convex envelope of the Pareto front can be found. Indeed, let  $x^1 = (1, 3)$ ,  $x^2 = (3, 1)$  and  $x^3 = (1.5, 1.5)$  be three Pareto optimal solutions. No weights  $w_1$  and  $w_2$  exist such that  $w_1 * x_1^3 + w_2 * x_2^3 > w_1 * x_1^i + w_2 * x_2^i$  with  $i \in \{1, 2\}$ .

Fortunately, in this work, the number of solutions to compare in each node in order to find the Pareto optimal ones is at most equal to the number of arguments. It can therefore be computed easily.

## 5 Dimensions of the Dialogue Problem

In this paper, we illustrate the use of multi-criteria decision making with four dimensions which we explain below: the valuation of the dialogue, the belief in the goal, the aggregated valence and the aggregated arousal for the dialogue. Therefore, each solution in the multi-criteria decision making problem is a vector with four dimensions.

### 5.1 Valuation

The valuation of a dialogue is a real value  $v \in [0, 1]$  representing how desirable this execution is with respect to the persuasion goal of the persuader. Therefore, they are computed only for the persuader, from her point of view.

For the sake of simplicity, in this paper, the value  $v_i$  of dialogue  $i$  is a value of 1 if the goal has been posited in dialogue  $i$  or 0 otherwise. This function can be replaced to take into account several goals and interactions between them (*e.g.*, synergies). Note that Dung’s [13] dialectical semantics can also be used. For instance, the value of the goal can be 1 if it is in the grounded extension, 0 otherwise.

Also, this value can be non-binary. Gradual valuations can be used, taking into account the interactions between arguments in the graph (*e.g.*, [31]) or argument strength (*e.g.*, [32, 33]).

### 5.2 Belief

Starting from an initial value for each argument, the belief has to be updated at each step depending on the arguments played. The ambivalent method as proposed by Hunter [34] allows for belief in an argument to increase when it has been posited, and no attacker of it is believed (*i.e.*, when  $\forall (B, A) \in \text{Attacks}(G), B_{i-1}(B) \leq 0.5$ ). In this work, we use a modification of the ambivalent method that is equivalent but faster to compute (defined below). The updated belief in  $A$  (*i.e.*,  $B_i(A)$ ) is the original belief (*i.e.*,  $B_{i-1}(A)$ ) plus  $k$  times the belief in its complement (*i.e.*,  $1 - B_{i-1}(A)$ ). The  $k$  coefficient is a value in the unit interval that allows for only part of the belief to be transferred and thereby for modelling agents who do not completely believe a proposition when posited, and do not completely disbelieve a proposition when it is defeated. In addition, belief in each attackee  $C$  of  $A$  is reduced to  $k \times B_{i-1}(C)$ .

**Definition 2.** (*Fast ambivalent method*) At step  $i$  in the dialogue,  $B_i$  is generated from  $B_{i-1}$  as follows if  $D(i) = A$  and  $\forall (B, A) \in \text{Attacks}(G), B_{i-1}(B) \leq 0.5$ :

1.  $B_i(A) = B_{i-1}(A) + k \times (1 - B_{i-1}(A))$ ,

$$2. \forall C \text{ s.t. } (A, C) \in \text{Attacks}(G), B_i(C) = k \times B_{i-1}(C)$$

Whilst we use the ambivalent method in this paper, it can be replaced by any update method (see [34] for more methods). The aim of this flexibility is to model different kinds of persuadee with different kinds of behaviours, some of which are not rational.

### 5.3 Affective Norm

We now consider how we can harness the resources for affective norms in multi-criteria decision making for persuasion. The values for valence and arousal are given for singleton words and are not considered as part of a sentence. As we deal with complete arguments, we thus first need to aggregate the values to have a valence and an arousal for an argument and for a sequence of arguments. As we have seen, valence and arousal interact in a specific way. Following our discussion of the psychology literature in Section 3, in this paper we aggregate the values for an argument by taking the minimum valence and the maximum arousal across all words of this argument.

*Example 1.* Consider the arguments A1 = “Smoking causes lung cancer” and A2 = “Smoking causes serious disease”, the aggregated affective norms are as follows. So A2 is more arousing and less pleasant, and hence indicates it would be a better argument against smoking than A1.

Word/Argument	Arousal	Valence
Smoke	5	3.44
Causes	3.48	5.14
Lung	2.64	4.84
Cancer	5.14	1.9
Serious	4.05	5.88
Disease	5.5	1.68
A1	5.14	1.9
A2	5.5	1.68

Aggregating scores for valence and arousal of the keywords in the argument as the score for valence and arousal of the argument follows widespread use of keywords in text to provide the semantics of the text. This method is only meant as a simple proposal for initiating the consideration of affective norms in argumentation. By drawing on developments in affective computing (see, *e.g.*, [35]), including sentiment analysis, we could obtain a deeper understanding of the affective nature of phrase (*e.g.*, taking the use of negation into account). In this work, we aggregate across the sequence of arguments using the minimum value over each dimension.

**Table 1.** Decision rules for each dimension where  $x_v$  is the valuation,  $x_b$  the belief,  $x_{v'}$  the valence and  $x_a$  the arousal

Dim. $(x_v, x_b, x_{v'}, x_a)$	$x^1 \in X$ is preferred iff
Valuation	$x_v^1 = \arg \max_{x' \in X} x'_v$
Belief	$x_b^1 = \arg \max_{x' \in X} x'_b$
Affective norm if:	
$\forall x^i \in X, x_{v'}^i > 0$	$x_a^1 = \arg \min_{x' \in X}  x'_a $
$\forall x^i \in X, x_{v'}^i < 0$	$x_a^1 = \arg \max_{x' \in X} x'_a$
Mixed valence	$x_a^1 = \arg \max_{x' \in X} x'_a$

#### 5.4 Comparing Dimensions

In order to calculate the Pareto front, we need to compare each dimension for each solution  $x$  in the set of solutions  $X$ . For this, we need to define the preference function for each dimension. They are all summarized in Table 1.

*Valuation and beliefs* Both the valuation and the belief are real valued dimensions. Comparing them for two solutions is straightforward as we want to maximize them in this work. Note that it depends on the aim of the dialogue for the persuader. In another situation, we might want to minimize the belief (for instance, to discredit a political opponent).

*Affective norm* Because of the interactions between them, we treat the two affective norm dimensions (valence and arousal) as a pair of dimensions. We need to have different sets of rules if we want to manage different situations. If, for a given argument, all the possible words have a positive valence, we shall neutralize the arousal and select the word with an arousal value close to the mean. Conversely, if all the possible words have a negative valence, we shall maximize the arousal and select the word with the highest arousal value. When the possible words have either positive or negative valence, we shall apply the rule corresponding to the negative valence. Indeed, the literature shows that there is a “positive-negative asymmetry in evaluation” [36, 37]. This means that negative stimuli are treated more fluently than positive stimuli with a comparable arousal value.

Note that, in practice, when different arguments are considered, they will more likely fall under the “mixed valence” condition. However, in some situations, we want to compare arguments that have the same meaning in order to choose which one is better phrased depending on the application and the persuadee. In this case, the arguments are only differing by a few words (synonyms). Therefore, the “only positive valence” and “only negative valence” conditions are important. Once the version of the argument (out of those with the same meaning) is chosen, it is then compared with the other unrelated arguments.

## 5.5 Selection Method

In order to compute an optimal policy we need to start from the leaves of the tree and recursively compute the Pareto optimal solutions in the decision nodes. However, a deterministic policy only gives one action to perform in each possible state. Therefore, when several incomparable Pareto optimal solutions are valid in a state, we need to carefully pick one.

This choice depends on the strategy the persuader wants to apply. Indeed, in a one-shot situation, when there is only one dialogue, the persuader may want to maximize the belief at all cost, meaning without any consideration to the emotions induced by the dialogue. In this case, we pick the argument with the maximum belief amongst the argument in the Pareto front. On the other hand, when several dialogues are planned (for instance, in a doctor/patient situation), the persuader might want to sacrifice the belief in order to leave a good impression in the first dialogues, in the hope of increasing the efficiency of the future dialogues.

## 6 Size-reducing Constraints

The low number of Pareto optimal solutions for each dialogue can be handled efficiently. However, the computational difficulty comes from the large number of possible dialogues. In the most general case where we allow the arguments to occur several times in the dialogue without a bounded horizon, this number is infinite. To reduce it, we consider three size-reducing constraints that can be independently used or combined:

**Constraint 1** no repetition of arguments in the dialogue,

**Constraint 2** no direct attacker of the goal is allowed as a proponent move,

**Constraint 3** only relevant arguments can be played.

A relevant argument is the first one of a dialogue or any argument connected to the ones already posited.

### 6.1 Theoretical Results

We now consider the theoretical benefits of size-reducing constraints by identifying the theoretical numbers of dialogues  $n$  depending on the different constraints applied. In the following,  $a$  is the number of arguments,  $h$  the horizon and  $k$  the number of direct attackers to the goal argument. We make the assumptions that  $h < k < a$ , for Propositions 1 to 5, and that  $h < a - k$ , for Proposition 3, to ensure that the dialogues are of size  $h$  in any situation. Without these assumptions the number of dialogues of size  $h$  is  $n' \leq n$ .

We start by giving the naive constraint which allows any argument to be used at each step of the dialogue.

**Proposition 1.** (*Naive constraint*) *In the general case, the number of dialogues is  $n = a^h$ , i.e.,  $a$  choices at each of the  $h$  steps.*

We consider constraint 1 and constraints 1 and 2 together.

**Proposition 2.** *(No repetition constraint) When arguments cannot be repeated in the dialogue,  $n = \frac{a!}{(a-h)!}$ , i.e.,  $a \times (a-1) \times \dots \times (a-h+1)$ .*

**Proposition 3.** *(No attacking arguments + no repetition) In this case, the proponent cannot play arguments that directly attack her goal argument:*

$$n = \sum_{i=0}^{\lfloor \frac{h}{2} \rfloor} \binom{\lfloor \frac{h}{2} \rfloor}{i} \times \frac{(a-k)!}{(a-k - \lfloor \frac{h}{2} \rfloor - i)!} \times \frac{k!}{(k - (\lfloor \frac{h}{2} \rfloor - i))!} .$$

*Proof.* (Outline) For horizon  $h = 3$ , the numbers of possible choices at each step are (depending on the opponent playing a direct attacker of the goal or not):

$$\begin{aligned} n &= (a-k) \times k \times (a-k-1) \\ &\quad + (a-k) \times (a-k-1) \times (a-k-2) \\ n &= 1 \times \frac{(a-k)!}{(a-k-2-0)!} \times \frac{k!}{(k-(1-0))!} \\ &\quad + 1 \times \frac{(a-k)!}{(a-k-2-1)!} \times \frac{k!}{(k-(1-1))!} \\ n &= \sum_{i=0}^1 \binom{1}{i} \times \frac{(a-k)!}{(a-k-2-i)!} \times \frac{k!}{(k-(1-i))!} . \end{aligned}$$

For horizon  $h = 4$ ,

$$\begin{aligned} n &= (a-k) \times (a-k-1) \times k \times (k-1) \\ &\quad + 2 \times (a-k) \times (a-k-1) \times (a-k-2) \times k \\ &\quad + (a-k) \times (a-k-1) \times (a-k-2) \times (a-k-3) \\ n &= 1 \times \frac{(a-k)!}{(a-k-2-0)!} \times \frac{k!}{(k-(2-0))!} \\ &\quad + 2 \times \frac{(a-k)!}{(a-k-2-1)!} \times \frac{k!}{(k-(2-1))!} \\ &\quad + 1 \times \frac{(a-k)!}{(a-k-2-2)!} \times \frac{k!}{(k-(2-2))!} \\ n &= \sum_{i=0}^2 \binom{2}{i} \times \frac{(a-k)!}{(a-k-2-i)!} \times \frac{k!}{(k-(2-i))!} . \end{aligned}$$

We can see the general expression from the examples.

The following result shows that we get improvement from the use of the constraints considered in the results above.

**Proposition 4.** *Given values for the number of arguments  $a$ , the horizon  $h$ , and the number of direct attackers to the goal argument  $k$ , if  $n_1$  is the number of dialogues obtained by proposition 1,  $n_2$  is the number of dialogues obtained by proposition 2, and  $n_3$  is the number of dialogues obtained by proposition 3, then  $n_1 \geq n_2$  and  $n_2 \geq n_3$ .*

In Table 2, we show the percentage of reduction in using the constraints as calculated by Propositions 1, 2, and 3.

**Table 2.** The fraction of all dialogues (as calculated by Proposition 1) obtained by using constraint 1 (*i.e.*, no repetition constraint) or constraint 1 and 2 (*i.e.*, no attacking arguments and no repetition), and where  $h = 10$  and  $k = 15$ .

Constraint	Number of arguments				
	30	40	50	60	70
1	0.184	0.293	0.381	0.452	0.509
1 + 2	0.004	0.023	0.058	0.101	0.146

**Table 3.** Number of nodes for a 7-argument graph

# arcs removed	0	5	10	15
# nodes	2372	1182	319	28

We now consider the third size-reducing constraint which states that only relevant arguments can be played. The number of dialogues using this constraint is highly dependent on the structure of the graph, and so we proceed by considering the worst case scenario, which is when the graph is a complete graph, and best case scenario, which is when the graph is a chain of arguments.

**Proposition 5.** (*Relevant arguments constraint*) (*Worst case*) *The number of dialogues is the same as in Proposition 2 (if we consider the no repetition constraint at the same time).* (*Best case*) *The number of dialogues is  $n = a - (h - 1)$ , i.e., one possible dialogue for each possible starting point leading to a sequence of length  $h$ .*

Table 3 shows the number of nodes in the decision tree generated from a 7-argument graph, with the “relevant arguments” and “no repetition” constraints applied. The top values represent the number of arcs removed from the complete graph where 0 denotes the complete graph and 15 a chain. We can see that the number of nodes in the tree exponentially decreases. Also note that the “no goal direct attackers” constraint is not applied here as it is trivial in the case of the complete graph. Indeed, in this case the only dialogues are of size 2, the

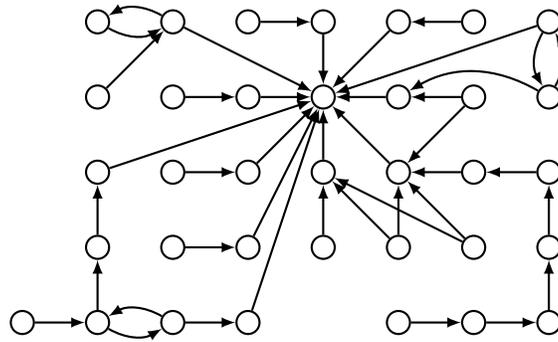
goal argument for the proponent and any of the remaining arguments for the opponent. The proponent cannot play at step 3 without attacking her goal.

## 7 Computational Evaluation

In order to study the computational efficiency of our method, we use an argument graph on the topic of the annual flu vaccination for hospital staff, developed with healthcare professionals. It contains 35 arguments including a persuasion goal for taking the vaccine and various counterarguments. Some examples of arguments are: “NHS staff can transmit infections to the patients.” (pro-vaccine), “The flu vaccine weakens the immunitary system.” (anti-vaccine) and “The flu vaccine is useless because the strain is guessed.” (anti-vaccine).

Note that, the creation of the argument graph is context dependent. In some cases, information is easily available and in others we need to rely on experts. Detailed explanations on how we can model a domain for behaviour change applications can be found in [38].

We also add synonym arguments (arguments with same meaning but different wording) for 15 of the arguments in the graph, carrying the same meaning but using different words and so different values for the affective norm. The final graph has thus 50 arguments<sup>3</sup>. Figure 2 depicts the graph without the synonym arguments. We can see that the graph is not a tree, it contains cycles and multiple paths from nodes to the root, making it a non-trivial example. The experiments have been run on an Intel i5-6600 at 3.30GHz with 8GB of RAM.



**Fig. 2.** Argument graph without synonym arguments

Using different affective norm values gives, as expected, different optimal sequences of arguments for each player. For instance, using the norm associated with older people yields the sequence with the following argument names: “annualvaccine notconcerned strongsense nocare cantransmit dontworry nodirect”

<sup>3</sup> The code, the graph and the mapping to the actual arguments can be found at <https://github.com/ComputationalPersuasion/stardec>.

**Table 4.** Computation time for horizons 1 to 8 (in sec)

1	2	3	4	5	6	7	8
0.08	0.08	0.08	0.12	0.26	4	28	597

while using the norm associated to younger people gives: “annualvaccine nocare strongsense dontworry caninfect noface2face nodirect”.

Table 4 shows the computation time from the creation of the tree to the computation of an optimal policy, with constraints 1 to 3 enabled, for horizons from 1 to 8. As we can see, the time grows exponentially with the horizon. However, we argue that 8 is a good length for the horizon as it allows several exchanges of arguments between both agents without incurring a too high risk of disengagement from the persuadee. Note that most of the time is spent on creating the tree. Using methods such as Branch-and-bound can highly improve the computation time.

## 8 Study with Participants

The aim of this empirical study is to investigate the computational viability of evaluating emotion and belief associated with an argument by an opponent and using this to determine the best policy for the proponent. The purpose is also to validate the choices of decision rules (described in Table 1) when instantiating with the psychology literature.

### 8.1 Preliminary Experimentation

As a first step, we chose a set of 13 arguments amongst the 35 from the same graph as in Section 7. Some are pro arguments the other are counterarguments. We crowdsourced the creation of synonyms. Using the Crowdfunder<sup>4</sup> platform, we gave all 13 arguments and asked for one or more synonyms for a given word in the argument. This was run using 25 participants, with a sufficient knowledge of the English language. After cleaning the results (removing aberrant answers such as “planetary system” when asking for a synonym of “mercury” in a vaccine, or answers in a different language), we ended up with between 2 and 7 alternative words, depending on the argument.

### 8.2 Participants, Material and Procedure

We recruited 100 participants via the Prolific<sup>5</sup> platform. The only screening criterion we used was the language of participants which we limited to English. We presented each of the 13 arguments considered in the preliminary experiment

<sup>4</sup> <https://www.crowdfunder.com>

<sup>5</sup> <https://www.prolific.ac>

to each participant. For each argument, we gave the synonyms obtained in the preliminary experiment in the form of a menu, and asked the participant to select the synonym she deems the best at “conveying its message”. We were therefore asking each participant which was the best version of each argument.

### 8.3 Results

Before collating the results, we removed some arguments and participants from further consideration as follows. We removed two arguments where the choice of word was highly constrained by the domain (“complications” cannot be replaced with *e.g.*, “problems” when talking about unfortunate consequences of a vaccine, and the word “injection” is almost never used when “vaccine” can be picked). We removed the 12 participants choosing the wrong answer for these two arguments. We also removed a third argument because the set of synonyms was almost the same as for another argument, leaving us with 10 arguments.

We applied our set of rules to the same set of arguments with the same sets of synonyms. We compared the results with the answers given by the participants. Our set of rules (Table 1) gives us the most chosen word for 8 arguments out of 10 and the second most chosen for the remaining 2. It means that these rules (the second half of Table 1) approximate efficiently the relation the majority of our participants have with emotions induced by the arguments.

**Table 5.** Cumulative score for the prediction of the answers

9	7	6	5	3	1
1.1%	13.2%	36.3%	52.7%	95.6%	100%

Note that while we accurately represent the choice of the majority for each argument, only a few participants are part of the majority for all arguments at the same time. Table 5 shows the cumulative sum of good predictions when analyzing the individual participant’s answers. Each percentage is the percentage of the participants we have managed to predict correctly at least  $x$  answers. For instance, for  $x = 9$ , 1.1% means that we have correctly predicted at least 9 choices out of 10 for 1.1% of the participants. Interestingly, our method predicted correctly more than half the answers for more than half of the population: 52.7% for 5 or more answers.

This user experiment has two conclusions. First it shows that the psychology literature reflects the choices made by the participants. Second it demonstrates that we can transform the psychological principles into logical and mathematical formulae and predict the choices with good results.

## 9 Conclusion

In this paper, we have shown how decision-theoretic methods for multi-criteria decision making can be used to identify an optimal policy for dialogical argumentation. We have presented a framework for modelling a proponent and opponent in a dialogue, and for handling multiple criteria such as the degree the opponent believes an argument, and the degree of valence and arousal, that could be evoked by the argument in the opponent. Our hypotheses are supported by the psychological literature and backed by user experiments. These are only indicative of the dimensions that might be used in an application. In future work, we intend to develop the treatment of emotion in the framework as the topic has received little attention in the computational argumentation literature. Exceptions are [39] which provide rules for specifying scenarios where empathy is given or received, and [40] which investigates relationships between emotions that participants feel during a debate (measured physiologically) and arguments. In contrast, affective computing has put emotion at the centre of the relationship between users and computing systems [41].

## Acknowledgements

This research is part funded by EPSRC Project EP/N008294/1 (Framework for Computational Persuasion).

## References

1. Prakken, H.: Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation* **15**(6) (2005) 1009–1040
2. Prakken, H.: Formal systems for persuasion dialogue. *Knowledge Engineering Review* **21**(2) (2006) 163–188
3. Fan, X., Toni, F.: Assumption-based argumentation dialogues. In: *Proc IJCAI'11*. (2011) 198–203
4. Caminada, M., Podlaskowski, M.: Grounded semantics as persuasion dialogue. In: *Proc. COMMA'12*. (2012) 478–485
5. Thimm, M.: Strategic argumentation in multi-agent systems. *Kunstliche Intelligenz* **28** (2014) 159–168
6. Rahwan, I., Larson, K.: Pareto optimality in abstract argumentation. In: *Proc. AAAI'08*. (2008) 150–155
7. Fan, X., Toni, F.: Mechanism design for argumentation-based persuasion. In: *Proc. COMMA'12*. (2012) 322–333
8. Rienstra, T., Thimm, M., Oren, N.: Opponent models with uncertainty for strategic argumentation. In: *Proc. IJCAI'13*. (2013) 332–338
9. Hadjinikolis, C., Siantos, Y., Modgil, S., Black, E., McBurney, P.: Opponent modelling in persuasion dialogues. In: *Proc. IJCAI'13*. (2013) 164–170
10. Hadoux, E., Beynier, A., Maudet, N., Weng, P., Hunter, A.: Optimization of probabilistic argumentation with Markov decision models. In: *Proc. IJCAI'15*. (2015) 2004–2010

11. Black, E., Coles, A., Bernardini, S.: Automated planning of simple persuasion dialogues. In: Proc. CLIMA'14. Volume 8624 of LNCS., Springer (2014) 87–104
12. Hadoux, E., Hunter, A.: Strategic sequences of arguments for persuasion using decision trees. In: Proc. AAAI'17. (2017)
13. Dung, P.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. *Artificial Intelligence* **77** (1995) 321–357
14. Thimm, M.: A probabilistic semantics for abstract argumentation. In: Proc. ECAI'12. (2012) 750–755
15. Hunter, A.: A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning* **54**(1) (2013) 47–81
16. Hunter, A., Thimm, M.: Probabilistic argumentation with incomplete information. In: Proc. ECAI'14. (2014) 1033–1034
17. Baroni, P., Giacomin, M., Vicig, P.: On rationality conditions for epistemic probabilities in abstract argumentation. In: Proc. COMMA'14. (2014) 121–132
18. Lazarus, R.S.: Progress on a cognitive-motivational-relational theory of emotion. *American psychologist* **46**(8) (1991) 819
19. Scherer, R., K.: Appraisal considered as a process of multilevel sequential checking. *Appraisal processes in emotion: Theory, methods, reasearch.* **92**(120) (2001) 57
20. Duhachek, A., Agrawal, N., Han, D.: Guilt versus shame: coping, fluency, and framing in the effectiveness of responsible drinking messages. *Journal of Marketing Research* **49**(6) (2012) 928–941
21. Tversky, A., Kahneman, D.: The framing of decisions and the psychology of choice. *Science* **211**(4481) (1981) 453–458
22. Ekman, P.: An argument for basic emotions. *Cognition & emotion* **6**(3-4) (1992) 169–200
23. Fulladoza Dalibón, S., Martinez, D., Simari, G.: Emotion-directed argument awareness for autonomous agent reasoning. *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial* **15**(50) (2012)
24. Lloyd-Kelly, M., Wyner, A.: Arguing about emotion. In: *International Conference on User Modeling, Adaptation, and Personalization*, Springer (2011) 355–367
25. Nawwab, F., Dunne, P., Bench-Capon, T.: Exploring the role of emotions in rational decision making. In: COMMA. (2010) 367–378
26. Bradley, M., Lang, P.: Affective norms for english words (ANEW): Instruction manual and affective ratings. Technical report, The Center for Research in Psychophysiology, University of Florida. (1999)
27. Warriner, A., Kuperman, V., Brysbaert, M.: Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior Research Methods* **45**(4) (2013) 11911207
28. Eder, A.B., Rothermund, K.: Automatic influence of arousal information on evaluative processing: Valence-arousal interactions in an affective Simon task. *Cognition & Emotion* **24**(6) (September 2010) 1053–1061
29. Jefferies, L.N., Smilek, D., Eich, E., Enns, J.T.: Emotional valence and arousal interact in attentional control. *Psychological Science* **19**(3) (2008) 290–295
30. Robinson, M.D.: Watch out! that could be dangerous: Valence-arousal interactions in evaluative processing. *Personality and Social Psychology Bulletin* **30**(11) (November 2004) 1472–1484
31. Cayrol, C., Lagasque-Schiex, M.: Gradual valuation for bipolar argumentation frameworks. In: Proc. ECSQARU'05, Springer (2005) 366–377
32. Amgoud, L., Ben-Naim, J.: Axiomatic foundations of acceptability semantics. In: Proc. KR'16. (2016)

33. Bonzon, E., Delobelle, J., Konieczny, S., Maudet, N.: A comparative study of ranking-based semantics for abstract argumentation. In: Proc. AAI'16. (2016) 914–920
34. Hunter, A.: Modelling the persuadee in asymmetric argumentation dialogues for persuasion. In: Proc. IJCAI'15. (2015) 3055–3061
35. Mohammad, S.: Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In: Emotion Management. Elsevier (2016) 201–238
36. Peeters, G., Czapinski, J.: Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European review of social psychology* **1**(1) (1990) 33–60
37. Baumeister, R.F., Bratslavsky, E., Finkenauer, C., Vohs, K.D.: Bad is stronger than good. *Review of General Psychology* **5**(4) (2001) 323–370
38. Chalaguine, L., Hadoux, E., Hamilton, F., Hayward, A., Hunter, A., Polberg, S., Potts, H.W.W.: Domain modelling in computational persuasion for behaviour change in healthcare. CoRR [abs/1802.10054](https://arxiv.org/abs/1802.10054) (2018)
39. Martinovski, B., Mao, W.: Emotion as an argumentation engine: Modeling the role of emotion in negotiation. *Group Decision and Negotiation* **18** (2009) 235–259
40. Benlamine, S., Chaouachi, M., Villata, S., Cabrio, E., Gandon, C.F.F.: Emotions in argumentation: an empirical evaluation. In: Proc. IJCAI'15. (2015) 156–163
41. Calvo, R., D'Mello, S.: Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Aff. Computing* **1**(1) (2010) 18–37