

# Optimization of Dialectical Outcomes in Dialogical Argumentation\*

Anthony Hunter<sup>†</sup> and Matthias Thimm<sup>‡</sup>

July 13, 2016

## Abstract

When informal arguments are presented, there may be imprecision in the language used, and so the audience may be uncertain as to the structure of the argument graph as intended by the presenter of the arguments. For a presenter of arguments, it is useful to know the audience's argument graph, but the presenter may be uncertain as to the structure of it. To model the uncertainty as to the structure of the argument graph in situations such as these, we can use probabilistic argument graphs. The set of subgraphs of an argument graph is a sample space. A probability value is assigned to each subgraph such that the sum is 1, thereby reflecting the uncertainty over which is the actual subgraph. We can then determine the probability that a particular set of arguments is included or excluded from an extension according to a particular Dung semantics. We represent and reason with extensions from a graph and from its subgraphs, using a *logic of dialectical outcomes* that we present. We harness this to define the notion of an *argumentation lottery*, which can be used by the audience to determine the expected utility of a debate, and can be used by the presenter to decide which arguments to present by choosing those that maximize expected utility. We investigate some of the options for using argumentation lotteries, and provide a computational evaluation.

## 1 Introduction

Computational models of argument aim to reflect how human argumentation uses conflicting information to construct and analyze arguments. There is a number of frameworks for computational models of argumentation. They incorporate a formal representation of individual arguments and techniques for comparing conflicting arguments (for reviews see [BCD07, BH08]).

---

\*This is an extended version of A Hunter and M Thimm (2014): Probabilistic Argument Graphs for Argumentation Lotteries. Computational Models of Argument (COMMA'14), IOS Press.

<sup>†</sup>Department of Computer Science, University College London, UK

<sup>‡</sup>Institute for Web Science and Technology, University of Koblenz-Landau, Germany

In abstract argumentation, a graph is used to represent a set of arguments and counterarguments. Each node is an argument and each arc from an argument  $\alpha$  to an argument  $\beta$  denotes an attack by  $\alpha$  on  $\beta$ . It is a well-established and intuitive approach to modelling argumentation, and it offers a valuable starting point for theoretical analysis of argumentation [Dun95]. However, including an argument  $\alpha$  in a graph usually means that one is sure that  $\alpha$  is a *justifiable* argument, i. e., that it is an argument that makes sense (independently of whether it can be *accepted* after relating it to other arguments). Abstract argumentation does not explicitly consider whether (or to what degree) an argument is believed to be justifiable or whether (or to what degree) an attack by an argument is believed to be justifiable. It only represents the existence of arguments and counterarguments in a strict manner.

To address the need to represent and reason with (quantified) uncertainty, it has been proposed to use a probability assignment to arguments and to attacks [BGW05]. This can be used to give a probability distribution over the subgraphs of the argument graph, and this can then be used to give a probability assignment for a set of arguments being an admissible set or extension of the argument graph [LON11, Hun12, Hun13c]. The probability distribution over subgraphs denotes the uncertainty over which subgraph is the actual graph that should be used. We refer to an argument graph with a probability distribution over subgraphs as a *probabilistic argument graph*.

Consider a typical argumentation scenario for persuasion [WK95, Ben03], i. e., a scenario where one or more agents are presenting arguments in front of an audience, with the aim of each participant being to persuade the audience to adopt a certain statement. We assume that each participant and the audience have some argument graph in mind but are willing to incorporate new arguments and attacks into it. In this context, we believe the following are two important applications for probabilistic argument graphs:

- **From an audience's perspective**, there may be uncertainty as to what the actual argument graph is. The audience may hear various comments in a debate, for example, but they are not sure about the exact set of arguments and attacks that are being put forward. For instance, there may be uncertainty about whether someone has put forward a complex multifaceted argument, or a number of smaller more focused arguments or there may doubt about whether some arguments are just rephrasings of previous arguments. There may be uncertainty about which arguments are meant to be attacked by some argument, which occurs frequently when enthymemes (incomplete arguments) are presented. So the audience can collate all the candidates for arguments and attacks, and construct the graph containing them all, and then identify a probability distribution over its subgraphs that reflects their uncertainty about which is the actual graph.
- **From a participant's perspective** (i. e. from the perspective of someone who is about to present arguments and/or attacks to some monological or dialogical argumentation), there may be uncertainty about what the audience regards as the argument graph. When a participant (such as a politician) considers presenting arguments to an audience, the participant might not know for sure which arguments and attacks the audience has in mind. In other words, even before a participant has started, the audience may already have an argument graph in

mind and the participant will be adding to that graph in the audience’s mind. To handle this, the participant may have an argument graph which he/she assumes will subsume the possibilities for the argument graph held by the audience, and then the participant might identify a probability distribution over subgraphs of the argument graph to reflect the uncertainty as judged by the participant over which is the subgraph being used by the audience.

In this paper we investigate the use of probabilistic argument graphs as the underlying knowledge representation formalism in dialogical argumentation scenarios. Besides the rich expressivity of probabilistic argument graphs, this formalism can also be used for the problem of rational action selection, i. e., which arguments to disclose in order to maximize the utility of the result of the dialogue. As we will see, we can utilize probabilistic argument graphs to determine the probability of possible outcomes of an argument graph.

We define these outcomes as formulae of an expressive logic of dialectical outcomes for reasoning about subgraphs and extensions of argument graphs. This logic allows the representation of complex statements about argument graphs—such as “there is a subgraph of the graph where all preferred extensions contain either  $\alpha$  or  $\beta$ ”—and can thus be used to represent desired outcomes and means to reach them. For probabilistic argument graphs, we can determine the probability of those formulae. From an audience’s perspective, this gives a better understanding of the consequences of the debate that they are observing, and from a participant’s perspective, it gives a better understanding of whether s/he will get the desired outcomes from his/her contributions to the argumentation.

We further exploit probabilistic argument graphs, by introducing the notion of *lotteries* for argumentation. Lotteries are an important approach to decision-making with uncertainty. In a lottery, there are a number of outcomes, and probability associated with each of them. For example, if we buy a lottery ticket for 1 Euro, with the prize being 500 Euros, and there are 1000 tickets, then we have the outcome “win” with the probability  $1/1000$  and the outcome “lose” with a probability of  $999/1000$ . We can then measure the utility of each outcome. For instance, the utility of “win” could be 500 for the prize minus 1 for the cost of entering (i. e. net utility is 499), and the utility of “lose” is -1. The expected utility of buying the ticket is then  $(499 \times 1/1000) + (-1 \times 999/1000)$  is  $-1/2$  Euro, whereas the expected utility of not buying the ticket is 0 Euro, which suggests it is not a good decision to buy the ticket.

Assume that during a discussion, a debater wants to identify a good argument to bring into the discussion and that the audience of the discussion is considering some subgraph of  $G$  as the *true* argument graph. The debater does not know for sure which subgraph is the correct one but he can identify a probability distribution over the subgraphs. Now, suppose he is keen that arguments  $\alpha$  and  $\beta$  are accepted by the audience (e. g. they are both in the grounded extension of whichever subgraph the audience is using). So the outcome we want is that  $\alpha$  and  $\beta$  are included in the grounded extension. If this is not possible, then perhaps he wants the outcome where  $\alpha$  is included and  $\beta$  excluded. Suppose any other outcome is inferior to these two outcomes. By using the probabilistic argument graphs, we are able to determine a probability for each of these outcomes, and we can construct a lottery containing these arguments. If we identify a

utility function over outcomes, we can apply utility theory to determine the expected utility. Furthermore, if we then consider further arguments that we can add to the discussion, we can evaluate the expected utility of each choice of further arguments to put forward. We can then determine which actions (i. e. which arguments to add to the discussion) will offer the maximum expected utility.

One further issue in the above outlined approach is *how* to adapt a probabilistic model when new information—that is, additional arguments and attacks—becomes available. The basic approach for doing this is Bayesian conditioning, as it is usually done in approaches to probabilistic reasoning [Pea88]. For example, if  $P$  is a probability distribution over the possible graphs the audience has in mind, we expect that after disclosing a set of arguments and attacks  $e$  the new probability distribution is  $P(\cdot | e)$ . In particular, all argument graphs that do not contain  $e$  will receive probability zero after  $e$  has been disclosed. This approach assumes that the audience accepts disclosed arguments and attacks. However, this is not necessarily the case as discussed above. The audience may not fully accept all proposed arguments or attacks and may simply ignore parts of the disclosed information. This issue has also been discussed within the field of belief revision [Han01] under the term of *non-prioritized belief revision* [Han99]. In order to address this issue for the purpose of selecting the best move to make, we also introduce two further approaches to update a probabilistic model on an argument graph that take this uncertainty into account. In order to analyze the behaviour of these different update approaches and the general approach of argumentation lotteries for move selection in argumentation dialogues, we also report on an experimental analysis that compares these approaches in actual argumentation dialogues.

In summary, the aim of this paper is to develop the use of probabilistic argument graphs and to apply them to argumentation lotteries. For this, we make the following contributions:

1. we introduce a logic of dialectical outcomes for representing and reasoning with outcomes that hold for an argument graph (Section 3);
2. we adapt previous works on probabilistic argument graphs and extend them to our logic of dialectical outcomes (Section 4);
3. we introduce the notion of an argumentation lottery, which we show can be used to determine the expected utility of outcomes of a probabilistic argument graph (Section 5);
4. we show how argumentation lotteries can be used in dialogical argumentation and, in particular, for enabling an agent to determine what would be the best contribution to make in monological or dialogical argumentation in order to maximize its expected utility; moreover, we provide three concrete approaches how probabilistic argument graphs can be updated to incorporate new information (Section 6); and
5. we conduct an extensive empirical evaluation in order to analyze the usefulness of argumentation lotteries for action selection in dialogical argumentation (Section 7).

We introduce some necessary preliminaries in Section 2, discuss related work in Section 8, and conclude in Section 9.

This paper extends [HT14a] by introducing the notion of a logic of dialectical outcomes (which we use to give a much more expressive way of defining outcomes in lotteries), by introducing the notion of redistribution which takes into account how the probability distribution needs to be defined when considering possible actions (which is necessary if we are to have a full account of how lotteries can be used to make rational choices of what argumentation move to make), and by undertaking an empirical evaluation.

## 2 Preliminaries

An **abstract argument graph**  $G$  is a pair  $G = (\mathcal{A}, \mathcal{R})$  where  $\mathcal{A}$  is a set and  $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$  [Dun95]. Each element  $\alpha \in \mathcal{A}$  is called an **argument** and  $(\alpha, \beta) \in \mathcal{R}$  means that  $\alpha$  **attacks**  $\beta$  (accordingly,  $\alpha$  is said to be an **attacker** or a **counterargument** for  $\beta$ ). A set of arguments  $S \subseteq \mathcal{A}$  **attacks**  $\beta \in \mathcal{A}$  iff there is an argument  $\alpha \in S$  such that  $\alpha$  attacks  $\beta$ . Also,  $S$  **defends**  $\alpha' \in \mathcal{A}$  iff for each argument  $\beta \in \mathcal{A}$ , if  $\beta$  attacks  $\alpha'$  then  $S$  attacks  $\beta$ . A set  $S \subseteq \mathcal{A}$  of arguments is **conflict-free** iff there are no arguments  $\alpha, \alpha' \in S$  such that  $\alpha$  attacks  $\alpha'$ . Let  $\wp(X)$  be the power set of a set  $X$ . Let  $\Gamma$  be a conflict-free set of arguments, and let  $\text{Defended}_{(\mathcal{A}, \mathcal{R})} : \wp(\mathcal{A}) \rightarrow \wp(\mathcal{A})$  be a function such that  $\text{Defended}_{(\mathcal{A}, \mathcal{R})}(\Gamma) = \{\alpha \mid \Gamma \text{ defends } \alpha\}$ .

Semantics are given to abstract argument graphs by *extensions*, i. e., sets of arguments that are considered to be jointly acceptable. We consider the following types of extensions (let  $\Gamma \subseteq \mathcal{A}$ )

- $\Gamma$  is a **complete extension** (co) iff  $\Gamma = \text{Defended}_{(\mathcal{A}, \mathcal{R})}(\Gamma)$ ,
- $\Gamma$  is a **grounded extension** (gr) iff it is the (uniquely determined) minimal (wrt. set inclusion) complete extension,
- $\Gamma$  is a **preferred extension** (pr) iff it is a maximal (wrt. set inclusion) complete extension, and
- $\Gamma$  is a **stable extension** (st) iff it is a preferred extension such that  $\Gamma$  attacks  $\beta$  for each argument  $\beta \in \Gamma \setminus \mathcal{A}$ .

For  $G = (\mathcal{A}, \mathcal{R})$ , let  $\text{Extensions}_X(G)$  be the set of extensions of  $G$  according to semantics  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ .

In order to present our framework of probabilistic argument graphs we need to introduce some notions for subgraphs of an argument graph. Let  $\mathcal{R} \otimes \mathcal{A}'$  be the subset of  $\mathcal{R}$  involving just the arguments in  $\mathcal{A}' \subseteq \mathcal{A}$ , i. e.,  $\mathcal{R} \otimes \mathcal{A}' = \{(\alpha, \beta) \in \mathcal{R} \mid \alpha, \beta \in \mathcal{A}'\}$ . Also let  $G_\emptyset$  denote the empty graph. For argument graphs  $G = (\mathcal{A}, \mathcal{R})$  and  $G' = (\mathcal{A}', \mathcal{R}')$  we say that  $G'$  is a **subgraph** of  $G$ , denoted  $G' \sqsubseteq G$ , iff  $\mathcal{A}' \subseteq \mathcal{A}$  and  $\mathcal{R}' \subseteq \mathcal{R} \otimes \mathcal{A}'$ . Note that  $G$  is also a subgraph of itself. For any argument graph  $G$ , let  $\text{Sub}(G)$  denote the set of subgraphs of  $G$ , i. e.,  $\text{Sub}(G) = \{G' \mid G' \sqsubseteq G\}$ . Also for any argument graph  $G = (\mathcal{A}, \mathcal{R})$ , let  $\text{Nodes}(G) = \mathcal{A}$  and  $\text{Arcs}(G) = \mathcal{R}$ . Furthermore, for any argument graph  $G$ , let  $\text{SpanningSub}(G) \subseteq \text{Sub}(G)$  denote the set of spanning

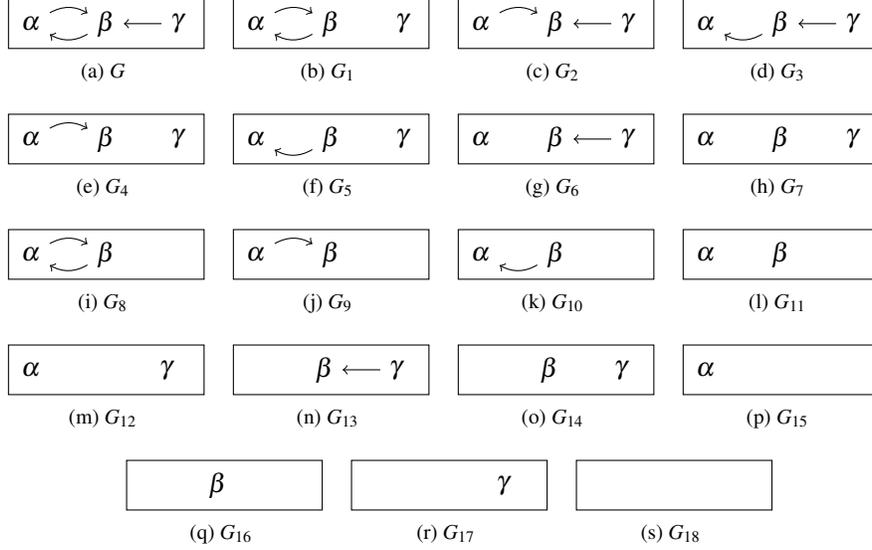


Figure 1: A simple argument graph  $G$  and its subgraphs.

subgraphs of  $G$  (graphs that contain a subset of nodes of  $G$  but all attacks on this subset), i. e.  $\text{SpanningSub}(G) = \{G' \in \text{Sub}(G) \mid \text{Arcs}(G') = \text{Arcs}(G) \otimes \text{Nodes}(G')\}$ .

**Example 1.** Consider the argument graph  $G$  depicted in Figure 1a and its subgraphs depicted in Figures 1a to 1s. Here we have  $G = (\mathcal{A}, \mathcal{R})$  with

$$\begin{aligned}\mathcal{A} &= \{\alpha, \beta, \gamma\} \\ \mathcal{R} &= \{(\alpha, \beta), (\beta, \alpha), (\gamma, \beta)\}\end{aligned}$$

and  $\text{Sub}(G) = \{G, G_1, \dots, G_{18}\}$  and  $\text{SpanningSub}(G) = \{G, G_8, G_{12}, G_{13}, G_{15}, G_{16}, G_{17}, G_{18}\}$ . We will use  $G$  throughout the paper as a running example.

In the following, we will use the subgraphs of a graph to model uncertainty in the original graph.

### 3 Logic for dialectical outcomes

We may view the construction of an argument graph as an important part of the process of argumentation. This may be a graph that all agents see or it may be a graph in the mind of one agent that is used to model the arguments of another agent. However, when we construct the argument graph, it is useful to consider the outcomes of the argument graph. These outcomes, which we call dialectical outcomes, may include what arguments are included or excluded from the extensions of the graph and of the subgraphs according to some dialectical semantics.

Once we consider dialectical outcomes, we may want a language to specify them. This is the aim of *the logic of dialectical outcomes* which we introduce in this section. We present it as a logic so that we have a precise semantics for it. This means we know exactly what is meant by any statement of the language. Furthermore, we can reason with statements of the language using the definition of the semantics in order to answer questions such as “is this set of statements consistent” or “does this set of statements imply another statement”. As with some other logics such as some many-valued logics, the logic can be useful for representation and reasoning without having a proof theory though we may develop a proof theory in future work.

The reason we introduce the logic of dialectical outcomes in this paper is that we can use it to specify outcomes in the argumentation lottery. So we can specify desirable and undesirable outcomes of an argument graph in this language, and these are the outcomes we consider in a lottery. Part of the role of the logic is to ensure that the outcomes in the lottery are disjoint and exhaustive.

Some of the ideas behind the logic of dialectical outcomes are inspired by similar approaches to argumentation logics such as e. g. [BKRvdT13]. We undertake a comparison with the literature in Section 8. In the following presentation of the logic of dialectical outcomes, let  $X$  be some fixed argumentation semantics,  $\Pi$  a set of arguments, and  $G$  be a fixed argument graph with  $\text{Nodes}(G) \subseteq \Pi$ .

**Definition 1.** *Let  $\Pi$  be a set of arguments. Let  $\mathcal{L}_\Pi$  be the modal language defined as follows:  $\phi \in \mathcal{L}_\Pi$  iff it is defined via the following BNF*

$$\phi ::= \alpha \mid \neg\phi \mid \phi \vee \phi \mid \phi \wedge \phi \mid \Box\phi \mid \Diamond\phi \mid \blacksquare_{S'}^S\phi \mid \blacklozenge_{S'}^S\phi$$

for all  $\alpha \in \Pi$  and  $S, S' \subseteq \Pi$  with  $S' \subseteq S$ . If  $S' = \emptyset$  we also write  $\blacksquare^S$  ( $\blacklozenge^S$ ) instead of  $\blacksquare_{S'}^S$  ( $\blacklozenge_{S'}^S$ ). If  $S = \Pi$  we also write  $\blacksquare$  ( $\blacklozenge$ ) instead of  $\blacksquare_{S'}^S$  ( $\blacklozenge_{S'}^S$ ). If both  $S' = \emptyset$  and  $S = \Pi$  we simply write  $\blacksquare$  ( $\blacklozenge$ ) instead of  $\blacksquare_{S'}^S$  ( $\blacklozenge_{S'}^S$ ).

This language is based on each atom representing an argument. We allow Boolean combinations of arguments using negation, disjunction, and conjunction. We also introduce modal operators where the white operators (i. e.  $\Box$  and  $\Diamond$ ) are concerned with making statements about extensions of graphs whereas the black operators (i. e.  $\blacksquare$  and  $\blacklozenge$  with or without sub- or super-script) are concerned with making statements about subgraphs. The sub- and super-scripts of  $\blacksquare_{S'}^S$  and  $\blacklozenge_{S'}^S$  are used to constrain the subgraphs considered in the nested expression. More precisely, only subgraphs containing the nodes in  $S$  but containing no nodes outside of  $S'$  are to be considered. Diamond-shaped operators ( $\Diamond$  and  $\blacklozenge$ ) are used for existential quantification (“There is some extension/subgraph that satisfies the nested formula”) while box-shaped operators ( $\Box$  and  $\blacksquare$ ) are used for universal quantification (“All extensions/subgraphs satisfy the nested formula”). We formalize the meaning of these formulae in the next subsection.

### 3.1 Semantics for the logic of dialectical outcomes

In order to introduce the semantics in a general way, we require the following subsidiary definition. Let  $\text{Graphs}(\Pi)$  be the set of graphs that can be constructed from the

arguments in  $\Pi$ :

$$\text{Graphs}(\Pi) = \{G \mid \text{Nodes}(G) \subseteq \Pi \text{ and } \text{Arcs}(G) \subseteq \text{Nodes}(G) \times \text{Nodes}(G)\}$$

The graph  $G$  is the **universal graph** for  $\Pi$  iff  $\text{Nodes}(G) = \Pi$  and  $\text{Arcs}(G) = \Pi \times \Pi$ . So if  $G$  is the universal graph for  $\Pi$ , then  $\text{Graphs}(\Pi) = \text{Sub}(G)$ .

The semantics for our logic is based on interpretations of the form  $(G, E)$  where  $G$  is an argument graph and  $E$  is a set of arguments.

**Definition 2.** *The set of interpretations for the language  $\mathcal{L}_\Pi$  is as follows*

$$\text{Interpretations}(\mathcal{L}_\Pi) = \{(G, E) \mid G \in \text{Graphs}(\Pi), E \subseteq \Pi\}$$

We proceed by defining the satisfaction relation for the logic.

**Definition 3.** *Let  $(G, E) \in \text{Interpretations}(\mathcal{L}_\Pi)$  and  $\psi, \phi \in \mathcal{L}_\Pi$ . The satisfaction relation for a semantics  $X$ , denoted  $\models_X$ , is defined as follows:*

$$\begin{aligned} (G, E) \models_X \alpha & \text{ iff } \alpha \in E \\ (G, E) \models_X \Box \phi & \text{ iff for all } E' \in \text{Extensions}_X(G) \text{ we have } (G, E') \models_X \phi \\ (G, E) \models_X \Diamond \phi & \text{ iff for some } E' \in \text{Extensions}_X(G) \text{ we have } (G, E') \models_X \phi \\ (G, E) \models_X \blacksquare_{S'}^S \phi & \text{ iff for all } G' \sqsubseteq G \text{ with } S' \subseteq \text{Nodes}(G') \subseteq S \text{ we have } (G', E) \models_X \phi \\ (G, E) \models_X \blacklozenge_{S'}^S \phi & \text{ iff for some } G' \sqsubseteq G \text{ with } S' \subseteq \text{Nodes}(G') \subseteq S \text{ we have } (G', E) \models_X \phi \\ (G, E) \models_X \phi \wedge \psi & \text{ iff } G, E \models_X \phi \text{ and } (G, E) \models_X \psi \\ (G, E) \models_X \phi \vee \psi & \text{ iff } G, E \models_X \phi \text{ or } (G, E) \models_X \psi \\ (G, E) \models_X \neg \phi & \text{ iff } G, E \not\models_X \phi \end{aligned}$$

Furthermore define

$$G \models_X \phi \text{ iff for all } E \in \text{Extensions}_X(G) \text{ we have } (G, E) \models_X \phi$$

For  $\psi, \phi \in \mathcal{L}_\Pi$  we also define  $\phi \models_X \psi$  iff for all  $G$  we have  $G \models_X \phi$  implies  $G \models_X \psi$ . We also write  $\phi \equiv_X \psi$  if both  $\phi \models_X \psi$  and  $\psi \models_X \phi$ . We denote a tautology by  $\top$ , i. e.  $G \models_X \top$  for every  $G$ , and contradiction by  $\perp$ , i. e.  $G \not\models_X \perp$  for every  $G$ .

**Example 2.** *We continue Example 1. The following are valid inferences wrt. our logic of dialectical outcomes:*

- $G \models_{\text{gr}} \alpha$  ( $\alpha$  is included in the grounded extension of  $G$ )
- $G \models_{\text{st}} \blacklozenge \Box \alpha$  (there is a subgraph  $G'$  of  $G$  such that all stable extensions of  $G'$  include  $\alpha$ )
- $G \models_{\text{st}} \neg \blacksquare_{\{\alpha\}} \Box \alpha$  (it is not the case that for all subgraphs  $G'$  of  $G$  that include  $\alpha$ , all stable extensions include  $\alpha$ )
- $G \models_{\text{st}} \blacklozenge \Diamond \beta$  (there is a subgraph  $G'$  of  $G$  that has a stable extension that contains  $\beta$ )

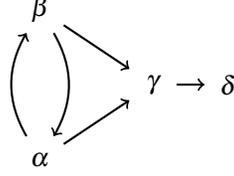


Figure 2: Argument graph for Example 3

- $G \models_{\text{st}} \blacksquare_{\{\alpha, \beta\}} \diamond(\alpha \vee \beta)$  (all subgraphs  $G'$  of  $G$  containing both  $\alpha$  and  $\beta$  have a stable extension that include either  $\alpha$  or  $\beta$ )

Our logic of dialectical outcomes is expressive enough to describe a variety of scenarios concerning arguments, extensions, and subgraphs. Before we continue with using our logic for the purpose of lotteries we have a first look at its logical properties.

**Proposition 1.** For each semantics  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ ,

1.  $\blacksquare\blacksquare\phi \equiv_X \blacksquare\phi$ ,  $\blacklozenge\blacklozenge\phi \equiv_X \blacklozenge\phi$ ,  $\square\square\phi \equiv_X \square\phi$ ,  $\diamond\diamond\phi \equiv_X \diamond\phi$
2.  $\neg\blacksquare\phi \equiv_X \blacklozenge\neg\phi$ ,  $\neg\square\phi \equiv_X \diamond\neg\phi$
3.  $\alpha \wedge \neg\alpha \equiv_X \perp$  for every  $\alpha \in \Pi$ .
4.  $\alpha \vee \neg\alpha \equiv_X \top$  for every  $\alpha \in \Pi$ .
5. for every extension  $E \in \text{Extensions}_X(G)$  we have  $G \models_X \diamond(\bigwedge_{\alpha \in E} \alpha \wedge \bigwedge_{\alpha \notin E} \neg\alpha)$ .
6. If  $X \neq \text{st}$ ,  $G \models_X \square\phi$  implies  $G \models_X \diamond\phi$ .
7.  $G \models_{\text{gr}} \square\phi$  whenever  $G \models_{\text{gr}} \diamond\phi$ .

*Proof.* Items 1–4 and 6 follow directly from Definition 3 and the fact that every argument graph possesses at least one grounded, preferred, and complete extension. As for 5, Let  $E \in \text{Extensions}_X(G)$  and observe  $G, E \models_X \alpha$  for every  $\alpha \in E$  and  $G, E \models_X \neg\alpha$  for every  $\alpha \notin E$ . Therefore  $G, E \models_X \bigwedge_{\alpha \in E} \alpha \wedge \bigwedge_{\alpha \notin E} \neg\alpha$  and by definition of  $\diamond$  we get the statement 5. For 7, note that grounded semantics is a unique-status semantics, so there is always exactly one grounded extension [Dun95]. Therefore,  $\square\phi$  and  $\diamond\phi$  are equivalent.  $\square$

Using our logic we can phrase formulae that describe if some arguments are included in an extension and some arguments are excluded in that extension for a given argument graph. For instance, for a graph  $G$  containing numerous arguments including arguments  $\alpha$  and  $\beta$ , we may want to know whether argument  $\alpha$  is included and  $\beta$  is excluded from the grounded extension of  $G$ . We might be unconcerned about the other arguments in  $G$ . Therefore, we want to know if  $\diamond(\alpha \wedge \neg\beta)$  is a valid inference from  $G$  wrt. the grounded semantics.

**Example 3.** Consider the argument graph  $G$  in Figure 2. Some formulae that hold include the following.

$$\begin{aligned} G &\models_{\text{st}} \diamond(\alpha \wedge \delta \wedge \neg\beta \wedge \neg\gamma) \\ G &\models_{\text{st}} \blacklozenge(\beta \wedge \gamma \wedge \neg\alpha \wedge \neg\delta) \\ G &\models_{\text{st}} \square(\delta \wedge \neg\gamma) \\ G &\models_{\text{gr}} \blacklozenge_{\{\alpha,\beta\}} \diamond(\neg\alpha \wedge \neg\beta) \end{aligned}$$

**Definition 4.** The set of **models** for a formula  $\phi \in \mathcal{L}_\Pi$  is defined via

$$\text{Models}_X^\Pi(\phi) = \{(G, E) \in \text{Interpretations}(\mathcal{L}_\Pi) \mid (G, E) \models_X \phi\}$$

A question we could ask of our logic of dialectical outcomes is whether we can represent argument graphs as formulae of the language. To address this question, the next two lemmas show that we can represent individual attacks, and then the subsequent proposition (i. e. Proposition 2) shows that it is possible to represent any argument graph as a formula.

**Lemma 1.** For  $\alpha \neq \beta$ , it holds  $(\alpha, \beta) \in \text{Arcs}(G)$  if and only if  $G \models_X \blacklozenge_{\{\alpha,\beta\}}^{\{\alpha,\beta\}} \left( \blacksquare_{\{\beta\}} \square\beta \wedge \square\neg\beta \right)$  for any  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ .

*Proof.* “ $\Rightarrow$ ” Let  $G$  be a graph with  $(\alpha, \beta) \in \text{Arcs}(G)$ . Then there is a subgraph  $G' \sqsubseteq G$  containing only the nodes  $\alpha, \beta$  and the attack  $(\alpha, \beta)$  (this corresponds to  $\blacklozenge_{\{\alpha,\beta\}}^{\{\alpha,\beta\}}$ ). In  $G'$ , we have that all extensions—regardless of the semantics, all semantics considered here have this feature—do not contain  $\beta$  ( $\square\neg\beta$ ). Furthermore, all subgraphs  $G'' \sqsubseteq G'$  that contain only node  $\beta$  contain no further attack ( $\blacksquare_{\{\beta\}}^{\{\beta\}}$ ). In  $G''$  all extensions contain  $\beta$  ( $\square\beta$ ). Therefore  $G \models_X \blacklozenge_{\{\alpha,\beta\}}^{\{\alpha,\beta\}} \left( \blacksquare_{\{\beta\}} \square\beta \wedge \square\neg\beta \right)$ .

“ $\Leftarrow$ ” Let  $G$  be such that  $G \models_X \blacklozenge_{\{\alpha,\beta\}}^{\{\alpha,\beta\}} \left( \blacksquare_{\{\beta\}} \square\beta \wedge \square\neg\beta \right)$  is valid and assume  $(\alpha, \beta) \notin \text{Arcs}(G)$ . Consider any subgraph  $G' \sqsubseteq G$  that contains both  $\alpha$  and  $\beta$ . In order for  $\square\neg\beta$  to hold there must be an attack on  $\beta$  in  $G'$ . As  $(\alpha, \beta) \notin \text{Arcs}(G)$  is assumed, there must be an attack from  $\beta$  on  $\beta$ , i. e.,  $(\beta, \beta) \in \text{Arcs}(G)$ . But then there is a subgraph  $G'' \sqsubseteq G'$  containing only node  $\beta$  where  $\beta$  is not in an extension (the subgraph that contains only  $\beta$  and the self-attack). Therefore  $\blacksquare_{\{\beta\}} \square\beta$  is not valid, contradicting the assumption.  $\square$

For any  $\alpha, \beta$  we abbreviate  $\xi_{\alpha,\beta} = \blacklozenge_{\{\alpha,\beta\}}^{\{\alpha,\beta\}} \left( \blacksquare_{\{\beta\}} \square\beta \wedge \square\neg\beta \right)$ .

**Lemma 2.** It holds  $(\alpha, \alpha) \in \text{Arcs}(G)$  if and only if  $G \models_X \blacklozenge_{\{\alpha\}}^{\{\alpha\}} \square\neg\alpha$  for any  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ .

*Proof.* “ $\Rightarrow$ ” Let  $G$  be a graph with  $(\alpha, \alpha) \in \text{Arcs}(G)$ . Then there is a subgraph  $G' \sqsubseteq G$  of  $G$  containing only  $\alpha$  and its self-attack. As a self-attacking argument cannot be in an extension it follows  $\square\neg\alpha$  is valid for  $G'$  and therefore  $G \models_X \blacklozenge_{\{\alpha\}}^{\{\alpha\}} \square\neg\alpha$ .

“ $\Leftarrow$ ” Let  $G$  be such that  $G \models_X \blacklozenge_{\{\alpha\}}^{\{\alpha\}} \square\neg\alpha$ . This means there is a subgraph  $G' \sqsubseteq G$  that contains only  $\alpha$  but no extension of  $G'$  contains  $\alpha$ . This is equivalent to having an attack from  $\alpha$  to  $\alpha$ .  $\square$

For any  $\alpha$  we abbreviate  $\xi_\alpha = \diamond_{\{\alpha\}} \square \neg \alpha$  and, for reasons of simplicity, identify  $\xi_{\alpha,\alpha}$  with  $\xi_\alpha$ .

**Proposition 2.** *For finite  $\Phi \subseteq \text{Interpretations}(\mathcal{L}_\Pi)$ , there is a  $\phi \in \mathcal{L}_\Pi$  such that  $\text{Models}_X^\Pi(\phi) = \Phi$ .*

*Proof.* Let  $\Phi \subseteq \text{Interpretations}(\mathcal{L}_\Pi)$  with  $\Phi = \{(G_1, E_1), \dots, (G_n, E_n)\}$  and define  $\phi = \phi_1 \vee \dots \vee \phi_n$  through

$$\begin{aligned} \phi_i = & \bigwedge_{\alpha \in E_i} \alpha \wedge \bigwedge_{\alpha \notin E_i} \neg \alpha \wedge \bigwedge_{\alpha \in \text{Nodes}(G_i)} \diamond \diamond \alpha \wedge \bigwedge_{\alpha \notin \text{Nodes}(G_i)} \neg \diamond \diamond \alpha \wedge \\ & \bigwedge_{(\alpha, \beta) \in \text{Arcs}(G_i)} \xi_{\alpha, \beta} \wedge \bigwedge_{(\alpha, \beta) \notin \text{Arcs}(G_i)} \neg \xi_{\alpha, \beta} \end{aligned}$$

for  $i = 1, \dots, n$ . In order to show  $\text{Models}_X^\Pi(\phi) = \Phi$  we show 1.)  $\text{Models}_X^\Pi(\phi) \subseteq \Phi$  and 2.)  $\text{Models}_X^\Pi(\phi) \supseteq \Phi$ .

1. Let  $(G, E) \in \text{Models}_X^\Pi(\phi)$ . By construction of  $\phi$  there is  $i \in \{1, \dots, n\}$  such that  $(G, E) \models_X \phi_i$ . Due to the first two terms of the definition of  $\phi_i$  it has to hold  $E = E_i$ . Observe that the term  $\diamond \diamond \alpha$  reads as “there is a subgraph that has an extension which contains  $\alpha$ ” which translates to “ $\alpha$  has to be in the graph”. Conversely,  $\neg \diamond \diamond \alpha$  means that  $\alpha$  cannot be in the graph (otherwise there is always the subgraph just containing  $\alpha$  without any edges which has an extension containing  $\alpha$ ). So, due to the third and fourth terms,  $\text{Nodes}(G) = \text{Nodes}(G_i)$ . Due to the fifth and sixth terms and Lemmas 1 and 2,  $\text{Arcs}(G) = \text{Arcs}(G_i)$  and therefore  $G = G_i$ .

2. For  $i = 1, \dots, n$  note that  $(G_i, E_i) \models_X \phi_i$  by construction. □

**Definition 5.** *A formula  $\phi \in \mathcal{L}_\Pi$  is **unsatisfiable** iff  $\text{Models}_X^\Pi(\phi) = \emptyset$ .*

Obviously, for all arguments  $\alpha$ , the formula  $\alpha \wedge \neg \alpha$  is unsatisfiable. More generally, any conjunction of literals that includes complementary literals is unsatisfiable.

### 3.2 Representing structure and extensions

The next definition is for a formula that captures exactly the extensions of a graph  $G$ . The models of this formula are the pairs  $(G, E)$  where  $E$  is an extension of  $G$  with respect to the semantics  $X$ .

**Definition 6.** *For all  $G$ ,  $\phi \in \mathcal{L}_\Pi$  **reflects**  $G$  wrt.  $X$  iff  $\text{Models}_X^{\text{Nodes}(G)}(\phi) = \{(G, E) \mid E \in \text{Extensions}_X(G)\}$ .*

**Proposition 3.** *For  $G \in \text{Graphs}(\Pi)$ , and for  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ , there is a  $\phi \in \mathcal{L}_\Pi$  such that  $\phi$  reflects  $G$  wrt.  $X$ .*

*Proof.* This follows directly from Proposition 2. □

**Example 4.** *Consider the graph  $G$  in Figure 3a and let  $\Gamma = \diamond \diamond (\alpha \wedge \beta) \wedge \xi_{\alpha, \beta} \wedge \xi_{\beta, \alpha} \wedge \neg \xi_\alpha \wedge \neg \xi_\beta$ .*

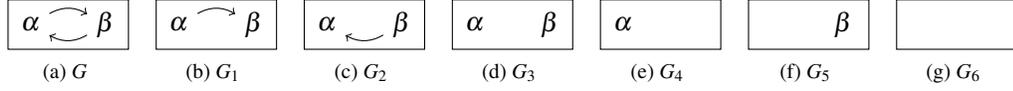


Figure 3: A simple argument graph  $G$  and its subgraphs.

- $\psi_1 = \neg\alpha \wedge \neg\beta \wedge \Gamma$  reflects  $G$  wrt. grounded semantics
- $\psi_2 = ((\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta)) \wedge \Gamma$  reflects  $G$  wrt. preferred semantics
- $\psi_3 = ((\neg\alpha \wedge \neg\beta) \vee (\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta)) \wedge \Gamma$  reflects  $G$  wrt. complete semantics

**Example 5.** Consider the graph  $G$  where  $\text{Nodes}(G) = \{\alpha, \beta\}$  and  $\text{Arcs}(G) = \{(\alpha, \alpha), (\beta, \beta)\}$ .

- $\neg\alpha \wedge \neg\beta \wedge \diamond(\alpha \wedge \beta) \wedge \xi_\alpha \wedge \xi_\beta \wedge \neg\xi_{\alpha,\beta} \wedge \neg\xi_{\beta,\alpha}$  reflects  $G$  wrt. grounded, preferred, and complete semantics.

**Proposition 4.** For  $G \in \text{Graphs}(\Pi)$ , for  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ , and for  $\phi \in \mathcal{L}_\Pi$ , if  $\phi$  reflects  $G$  wrt.  $X$ , then  $G \models_X \phi$ .

*Proof.* If  $\phi$  reflects  $G$  then  $\text{Models}_X(\phi) = \{(G, E) \mid E \in \text{Extensions}_X(G)\}$  which means  $(G, E) \models_X \phi$  for all  $E \in \text{Extensions}_X(G)$ . This is equivalent to stating  $G \models_X \phi$ .  $\square$

**Definition 7.** For all  $G, \phi \in \mathcal{L}_\Pi$  exactly reflects  $G' \sqsubseteq G$  in  $G$  wrt.  $X$  iff  $\text{Models}_X^{\text{Nodes}(G)}(\phi) = \{(G', E) \mid E \in \text{Extensions}_X(G')\}$

**Example 6.** For Figure 3,  $\psi_1$  from Example 4 exactly reflects  $G$  in  $G$  itself with respect to grounded semantics,  $\psi_2$  exactly reflects  $G$  in  $G$  itself with respect to preferred semantics, and  $\psi_3$  exactly reflects  $G$  in  $G$  itself with respect to complete semantics.

**Example 7.** Consider  $G$  and its subgraph  $G_1$  in Figure 4 and observe that

$$\phi = ((\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta)) \wedge \diamond(\alpha \wedge \neg\beta) \wedge \diamond(\neg\alpha \wedge \beta)$$

reflects  $G_1$  wrt. preferred semantics but does not exactly reflect  $G_1$  in  $G$  wrt. preferred semantics as  $(G, \{\alpha\}) \in \text{Models}_{\text{pr}}^{\{\alpha, \beta, \gamma\}}(\phi)$  as well. However, observe that

$$\phi' = \phi \wedge \blacksquare \neg\gamma$$

exactly reflects  $G_1$  in  $G$  wrt. preferred semantics.

The next definition of *constitutes* captures when a set of formulae reflects every subgraph of a graph. When a set of formulae minimally constitutes a graph, there is a bijection from the set of formulae to the set of subgraphs of the graph.

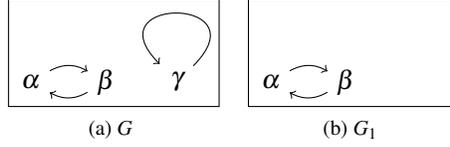


Figure 4: The argument graph  $G$  from Example 7 and one of its subgraphs.

**Definition 8.** For  $G \in \text{Graphs}(\Pi)$ , for  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ , and for  $\Phi \subseteq \mathcal{L}_\Pi$ ,  $\Phi$  **constitutes**  $G$  with respect to  $X$  iff for each subgraph  $G' \sqsubseteq G$ , there is a  $\phi \in \Phi$  such that  $\phi$  exactly reflects  $G'$  in  $G$  wrt.  $X$ . Furthermore,  $\Phi$  **minimally constitutes**  $G$  with respect to  $X$  iff  $\Phi$  constitutes  $G$  with respect to  $X$  and it is not the case that there is a  $\Phi' \subset \Phi$  such that  $\Phi'$  constitutes  $G$  with respect to  $X$ .

**Example 8.** Consider Figure 3. The following formulae minimally constitute  $G$  with respect to preferred semantics.

- $((\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta)) \wedge \diamond(\alpha \wedge \neg\beta) \wedge \diamond(\neg\alpha \wedge \beta)$  exactly reflects  $G$  in  $G$
- $\alpha \wedge \neg\beta \wedge \square(\alpha \wedge \neg\beta) \wedge \diamond\beta \wedge \neg\xi_\beta$  exactly reflects  $G_1$  in  $G$
- $\neg\alpha \wedge \beta \wedge \square(\neg\alpha \wedge \beta) \wedge \diamond\alpha \wedge \neg\xi_\alpha$  exactly reflects  $G_2$  in  $G$
- $\alpha \wedge \beta \wedge \square(\alpha \wedge \beta)$  exactly reflects  $G_3$  in  $G$
- $\alpha \wedge \neg\beta \wedge \square\alpha \wedge \blacksquare\neg\beta$  exactly reflects  $G_4$  in  $G$
- $\beta \wedge \neg\alpha \wedge \square\beta \wedge \blacksquare\neg\alpha$  exactly reflects  $G_5$  in  $G$
- $\neg\alpha \wedge \neg\beta \wedge \blacksquare\neg(\alpha \wedge \neg\beta)$  exactly reflects  $G_6$  in  $G$

**Proposition 5.** For  $G \in \text{Graphs}(\Pi)$ , for  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ , and for  $\Phi \subseteq \mathcal{L}_\Pi$ , if  $\Phi$  constitutes  $G$  with respect to  $X$ , then

$$\bigcup_{\phi \in \Phi} \text{Models}_X^{\text{Nodes}(G)}(\phi) = \{(G', E) \mid G' \sqsubseteq G \text{ and } E \in \text{Extensions}_X(G')\}$$

*Proof.* Assume  $\Phi$  constitutes  $G$  with respect to  $X$ . Therefore, by Definition 8, for each subgraph  $G' \sqsubseteq G$ , there is a  $\phi \in \Phi$  such that  $\phi$  exactly reflects  $G'$  in  $G$  wrt.  $X$ . Therefore, by Definition 6, for each subgraph  $G' \sqsubseteq G$ , there is a  $\phi \in \Phi$  such that  $\text{Models}_X^{\text{Nodes}(G)}(\phi) = \{(G', E) \mid E \in \text{Extensions}_X(G')\}$ . Therefore, if  $\Phi = \{\phi_1, \dots, \phi_k\}$ , then  $\text{Models}_X^{\text{Nodes}(G)}(\phi_1) \cup \dots \cup \text{Models}_X^{\text{Nodes}(G)}(\phi_k)$  is  $\{(G', E) \mid G' \sqsubseteq G \text{ and } E \in \text{Extensions}_X(G')\}$ .  $\square$

**Proposition 6.** For  $G \in \text{Graphs}(\Pi)$ , for  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ , and for  $\Phi \subseteq \mathcal{L}_\Pi$ , if  $\Phi$  minimally constitutes  $G$  with respect to  $X$ , then for all  $\phi_i, \phi_j \in \Phi$ ,  $\text{Models}_X^{\text{Nodes}(G)}(\phi_i) \cap \text{Models}_X^{\text{Nodes}(G)}(\phi_j) = \emptyset$ .

*Proof.* Assume  $\Phi$  minimally constitutes  $G$  with respect to  $X$ . Therefore, by Definition 8,  $\Phi$  constitutes  $G$  with respect to  $X$  and it is not the case that there is a  $\Phi' \subset \Phi$  such that  $\Phi'$  constitutes  $G$  with respect to  $X$ . Therefore, there is exactly one  $\phi \in \Phi$  for each  $G' \sqsubseteq G$  such that  $\text{Models}_X(\phi) = \{(G', E) \mid E \in \text{Extensions}_X(G')\}$ . Therefore, for each  $\phi_i, \phi_j \in \Phi$ , where  $\phi_i \neq \phi_j$ , let  $\text{Models}_X^{\text{Nodes}(G)}(\phi_i) = \{(G_i, E) \mid E \in \text{Extensions}_X(G')\}$  and  $\text{Models}_X^{\text{Nodes}(G)}(\phi_j) = \{(G_j, E) \mid E \in \text{Extensions}_X(G')\}$ . Since  $G_i \neq G_j$ , we have that  $\text{Models}_X^{\text{Nodes}(G)}(\phi_i) \cap \text{Models}_X^{\text{Nodes}(G)}(\phi_j) = \emptyset$ .  $\square$

In this subsection, we have shown how we can use a formula to characterise the structure and extensions of a graph. This means we can use formulae to directly discuss individual graphs, and individual subgraphs, and their extensions. We will harness this granularity when we use the set of subgraphs of a graph as a sample space.

### 3.3 Dividers

Given a graph  $G$  and a formula  $\phi \in \mathcal{L}_\Pi$  the set of *dividers* of  $\phi$  is the set of subgraphs that have  $\phi$  as an inference.

**Definition 9.** Let  $G$  be an argument graph and let  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$  be a semantics. A graph  $G' = (\mathcal{A}', \mathcal{R}') \sqsubseteq G$  is a **divider** for  $\phi \in \mathcal{L}_\Pi$  iff  $G' \models_X \phi$ . Let  $\text{Dividers}_X^G(\phi)$  be the set of dividers of  $\phi$  wrt.  $G$  and  $X$ .

**Example 9.** We continue Example 2. There we have among others

$$\begin{aligned} \text{Dividers}_{\text{gr}}^G(\alpha \wedge \neg\beta) &= \{G, G_2, G_3, G_4, G_6, G_9, G_{12}, G_{15}\} \\ \text{Dividers}_{\text{gr}}^G(\alpha) &= \{G, G_2, G_3, G_4, G_6, G_7, G_9, G_{11}, G_{12}, G_{15}\} \\ \text{Dividers}_{\text{gr}}^G(\beta \wedge \neg\alpha) &= \{G_5, G_{10}, G_{14}, G_{16}\} \\ \text{Dividers}_{\text{gr}}^G(\alpha \wedge \neg\alpha) &= \emptyset \\ \text{Dividers}_{\text{gr}}^G(\alpha \wedge \beta \wedge \neg\gamma) &= \{G_{11}\} \\ \text{Dividers}_{\text{gr}}^G(\alpha \vee \neg\alpha) &= \{G, G_1, \dots, G_{18}\} \\ \text{Dividers}_{\text{gr}}^G(\blacksquare_{\{\alpha, \beta\}}^{\{\alpha, \beta\}} \neg\alpha \wedge \neg\beta) &= \{G, G_1, G_8, G_{12}, \dots, G_{18}\} \end{aligned}$$

Note that all formulae appearing above may have been prefixed by either  $\square$  or  $\diamond$  without changing the dividers as the grounded semantics is a unique-extension semantics.

**Example 10.** Consider Figure 3. There we have among others

$$\begin{array}{ll}
\text{Dividers}_{\text{gr}}^G(\alpha) = \{G_1, G_3, G_4\} & \text{Dividers}_{\text{pr}}^G(\alpha) = \{G_1, G_3, G_4\} \\
\text{Dividers}_{\text{pr}}^G(\diamond\alpha) = \{G, G_1, G_3, G_4\} & \text{Dividers}_{\text{gr}}^G(\beta) = \{G_2, G_3, G_5\} \\
\text{Dividers}_{\text{pr}}^G(\beta) = \{G_2, G_3, G_5\} & \text{Dividers}_{\text{pr}}^G(\diamond\beta) = \{G, G_2, G_3, G_5\} \\
\text{Dividers}_{\text{gr}}^G(\alpha \wedge \beta) = \{G_3\} & \text{Dividers}_{\text{pr}}^G(\alpha \wedge \beta) = \{G_3\} \\
\text{Dividers}_{\text{pr}}^G(\diamond(\alpha \wedge \beta)) = \{G_3\} & \text{Dividers}_{\text{gr}}^G(\alpha \wedge \neg\beta) = \{G_1, G_4\} \\
\text{Dividers}_{\text{pr}}^G(\alpha \wedge \neg\beta) = \{G_1, G_4\} & \text{Dividers}_{\text{pr}}^G(\diamond(\alpha \wedge \neg\beta)) = \{G, G_1, G_4\} \\
\text{Dividers}_{\text{gr}}^G(\neg\alpha \wedge \beta) = \{G_2, G_5\} & \text{Dividers}_{\text{pr}}^G(\neg\alpha \wedge \beta) = \{G_2, G_5\} \\
\text{Dividers}_{\text{pr}}^G(\diamond(\neg\alpha \wedge \beta)) = \{G, G_2, G_5\} & \text{Dividers}_{\text{gr}}^G(\neg\alpha \wedge \neg\beta) = \{G, G_6\} \\
\text{Dividers}_{\text{gr}}^G(\neg\alpha \wedge \neg\beta) = \{G, G_6\} & \text{Dividers}_{\text{pr}}^G(\diamond(\neg\alpha \wedge \neg\beta)) = \{G_6\}
\end{array}$$

**Proposition 7.** For every argument graph  $G$  and each semantics  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ , if  $\psi \models_X \phi$  then  $\text{Dividers}_X^G(\psi) \subseteq \text{Dividers}_X^G(\phi)$ .

*Proof.* Let  $G \in \text{Dividers}_X^G(\psi)$ . By Definition 9 it follows  $G \models_X \psi$ . Assuming  $\psi \models_X \phi$  we conclude  $G \models_X \phi$  and therefore  $G \in \text{Dividers}_X^G(\phi)$ .  $\square$

We will use formulae in our logic of dialectical outcomes as our outcomes when we consider lotteries. For this, we need to ensure that each outcome is disjoint from the other outcomes. For this, we use the following definition.

**Definition 10.** Let  $G$  be an argument graph, and let  $\phi, \psi \in \mathcal{L}_\Pi$ .  $\phi$  and  $\psi$  are **disjoint** for  $G$  and wrt. semantics  $X$  iff  $\text{Dividers}_X^G(\phi) \cap \text{Dividers}_X^G(\psi) = \emptyset$ . A set  $\Phi \subseteq \mathcal{L}_\Pi$  is **pairwise disjoint** iff for each  $\phi, \psi \in \Phi$ ,  $\phi$  and  $\psi$  are disjoint.

**Example 11.** We continue Example 9 and grounded semantics. Then  $\alpha \wedge \neg\beta$  and  $\beta \wedge \neg\alpha$  are disjoint and  $\alpha$  and  $\gamma$  are not disjoint as, e. g.,  $G_{12}$  is a divider for both of them.

We can equivalently view disjointness between a pair of outcomes as inconsistency between these outcomes as shown in the next proposition.

**Proposition 8.** Let  $G$  be an argument graph, and let  $\phi, \psi \in \mathcal{L}_\Pi$ .  $\phi$  and  $\psi$  are disjoint for  $G$  and wrt. semantics  $X$  iff  $\phi \wedge \psi \equiv_X \perp$ .

*Proof.* Let  $\phi, \psi \in \mathcal{L}_\Pi$  and assume  $\Pi = \text{Nodes}(G)$ .

“ $\Rightarrow$ ” Let  $\phi$  and  $\psi$  be disjoint and assume  $\phi \wedge \psi \not\equiv_X \perp$ . Then there is  $(G', E) \in \text{Interpretations}(\mathcal{L}_\Pi)$  with  $(G', E) \models_X \phi$  and  $(G', E) \models_X \psi$ . But then  $G'$  would be a divider of both  $\phi$  and  $\psi$ .

“ $\Leftarrow$ ” If  $\phi \wedge \psi \equiv_X \perp$  then there is no  $(G', E) \in \text{Interpretations}(\mathcal{L}_\Pi)$  with  $(G', E) \models_X \phi$  and  $(G', E) \models_X \psi$ , i. e., there is no common divider of  $\phi$  and  $\psi$ .  $\square$

The next definition is for when a set of outcomes is exhaustive. Again, this is required for when we use outcomes in lotteries. The definition for exhaustive is specified as a constraint using a formula of our language for dialectical outcomes. In the subsequent proposition, we show that this definition is implied by to the union of the dividers of the formulae being the set of all subgraphs of the argument graph.

**Definition 11.** A set of formulae  $\{\phi_1, \dots, \phi_k\} \subseteq \mathcal{L}_\Pi$  is **exhaustive** for  $G$  wrt. semantics  $X$  iff  $G \models_X \blacksquare \Box(\phi_1 \vee \dots \vee \phi_k)$ .

**Proposition 9.** Let  $\{\phi_1, \dots, \phi_k\} \subseteq \mathcal{L}_\Pi$ . If

$$\text{Dividers}_X^G(\phi_1) \cup \dots \cup \text{Dividers}_X^G(\phi_k) = \text{Sub}(G) \quad (1)$$

then  $\{\phi_1, \dots, \phi_k\}$  is exhaustive for  $G$  wrt. semantics  $X$ .

*Proof.* Let  $\{\phi_1, \dots, \phi_k\} \subseteq \mathcal{L}_\Pi$  be a set of formulae. Note first that  $\text{Dividers}_X^G(\phi_1) \cup \dots \cup \text{Dividers}_X^G(\phi_k) \subseteq \text{Sub}(G)$  for any set  $\{\phi_1, \dots, \phi_k\} \subseteq \mathcal{L}_\Pi$  due to the definition of dividers. Then (1) is equivalent to

$$\begin{aligned} & \text{Dividers}_X^G(\phi_1) \cup \dots \cup \text{Dividers}_X^G(\phi_k) = \text{Sub}(G) \\ \iff & \forall G' \in \text{Sub}(G) : \exists i \in \{1, \dots, k\} : G' \in \text{Dividers}_X^G(\phi_i) \\ \iff & \forall G' \in \text{Sub}(G) : \exists i \in \{1, \dots, k\} : G' \models_X \phi_i \\ \iff & \forall G' \in \text{Sub}(G) : \exists i \in \{1, \dots, k\} : \forall E \in \text{Extensions}_X(G') : (G', E) \models_X \phi_i \quad (2) \end{aligned}$$

and furthermore from (2) follows

$$\begin{aligned} \implies & \forall G' \in \text{Sub}(G) : \exists i \in \{1, \dots, k\} : \forall E \in \text{Extensions}_X(G') : (G', E) \models_X \phi_1 \vee \dots \vee \phi_k \\ \iff & \forall G' \in \text{Sub}(G) : \forall E \in \text{Extensions}_X(G') : (G', E) \models_X \phi_1 \vee \dots \vee \phi_k \\ \iff & \forall G' \in \text{Sub}(G) : G' \models_X \Box(\phi_1 \vee \dots \vee \phi_k) \\ \iff & G \models_X \blacksquare \Box(\phi_1 \vee \dots \vee \phi_k) \end{aligned}$$

□

Note that the other direction, i. e., from exhaustiveness follows (1), is not true in general as the next example shows.

**Example 12.** Consider the graph  $G$  where  $\text{Nodes}(G) = \{\alpha, \beta\}$  and  $\text{Arcs}(G) = \{(\alpha, \beta), (\beta, \alpha)\}$  and stable semantics. Then  $\{\alpha, \neg\alpha\}$  is exhaustive:

$$\forall G' \in \text{Sub}(G) : \forall E \in \text{Extensions}_{st}(G') : (G', E) \models_X \alpha \vee \neg\alpha$$

as every extension of every subgraph either contains  $\alpha$  or not. However, observe that  $G \in \text{Sub}(G)$  is not a divider for  $\alpha$  nor  $\neg\alpha$  as  $G$  has one extension containing  $\alpha$  and one extension not containing  $\alpha$ .

**Example 13.** We continue Example 10 and therefore consider Figure 1 again. The following set of formulae  $\{\phi_1, \dots, \phi_8\}$  is exhaustive for  $G$  with respect to grounded semantics:

$$\begin{aligned} \phi_1 &= \alpha \wedge \beta \wedge \gamma & \phi_2 &= \alpha \wedge \beta \wedge \neg\gamma & \phi_3 &= \alpha \wedge \neg\beta \wedge \gamma & \phi_4 &= \neg\alpha \wedge \beta \wedge \gamma \\ \phi_5 &= \alpha \wedge \neg\beta \wedge \neg\gamma & \phi_6 &= \neg\alpha \wedge \beta \wedge \neg\gamma & \phi_7 &= \alpha \wedge \neg\beta \wedge \neg\gamma & \phi_8 &= \neg\alpha \wedge \neg\beta \wedge \neg\gamma \end{aligned}$$

Furthermore, the formulae are pairwise disjoint (i. e. for each  $i, j \in \{1, \dots, 8\}$ ,  $\phi_i$  and  $\phi_j$  are disjoint for grounded semantics).

**Example 14.** Consider Figure 3. The following set of formulae  $\{\phi_1, \dots, \phi_5\}$  is exhaustive for  $G$  with respect to preferred semantics:

$$\begin{array}{ll} \phi_1 = \diamond(\alpha \wedge \neg\beta) \wedge \diamond(\neg\alpha \wedge \beta) & \text{Dividers}(\phi_1) = \{G\} \\ \phi_2 = \alpha \wedge \neg\beta & \text{Dividers}(\phi_2) = \{G_1, G_4\} \\ \phi_3 = \neg\alpha \wedge \beta & \text{Dividers}(\phi_3) = \{G_2, G_5\} \\ \phi_4 = \neg\alpha \wedge \neg\beta & \text{Dividers}(\phi_4) = \{G_6\} \\ \phi_5 = \alpha \wedge \beta & \text{Dividers}(\phi_5) = \{G_3\} \end{array}$$

Furthermore, the formulae are pairwise disjoint for preferred semantics (i. e. for each  $i, j \in \{1, \dots, 5\}$ ,  $\phi_i$  and  $\phi_j$  are disjoint for preferred semantics).

**Proposition 10.** If a set of formulae  $\Phi = \{\phi_1, \dots, \phi_n\}$  is exhaustive and pairwise disjoint then  $\Phi \setminus \{\phi_i, \phi_j\} \cup \{\phi_i \vee \phi_j\}$  is also exhaustive and pairwise disjoint for every  $i, j = 1, \dots, n$ .

*Proof.* Let  $\Phi = \{\phi_1, \dots, \phi_n\}$  be exhaustive and pairwise disjoint. That is,  $G \models_X \blacksquare\Box(\phi_1 \vee \dots \vee \phi_n)$  and each  $\phi_i, \phi_j \in \Phi$  ( $i \neq j$ ) are disjoint.

Without loss of generality consider  $\Phi' = \{\phi_1 \vee \phi_2, \phi_3, \dots, \phi_n\}$ . Then  $\Phi'$  is exhaustive as  $G \models_X \blacksquare\Box(\phi_1 \vee \dots \vee \phi_n)$  is equivalent to  $G \models_X \blacksquare\Box((\phi_1 \vee \phi_2) \vee \phi_3 \vee \dots \vee \phi_n)$ . As  $\Phi$  is pairwise disjoint so is  $\{\phi_3, \dots, \phi_n\} \subseteq \Phi$ . Let now  $k = 3, \dots, n$  and consider  $\phi_1 \vee \phi_2$  and  $\phi_k$ . Assume  $\text{Dividers}_X^G(\phi_1 \vee \phi_2) \cap \text{Dividers}_X^G(\phi_k) \neq \emptyset$  and let  $G' \in \text{Dividers}_X^G(\phi_1 \vee \phi_2) \cap \text{Dividers}_X^G(\phi_k)$ . As  $G' \in \text{Dividers}_X^G(\phi_1 \vee \phi_2)$  it is either  $G' \in \text{Dividers}_X^G(\phi_1)$  or  $G' \in \text{Dividers}_X^G(\phi_2)$  (or both). In any case, this implies  $\Phi$  cannot be pairwise disjoint, contradicting the assumption.  $\square$

**Example 15.** We continue Example 13. As  $\phi_9 = \beta \equiv_X \phi_1 \vee \phi_2 \vee \phi_4 \vee \phi_6$  we have that  $\{\phi_3, \phi_5, \phi_7, \phi_8, \phi_9\}$  is exhaustive and pairwise disjoint.

The following result shows that any set of subgraphs of a graph can be characterized as the dividers of some formula. As we shall use the set of subgraphs of a graph as the sample space for a probability distribution, it means the language allows us to talk about the sample space.

**Proposition 11.** For any  $\Psi \subseteq \text{Sub}(G)$ , there is a formula  $\phi \in \mathcal{L}_\Pi$  such that

$$\text{Dividers}_X^G(\phi) = \Psi$$

*Proof.* Let  $\Psi \subseteq \text{Sub}(G)$  with  $\Psi = \{G_1, \dots, G_n\}$ . For each  $i = 1, \dots, n$ , let  $E^i$  be the set of X-extensions of  $G_i$  and consider the set  $M_i = \{(G_i, E) \mid E \in E^i\}$  of interpretations. By Proposition 2 there is  $\phi_i$  with  $\text{Models}_X(\phi_i) = M_i$  and therefore  $\text{Dividers}_X^G(\phi_i) = \{G_i\}$ . Define  $\phi = \phi_1 \vee \dots \vee \phi_n$  and observe  $\text{Dividers}_X^G(\phi) = \text{Dividers}_X^G(\phi_1) \cup \dots \cup \text{Dividers}_X^G(\phi_n) = \{G_1, \dots, G_n\} = \Psi$ .  $\square$

In a later section, we will use formulae as outcomes in argumentation lotteries. For this, we will want to avoid both tautologies and contradictions as outcomes in useful argumentation lotteries.

## 4 Probability distributions

Given an argument graph  $G$  we represent the uncertainty we may have over the arguments and/or attacks by using the set of subgraphs  $\text{Sub}(G)$  as the sample space. This means we are unsure which subgraph is the “correct” subgraph. Then using this sample space, we define a probability distribution as follows. Note, in previous work, we restricted consideration to the *spanning* subgraphs—i. e. subgraphs containing all attacks on a subset of arguments—thereby denoting uncertainty in the arguments [Hun12], or on the *full* subgraphs—i. e. subgraphs containing all arguments but a subset of attacks—thereby denoting uncertainty in the attacks [Hun13c]. Here we allow the representation of uncertainty in both arguments and attacks.

**Definition 12.** *Let  $G$  be an argument graph. A probability distribution  $P$  for  $G$  is a function  $P : \text{Sub}(G) \rightarrow [0, 1]$  such that  $\sum_{G' \sqsubseteq G} P(G') = 1$ .*

Given a probability distribution over subgraphs, we can obtain the probability of each argument and each attack as a marginal distribution. More specifically, for an argument graph  $G$  and a probability distribution  $P$ , the **marginal distribution for an argument  $\alpha$**  is

$$P(\alpha) = \sum_{G' \sqsubseteq G \text{ s.t. } \alpha \in \text{Nodes}(G')} P(G')$$

The **marginal distribution for an attack  $(\alpha, \beta)$**  is

$$P((\alpha, \beta)) = \sum_{G' \sqsubseteq G \text{ s.t. } (\alpha, \beta) \in \text{Arcs}(G')} P(G')$$

Note that these probabilities describe the uncertainty to which an argument or attack is believed to be justifiable, i. e. whether it is appropriate to consider this element to be present in the argument graph. In particular, a high probability of an argument does not necessarily imply that the argument is highly acceptable, see below.

**Example 16.** *We consider Figure 1. Define a probability distribution  $P$  on  $G$  via  $P(G) = P(G_5) = P(G_8) = 0.1$ ,  $P(G_9) = 0.7$ , and  $P(G') = 0$  for the remaining subgraphs  $G'$  of  $G$ . Then the marginal distributions are as follows:  $P(\alpha) = 1$ ,  $P(\beta) = 1$ ,  $P(\gamma) = 0.2$ ,  $P((\alpha, \beta)) = 0.9$ ,  $P((\beta, \alpha)) = 0.3$ , and  $P((\gamma, \beta)) = 0.1$ .*

If the probability distribution is restricted to a distribution over full subgraphs, then the arguments are certain, and the attacks may be uncertain, whereas if the probability distribution is restricted to a distribution over spanning subgraphs, then the attacks are certain, and the arguments may be uncertain.

Now we use the probability distribution over subgraphs to give a probability that a certain formula in  $\mathcal{L}_{\Pi}$  holds. As defined below, the probability that  $\phi \in \mathcal{L}_{\Pi}$  is true in  $G$  is the sum of the probabilities of the subgraphs for which  $\phi$  is true.

**Definition 13.** *Let  $G$  be an argument graph and let  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$  be a semantics. Also let  $P$  be a probability distribution for  $G$ . For a  $\phi \in \mathcal{L}_{\Pi}$ , the probability of  $\phi$  wrt.  $X$  is*

$$P_X(\phi) = \sum_{G' \in \text{Dividers}_X^G(\phi)} P(G')$$

**Example 17.** We continue Example 16. There we have  $P_{\text{gr}}(\alpha \wedge \neg\beta) = 0.8$  (as  $G$  and  $G_9$  are the only dividers with positive probabilities) and  $P_{\text{gr}}(\beta \wedge \neg\alpha) = 0.1$  (as  $G_5$  is the only divider with positive probability).

The notion of the probability of a formula subsumes the definition of the probability that a set of arguments is an extension, and it subsumes the definition for the probability that an argument is an inference, cf. [LON11, Hun12, Hun13c].

**Proposition 12.** For every argument graph  $G$  and each semantics  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ ,

1. If  $\phi \equiv_X \perp$ , then  $P_X(\phi) = 0$
2. If  $\phi \equiv_X \top$ , then  $P_X(\phi) = 1$
3. If  $\phi \models_X \psi$ , then  $P_X(\phi) \leq P_X(\psi)$
4. If  $\phi \wedge \psi$  is unsatisfiable, then  $P_X(\phi \vee \psi) = P_X(\phi) + P_X(\psi)$ .

*Proof.*

1. If  $\phi \equiv_X \perp$  then  $\text{Models}_X^\Pi(\phi) = \emptyset$  and therefore  $\text{Dividers}_X^G(\phi) = \emptyset$ . It follows  $P_X(\phi) = \sum_{G' \in \emptyset} P(G') = 0$ .
2. If  $\phi \equiv_X \top$  then  $\text{Models}_X^\Pi(\phi) = \text{Interpretations}(\mathcal{L}_\Pi)$  and therefore  $\text{Dividers}_X^G(\phi) = \{G' \mid G' \sqsubseteq G\}$ . It follows  $P_X(\phi) = \sum_{G' \sqsubseteq G} P(G') = 1$ .

3. If  $\phi \models_X \psi$  then  $\text{Dividers}_X^G(\phi) \subseteq \text{Dividers}_X^G(\psi)$  due to Proposition 7. It follows

$$P_X(\phi) = \sum_{G' \in \text{Dividers}_X^G(\phi)} P(G') \leq \sum_{G' \in \text{Dividers}_X^G(\psi)} P(G') = P_X(\psi)$$

4. Let  $\phi \wedge \psi$  be unsatisfiable, i. e.,  $\text{Models}_X^\Pi(\phi \wedge \psi) = \emptyset$  and  $P_X(\phi \wedge \psi) = 0$ . Then we have  $P_X(\phi \vee \psi) = P_X(\phi) + P_X(\psi) - P_X(\phi \wedge \psi) = P_X(\phi) + P_X(\psi)$ .  $\square$

**Proposition 13.** Let  $G$  be the argument graph, let  $\{\phi_1, \dots, \phi_k\} \subseteq \mathcal{L}_\Pi$ , and let  $P$  be a probability distribution. If  $\{\phi_1, \dots, \phi_k\}$  is exhaustive and pairwise disjoint for  $G$  wrt.  $X$ , then

$$\sum_{i=1}^k P_X(\phi_i) = 1$$

*Proof.* Consider

$$\sum_{i=1}^k P_X(\phi_i) = \sum_{i=1}^k \sum_{G' \in \text{Dividers}_X^G(\phi_i)} P(G')$$

As  $\{\phi_1, \dots, \phi_k\} \subseteq \mathcal{L}_\Pi$  is exhaustive and pairwise disjoint it follows that for every  $G' \sqsubseteq G$  the term  $P(G')$  appears exactly once in the above sum ( $G'$  is a divider for exactly one formula). This means the above sum becomes

$$\sum_{i=1}^k \sum_{G' \in \text{Dividers}_X^G(\phi_i)} P(G') = \sum_{G' \in \text{Sub}(G)} P(G') = 1$$

$\square$

In general, the probability distribution is on the set of all subgraphs of the argument graph. However, often the uncertainty will be more focused. For instance, it may be focused on some full subgraphs (thereby reflecting uncertainty in some of the arguments) or on some spanning subgraphs (thereby reflecting uncertainty in some of the attacks) or on some combination of arguments and attacks.

## 5 Argumentation as a lottery

In this section, we show how we can use the logic of dialectical outcomes, and the probability distribution over subgraphs of the argument graph, to give a definition for argumentation lotteries. Our primary aim is to have a mechanism for artificial agents to evaluate the options they have for arguments and/or attacks to present in dialogical argumentation. We are therefore drawing on the computational ability of an artificial agent to evaluate the extensions of an argument graph and of its subgraphs.

### 5.1 A brief review of lotteries

We start by briefly reviewing the notion of a lottery which is an idea that comes from decision theory. A **lottery** is a probability distribution over a set of possible outcomes that are assumed to be disjoint and exhaustive. A lottery with possible outcomes  $\phi_1, \dots, \phi_n$  that can occur with probabilities  $p_1, \dots, p_n$  is written  $[p_1, \phi_1; \dots; p_n, \phi_n]$ .

**Example 18.** Consider a gamble involving a die with 6 sides. Let these sides be called  $s_1, s_2, s_3, s_4, s_5,$  and  $s_6$ . Also assume that each side is equally to occur when the die is rolled. This can be represented by the following lottery.

$$[1/6, s_1; 1/6, s_2; 1/6, s_3; 1/6, s_4; 1/6, s_5; 1/6, s_6]$$

Individual agents may have preferences over the outcomes, which we model using *utility functions*. A utility function  $U$  assigns to each outcome a real value with the interpretation that larger values indicate larger utility. In the above example a utility function  $U$  defined via  $U(s_1) = \dots = U(s_5) = 1$  and  $U(s_6) = 2$  means that the outcome  $s_6$  is preferred to all other outcomes, which themselves are equally preferred. Note that the values of a utility function usually do not have any specific semantics (a value twice as large as another value does not necessarily mean that it is twice as preferred).

A utility function can be used to give an evaluation of how much an agent values an outcome (i. e. how much utility the agent will obtain from the outcome) and monetary values can be used for illustration purposes. For instance, in gambling, it is often clear what the monetary value of each outcome is. But many other outcomes can be quantified in terms of monetary value.

Given the probability distribution over outcomes (as specified in the lottery) and the utility function, an overall value can be calculated for the return that can be expected from a lottery as follows: For a utility function  $U$ , the **expected utility of a lottery**  $L$ , denoted  $E(L, U)$ , is given by

$$E(L, U) = \sum_{i=1}^n p_i U(\phi_i)$$

For decision making, each action has a resulting lottery that incorporates all the possible outcomes of the action, together with associated probability distribution. In utility theory, the **principle of maximum expected utility** states that a rational agent should choose an action that maximizes its expected utility.

**Example 19.** Consider that you are a contestant in a TV game show. The host offers you a \$1 million prize or a gamble on the flip of a coin. If the coins comes up heads, you win \$3 million, otherwise you end up with nothing. The lotteries could then be represented as follows

- For action of coin-flip gamble, the lottery  $L_{\text{coinflip}}$  is  $[1/2, \text{heads}; 1/2, \text{tails}]$
- For action of collect prize, the lottery  $L_{\text{collectprize}}$  is  $[1, \text{get prize}]$

If we take the absolute monetary values as the utility values, then we get the following utility function:  $U(\text{get prize}) = 1M$ ,  $U(\text{heads}) = 3M$ , and  $U(\text{tails}) = 0$ . So the expected utility of the lotteries are as follows

- For action of coin-flip gamble, we have the lottery  $L_{\text{coinflip}}$ , and hence  $E(L_{\text{coinflip}}, U) = 1.5M$
- For action of collect prize, we have the lottery  $L_{\text{collectprize}}$ , and hence  $E(L_{\text{collectprize}}, U) = 1M$

So according to the principle of maximum expected utility, the rational agent should choose the action of coin-flip gamble (if s/he accepts this utility function).

However, this does not mean that accepting the bet is a better or more rational decision for everyone. For instance, for most people considering this example, utility is not directly proportional to monetary value, because the utility (interpreted as the positive change in your lifestyle) for your first million is very high whereas the utility for an additional million is much smaller. This could be reflected by the utility function  $U'$ , where  $U'(\text{get prize}) = 1M$ ,  $U'(\text{heads}) = 1.5M$ , and  $U'(\text{tails}) = 0$ , and so the revised expected utility would be  $E(L_{\text{coinflip}}, U) = 750K$ . Hence, the rational agent who accepts utility function  $U'$  should choose the action of collect prize (and so not enter the coin-flip gamble).

Utility theory and decision theory are extensively studied subjects in economics, social sciences, philosophy, psychology, and more recently in artificial intelligence. For an introduction, see [Pet09]. We have only presented a simple and commonly used form of lottery here. There are numerous alternatives that have been developed to address specific concerns or perceived shortcomings of the simple lottery. Furthermore, we refer the reader to the literature on lotteries for methods for obtaining utility functions for applications.

## 5.2 Argumentation lotteries

We can view an argument graph  $G$  as invoking a lottery. For that we use formulae of  $\mathcal{L}_{\Pi}$  as outcomes and the probability of a formula holding as the probability of the outcome. Furthermore, it is quite natural to think of formulae as having utility. For

example, for an argument graph containing arguments  $\alpha$ ,  $\beta$ , and  $\gamma$ , and we prefer to have  $\alpha$  and  $\beta$  and to not have  $\gamma$ , otherwise we prefer either  $\alpha$  or  $\beta$  and not  $\gamma$ , otherwise we are indifferent about the outcome, then we have the preferences over outcomes where  $\alpha \wedge \beta \wedge \neg\gamma$  is most preferred,  $\alpha \wedge \neg\beta \wedge \neg\gamma$  and  $\neg\alpha \wedge \beta \wedge \neg\gamma$  are the second most preferred, and then  $\neg\alpha \wedge \neg\beta$  is the least preferred. Since, we can identify this preference ordering, we can identify a utility function to indicate the degree to which we prefer each of the options. For instance, we could let the utility function  $U$  be  $U(\alpha \wedge \beta \wedge \neg\gamma) = 10$ ,  $U(\alpha \wedge \neg\beta \wedge \neg\gamma) = 5$ ,  $U(\neg\alpha \wedge \beta \wedge \neg\gamma) = 5$  and  $U(\neg\alpha \wedge \neg\beta) = 0$ , and  $U(\phi) = -10$  for every other option  $\phi$ .

We formalize the construction of a lottery for argumentation as follow. For this we need to ensure that the formulae we use as outcomes in the lottery are pairwise disjoint and together they are exhaustive.

**Definition 14.** Let  $G$  be the argument graph, let  $\{\phi_1, \dots, \phi_k\} \subseteq \mathcal{L}_\Pi$ , and let  $P$  be a probability distribution. The tuple  $[P_X(\phi_1), \phi_1; \dots; P_X(\phi_k), \phi_k]$  is an **argumentation lottery** for  $G$  wrt. semantics  $X$  iff  $\{\phi_1, \dots, \phi_k\}$  is exhaustive and pairwise disjoint for  $G$  wrt.  $X$ .

**Example 20.** We continue Example 17. Recall that  $\{\phi_3, \phi_5, \phi_7, \phi_8, \phi_9\}$  is exhaustive and pairwise disjoint for  $G$  with respect to grounded semantics (see Example 14) and let  $P$  be as defined in Example 16. This gives the following argumentation lottery

$$[P(\phi_3), \phi_3; P(\phi_5), \phi_5; P(\phi_7), \phi_7; P(\phi_8), \phi_8; P(\phi_9), \phi_9] = [0.1, \phi_3; 0.7, \phi_5; 0, \phi_7; 0.1, \phi_8; 0.1, \phi_9]$$

Define a utility function  $U$  via  $U(\phi_3) = 10$ ,  $U(\phi_5) = 5$ ,  $U(\phi_7) = 5$ ,  $U(\phi_8) = 0$ , and for  $k \in \{1, 2, 4, 6, 9\}$ ,  $U(\phi_k) = -10$ . Observe that  $U$  favours  $\alpha$  and/or  $\gamma$ , but not  $\beta$ , in our grounded extension. Therefore the expected utility is  $(0.1 \cdot 10) + (0.7 \cdot 5) + (0.0 \cdot 5) + (0.1 \cdot 0) + (0.1 \cdot -10) = 3.5$ .

**Example 21.** We return to Figure 3. Consider the formulae  $\phi_1 = \diamond(\alpha \wedge \neg\beta)$  and  $\phi_2 = \neg(\diamond(\alpha \wedge \neg\beta))$ . These are exhaustive and disjoint. Suppose  $P(G) = 0.8$ ,  $P(G_1) = 0.1$ , and  $P(G_2) = 0.1$ . Therefore  $P_{\text{pr}}(\phi_1) = P(G) + P(G_1) = 0.9$  and  $P_{\text{pr}}(\phi_2) = P(G_2) = 0.1$ . This gives the following argumentation lottery

$$[P_{\text{pr}}(\phi_1), \phi_1; P_{\text{pr}}(\phi_2), \phi_2]$$

Let  $U$  be the utility function where  $U(\phi_1) = 10$ , and  $U(\phi_2) = -10$ . Observe that  $U$  favours to have an extension with  $\alpha$  and not  $\beta$  before anything else. Therefore the expected utility is  $(0.9 \cdot 10) + (0.1 \cdot -10) = 8$ .

The following result shows that there is always an argumentation lottery. Furthermore, with the use of subsumption, we can restructure the argumentation lottery to reduce the number of outcomes as illustrated in the above example.

**Proposition 14.** Let  $G$  be an argument graph,  $P$  a probability distribution on  $G$ , and  $\{\phi_1, \dots, \phi_k\}$  is the following set.

$$\begin{aligned} & \{\alpha_1 \wedge \dots \wedge \alpha_m \wedge \neg\beta_1 \wedge \dots \wedge \neg\beta_n \mid \{\alpha_1, \dots, \alpha_m\} \subseteq \text{Nodes}(G) \text{ and} \\ & \{\beta_1, \dots, \beta_n\} = \text{Nodes}(G) \setminus \{\alpha_1, \dots, \alpha_m\}\} \end{aligned}$$

Then  $[P(\phi_1), \phi_1; \dots; P(\phi_k), \phi_k]$  is an argumentation lottery for  $G$  with respect to grounded semantics.

*Proof.* We have to show that  $\Phi = \{\phi_1, \dots, \phi_k\}$  is exhaustive and pairwise disjoint for  $G$  wrt. grounded semantics.

(Exhaustive)  $\Phi$  is exhaustive iff  $G \models_{\text{gr}} \blacksquare \square (\phi_1 \vee \dots \vee \phi_k)$   
iff  $G, E \models_{\text{gr}} \blacksquare \square (\phi_1 \vee \dots \vee \phi_k)$  for all  $E \in \text{Extensions}_{\text{gr}}(G)$   
iff  $G', E \models_{\text{gr}} \square (\phi_1 \vee \dots \vee \phi_k)$  for all  $E \in \text{Extensions}_{\text{gr}}(G)$  and  $G' \sqsubseteq G$   
iff  $G', E' \models_{\text{gr}} \phi_1 \vee \dots \vee \phi_k$  for all  $E' \in \text{Extensions}_{\text{gr}}(G')$  and  $G' \sqsubseteq G$

Note that the last statement is true as there is  $i \in \{1, \dots, k\}$  such that  $\phi_i = \alpha_1 \wedge \dots \wedge \alpha_m \wedge \neg\beta_1 \wedge \dots \wedge \neg\beta_n$  with  $E' = \{\alpha_1, \dots, \alpha_m\}$ .

(Disjoint) For all  $i, j \in \{1, \dots, k\}$ ,  $i \neq j$  we have that  $\phi_i$  and  $\phi_j$  are disjoint as there is necessarily a  $\gamma \in \text{Nodes}(G)$  such that  $\phi_i \models_{\text{gr}} \gamma$  and  $\phi_j \models_{\text{gr}} \neg\gamma$  (or the other way around).  $\square$

**Example 22.** Let  $G$  be the argument graph in Figure 3. Also let  $P$  a probability distribution on  $G$ . Applying Proposition 14, if  $\{\phi_1, \dots, \phi_4\}$  is the set of the following formulae, then  $[P(\phi_1), \phi_1; \dots; P(\phi_4), \phi_4]$  is an argumentation lottery for  $G$  with respect to any semantics.

$$\phi_1 = \alpha \wedge \beta \quad \phi_2 = \neg\alpha \wedge \beta \quad \phi_3 = \alpha \wedge \neg\beta \quad \phi_4 = \neg\alpha \wedge \neg\beta$$

The next result draws on the notion of when a set of formulae minimally constitutes an argument graph (Definition 8). If a set of formulae minimally constitutes a set of formulae a graph with respect to a semantics, then we can construct an argumentation lottery using these formulae as the outcomes. In this case, there is a unique formula for each subgraph of the graph.

**Proposition 15.** Let  $G$  be an argument graph, let  $P$  be a probability distribution on  $G$ , let  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ , and  $\Phi \subseteq \mathcal{L}_{\Pi}$ . If  $\Phi$  minimally constitutes  $G$  with respect to  $X$ , where  $\Phi = \{\phi_1, \dots, \phi_k\}$ , then  $[P(\phi_1), \phi_1; \dots; P(\phi_k), \phi_k]$  is an argumentation lottery for  $G$  with respect to  $X$ .

*Proof.* Assume  $\Phi$  minimally constitutes  $G$  with respect to  $X$ .

- (Exhaustive)  $\text{Models}(\phi_1) \cup \dots \cup \text{Models}(\phi_k) = \{(G', E) \mid G' \sqsubseteq G \text{ and } E \in \text{Extensions}_X(G')\}$  by Proposition 5. For each  $\phi_i \in \Phi$ , let  $\text{Models}_X(\phi_i) = \{(G_i, E) \mid E \in \text{Extensions}_X(G_i)\}$ . Therefore, for each  $G_i \sqsubseteq G$ , there is a  $\phi_i \in \Phi$  such that  $G_i \models_X \phi_i$ . Therefore, by Proposition 9,  $\Phi$  is exhaustive.
- (Disjoint) By Proposition 6, for all  $\phi_i, \phi_j \in \Phi$ ,  $\text{Models}(\phi_i) \cap \text{Models}(\phi_j) = \emptyset$ . For each  $\phi_i \in \Phi$ , let  $\text{Models}_X(\phi_i) = \{(G_i, E) \mid E \in \text{Extensions}_X(G_i)\}$ . Therefore, for each  $G_i \sqsubseteq G$ ,  $G_i \models_X \phi_i$ , and for all  $\phi_j \in \Phi$ , if  $\phi_i \neq \phi_j$ , then  $G_i \not\models_X \phi_j$ . Therefore, for each  $G_i, G_j \sqsubseteq G$ ,  $\text{Dividers}(G_i) \cap \text{Dividers}(G_j) = \emptyset$ . Therefore, by Definition 10,  $\Phi$  is pairwise disjoint for  $G$  with respect to  $X$ .

Since,  $\Phi$  is exhaustive and pairwise disjoint for  $G$  with respect to  $X$ ,  $[P(\phi_1), \phi_1; \dots; P(\phi_k), \phi_k]$  is an argumentation lottery for  $G$  with respect to  $X$ .  $\square$

**Example 23.** Let  $G$  be the argument graph in Figure 3. Also let  $P$  a probability distribution on  $G$ . Applying Proposition 15, if  $\{\phi_1, \dots, \phi_7\}$  is the set of the following formulae, then  $[P(\phi_1), \phi_1; \dots; P(\phi_7), \phi_7]$  is an argumentation lottery for  $G$  with respect to grounded semantics.

$$\begin{aligned}
\phi_1 &= \neg\alpha \wedge \neg\beta \wedge \blacklozenge(\alpha \wedge \beta) \wedge \blacksquare_{\{\alpha\}}\alpha \wedge \blacksquare_{\{\beta\}}\beta & \phi_5 &= \alpha \wedge \blacksquare\neg\beta \\
\phi_2 &= \alpha \wedge \neg\beta \wedge \blacklozenge\beta \wedge \neg\xi_\beta & \phi_6 &= \blacksquare\neg\alpha \wedge \beta \\
\phi_3 &= \neg\alpha \wedge \beta \wedge \blacklozenge\alpha \wedge \neg\xi_\alpha & \phi_7 &= \neg\alpha \wedge \neg\beta \wedge \neg\blacklozenge\alpha \wedge \neg\blacklozenge\beta \\
\phi_4 &= \alpha \wedge \beta
\end{aligned}$$

Now that we have defined the notion of an argumentation lottery and shown how they can be constructed in general, we will consider in the following section how they can be used by participants in argumentation. Note, we are primarily considered with artificial agents who have the computational capacity to evaluate the probability of outcomes by identifying the extensions of the subgraphs of a graph, and thereby have the computational ability to calculate the expected utility of a lottery.

## 6 Using Lotteries in Dialogical Argumentation

Argument lotteries are a general tool for evaluating the uncertainty of outcomes in argument graphs. In this section, we will discuss two particular application scenarios of argumentation lotteries in scenarios of dialogical argumentation, one for a passive audience of an argumentation dialogue, and one for active participants in argumentation dialogue. We will only briefly discuss the first application scenario and focus on the second.

### 6.1 Audience evaluation of argumentation

We believe that expected utility is a useful formal tool for an audience to judge argumentation. From the audience's perspective, we are interested in modelling how a member of the audience may evaluate some arguments. For example, a member of the audience of a political speech may listen to the arguments and counterarguments that the politician has presented, or a member of the audience of a debate may hear the arguments and counterarguments exchanged by the participants. In each case, an argument graph is produced. The member of the audience then may look at the arguments and the attacks and she may be uncertain whether some of the arguments should be included in the graph (perhaps some arguments are rephrasing of previously expressed arguments), and/or whether some of the attacks hold (perhaps the arguments are enthymemes, and she doubts that the enthymemes can be decoded so that it can be attacked by the given counterarguments).

In order to represent the uncertainty in the arguments and attacks, the member of the audience identifies a probability distribution over the subgraphs. With this probability

distribution, she can determine the probability that specific arguments are included or excluded according to specific semantics. Furthermore, by determining the expected utility of the corresponding argumentation lottery, she can determine the worth of the consequences of the debate to her in utility-theoretic terms.

**Example 24.** Consider Figure 1 with the probability distribution where  $P(G) = 0.1$ ,  $P(G_5) = 0.1$ ,  $P(G_8) = 0.1$ , and  $P(G_9) = 0.7$ . Also consider the following formulae

$$\phi_1 = \alpha \wedge \neg\beta \wedge \gamma \quad \phi_2 = \alpha \wedge \neg\beta \wedge \neg\gamma \quad \phi_3 = \neg\alpha \wedge \neg\beta \wedge \gamma \quad \phi_4 = \neg\alpha \wedge \neg\beta \wedge \neg\gamma \quad \phi_5 = \beta$$

The above set of formulae is exhaustive and pairwise disjoint for  $G$  with respect to grounded semantics. Let the utility function be  $U(\phi_1) = 10$ ,  $U(\phi_2) = 5$ ,  $U(\phi_3) = 5$ ,  $U(\phi_4) = 0$ , and  $U(\phi_5) = -10$ . Observe that  $U$  favours  $\alpha$  and/or  $\gamma$ , but not  $\beta$ , in our grounded extension. Therefore the expected utility is  $= 3.5$ .

So what we have presented so far in this paper already supports our member of the audience. She constructs the graph  $G$  from the arguments and attacks presented, she identifies her probability distribution, and she determines the probabilistic outcomes. Moreover, by determining the expected utility of the lottery corresponding to an argument graph, she can determine what it is likely to get from the argument graph. In other words, given the outcomes of interest to the member of the audience, the utility function used for those outcomes, and the probability distribution it used to express the uncertainty about which is the actual argument graph, the member of the audience has an evaluation of how good the argument graph is for her.

## 6.2 Maximizing expected utility in argumentation

We now consider the argumentation lottery from the perspective of the participant. When an agent presents an argument  $\alpha$ , this can be viewed as a lottery by the agent since there is uncertainty about whether  $\alpha$  will be included or excluded from the viewpoint of the audience according to some semantics. If the agent's probability distribution is  $P$  then it can assess the outcome of presenting argument  $\alpha$  by evaluating the lottery

$$[P_X(\alpha), \alpha; P_X(\neg\alpha), \neg\alpha]$$

with respect to its utility function  $U$ . Now suppose the agent has a choice of arguments to present say  $\alpha_1$ ,  $\alpha_2$ , or  $\alpha_3$ , and for each of the arguments  $\alpha_i \in \{\alpha_1, \alpha_2, \alpha_3\}$ , if  $\alpha_i$  is presented, the agent is unsure whether  $\alpha_i$  will be in, out or undecided from the viewpoint of the audience according to, e. g., grounded semantics. So each  $\alpha_i$  is an option for an action with an associated lottery  $L_i$ , respectively, of the above form. Given these lotteries, we can choose the argument  $\alpha_i$  that maximizes expected utility. In the same way, suppose that an agent has a choice of which *sets of arguments* to present, say  $A_1$ ,  $A_2$ , or  $A_3$ , but the agent is concerned about another argument  $\beta$  in  $G$ . For instance, in a dialogue,  $\beta$  may have been given earlier, and the agent wants to know which would be the best arguments to add at this stage in order to get a particular outcome concerning  $\beta$ .

### 6.2.1 Basic framework for selecting contributions

We organize the ideas outlined above as follows. A dialogue (such as a discussion, debate, etc) is a sequence of moves involving two or more agents. We assume two agents for simplicity here and we assume the agents take it in turns. We also assume each move is a set of zero or more arguments and zero or more attacks. We therefore assume an abstract notion of a dialogue.

**Definition 15.** A **dialogue** is a sequence of moves  $D = [M^1, \dots, M^n]$  where each  $M^i \in \{M^1, \dots, M^n\}$  is a tuple of the form  $(R, S)$  where  $R$  is a set of arguments and  $S$  is a set of attacks. We use a function  $D$  as an equivalent representation of a dialogue. So for each  $i \in \{1, \dots, n\}$ ,  $D(i) = M^i$ .

In this paper, we are not concerned with the broader issues concerning dialogues such as protocols, results of information exchange, etc. [KJRM09, BA09, CS10]. Rather, we are concerned with the specific issue of strategic move selection [Thi14a], where at some point in the dialogue it is the turn of one of the agents to make the next move (i. e. to make the next contribution). Let the **proponent** be the agent who wants to make the next contribution to a dialogue, and let the **audience** be the person or people who will hear (or read) that contribution to the dialogue. So we are concerned with how the next move  $M^{n+1}$  can be added to the dialogue (assuming that the dialogue has not terminated at step  $n$ ).

We assume that the proponent has an argument graph  $G_P$  in mind (i. e. the proponent graph) that contains all the arguments and attacks that s/he is aware of. This contains all the arguments and attacks that the proponent may use in his/her contribution. So the proponent graph may include arguments and/or attacks that are only there because the audience may have them in their graph. We also assume that the proponent has a model of the audience  $G_A$  (i. e. the audience graph) that contains all the arguments and attacks that s/he thinks the audience is aware of. We assume that  $G_A \sqsubseteq G_P$ .

**Example 25.** Consider Figure 1. If  $G$  is the proponent graph, then any of  $G_1$  to  $G_{18}$  could be selected as the audience graph.

According to Definition 15, each move (i. e. contribution to the dialogue) is a pair  $(R, S)$  where  $R$  is a set of arguments and  $S$  is a set of attacks. When a proponent is choosing its move, it may have a number of options from which to pick its move. We assume the following constraint on options, and so we assume the move is a set of arguments and attacks obtained from the proponent graph.

**Definition 16.** An **option** for the proponent is a pair  $(R, S)$  where  $R \subseteq \text{Nodes}(G_P)$  and  $S \subseteq \text{Arcs}(G_P)$ .

**Example 26.** Continuing Example 25, any tuple in  $\wp(\{\alpha, \beta, \gamma\}) \times \wp(\{(\alpha, \beta), (\beta, \alpha), (\gamma, \beta)\})$  is an option.

Suppose the proponent can choose between a number of options for the next contribution. The proponent needs to choose one of them. For this, it needs to consider the effect of that contribution on its audience model. Let the options be  $C_1, \dots, C_k$ . Now suppose  $G_A$  is the argument graph that includes all the possible arguments and

attacks that the intended audience may currently entertain (i. e. the audience graph). For each option for a contribution  $C_i \in \{C_1, \dots, C_k\}$ , we let  $G_A + C_i$  be the argument graph obtained by augmenting  $G_A$  with the arguments and attacks in  $C_i$ .

**Definition 17.** Let  $G$  be an argument graph and let  $C$  be an option for the proponent. The **augmentation** of  $G$  by  $C$ , denoted  $G + C$ , is the following graph.

$$(\text{Nodes}(G) \cup \text{Nodes}(C), \text{Arcs}(G) \cup \text{Ok}(G, C))$$

where  $\text{Ok}(G, C) = \{(\alpha, \beta) \in \text{Arcs}(C) \mid \alpha \in \text{Nodes}(G) \cup \text{Nodes}(C) \text{ and } \beta \in \text{Nodes}(G) \cup \text{Nodes}(C)\}$ .

In augmentation, we need to check when adding arcs from the option that the source and destination nodes are in the augmented graph (i. e. the source and destination nodes need to be in  $\text{Nodes}(G) \cup \text{Nodes}(C)$ ). We do this check using the  $\text{Ok}(C)$  function.

**Example 27.** Consider the audience graph  $G_A$  where  $\text{Nodes}(G_A) = \{\alpha\}$  and  $\text{Arcs}(G_A) = \emptyset$ . Let  $C$  be the contribution where  $\text{Nodes}(C) = \{\beta\}$  and  $\text{Arcs}(C) = \{(\beta, \alpha), (\gamma, \alpha)\}$ . So  $G_A + C$  is the graph where  $\text{Nodes}(G_A + C) = \{\alpha, \beta\}$  and  $\text{Arcs}(G_A + C) = \{(\beta, \alpha)\}$ .

So far, we assumed that the proponent knows the exact nature of the audience graph. We now turn to the probabilistic setting where the proponent is uncertain to what the exact audience graph is. The only remaining assumption we entertain is that the actual audience graph is a subgraph of the proponent graph. More formally, let  $P : \text{Sub}(G) \rightarrow [0, 1]$  be a probability distribution such that  $P(G')$  is the degree of belief of the proponent that the audience graph is  $G' \sqsubseteq G$ . We first consider the issue of what happens with  $P$  when the proponent discloses its selected contribution  $C$ . There are various ways for how we could calculate this a posteriori probability distribution. We start with a simple method below, which corresponds to classical Bayesian conditioning [Pea88], and then later consider some alternatives. For that, we assume for now that the audience will believe in the contribution  $C$ . That is, subgraphs  $G'' \sqsubset G$  that do not contain  $C$  should have zero probability after disclosing  $C$ .

**Definition 18.** Let  $P$  be a probability distribution over  $G$  and let  $C$  be an option. We define the **simple redistribution** of  $P$  wrt.  $C$ , denoted  $P'$ , for all  $G' \sqsubseteq G$  via

$$P'(G') = \begin{cases} \sum_{G''=G'+C} P(G'') & \text{if } C \sqsubseteq G' \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

**Example 28.** Consider Figure 1. Let the proponent graph be  $G$ , and let  $P(G_{16}) = 0.8$ ,  $P(G_{18}) = 0.2$ , and  $P(G') = 0$  for all other subgraphs. Now consider contributions  $C_1$  where  $\text{Nodes}(C_1) = \{\alpha\}$  and  $\text{Arcs}(C_1) = \{(\alpha, \beta)\}$  and  $C_2$  where  $\text{Nodes}(C_2) = \{\gamma\}$  and  $\text{Arcs}(C_2) = \{(\gamma, \beta)\}$ . So  $G_{16} + C_1$  is  $G_9$ ,  $G_{18} + C_1$  is  $G_{15}$ ,  $G_{16} + C_2$  is  $G_{13}$ , and  $G_{18} + C_2$  is  $G_{17}$ . Therefore, we get  $P_1$  as the simple redistribution of  $P$  wrt.  $C_1$  and  $P_2$  as the simple redistribution of  $P$  wrt.  $C_2$  with.

$$P_1(G_9) = 0.8 \quad P_1(G_{15}) = 0.2 \quad P_2(G_{13}) = 0.8 \quad P_2(G_{17}) = 0.2$$

all  $P_1(G') = P_2(G') = 0$  for all other subgraphs, respectively.

**Proposition 16.** *Let  $P$  be a probability distribution over  $G$  and let  $C$  be an option. If  $P'$  is the simple redistribution of  $P$  wrt.  $C$ , then  $P'$  is a probability distribution over  $G$ .*

*Proof.* We have to show that  $\sum_{G':G' \sqsubseteq G} P'(G') = 1$ . Note that in (3) only those subgraphs  $G' \sqsubseteq G$  receive a positive probability in  $P'$  for which  $C \sqsubseteq G'$  holds. Therefore we have

$$\begin{aligned} \sum_{G':G' \sqsubseteq G} P'(G') &= \sum_{G':C \sqsubseteq G' \sqsubseteq G} \sum_{G'':G'=G''+C} P(G'') \\ &= \sum_{G'':C \sqsubseteq G''+C \sqsubseteq G} P(G'') \\ &= \sum_{G'':G'' \sqsubseteq G} P(G'') \\ &= 1 \end{aligned}$$

□

We can now define the argumentation lottery for selecting a move. Essentially, it is Definition 14 with the probability distribution obtained for the candidate.

**Definition 19.** *Let  $G_P$  be the proponent graph, let  $P$  be a probability distribution over  $G$ , and let  $P'$  be the simple probability distribution of  $P$  wrt. an option  $C$ . Also let  $\{\phi_1, \dots, \phi_k\} \subseteq \mathcal{L}_\Pi$  be exhaustive and pairwise disjoint for  $G$  wrt.  $X$ . The argumentation lottery  $[P'_X(\phi_1), \phi_1; \dots; P'_X(\phi_k), \phi_k]$  is the **candidate lottery** for  $C$  wrt.  $G$  and semantics  $X$ .*

**Example 29.** *Consider the proponent graph  $G_P$  in Figure 5. Assume that the outcomes of interest are  $\alpha$  and  $-\alpha$  where  $U(\alpha) = 10$  and  $U(-\alpha) = -10$  and that the a priori probability distribution  $P$  of the proponent on the audience graphs is as depicted in Table 1a (the subgraphs  $G_1, G_2, G_3, G_4$  of  $G_P$  have the probability indicated in Table 1a, all other subgraphs have zero probability). Hence, the expected utility of the argumentation lottery on  $G_P$  and  $P$  is  $(0.2 \times 10) + (0.8 \times (-10)) = -6$ .*

- *Now consider the contribution  $C^1 = (\{\delta\}, \{(\delta, \beta)\})$ , which yields the a posteriori probability distribution  $P_1$  depicted in Table 1b. Hence, the expected utility of the argumentation lottery on  $G_P$  and  $P_1$  is  $(0.8 \times 10) + (0.2 \times (-10)) = 6$ .*
- *Finally consider the contribution  $C^2 = (\{\varepsilon\}, \{(\varepsilon, \gamma)\})$ , which yields the a posteriori probability distribution  $P_2$  depicted in Table 1c. Hence, the expected utility of the argumentation lottery on  $G_P$  and  $P_2$  is  $(0.3 \times 10) + (0.7 \times (-10)) = -4$ .*

*So both contributions increase the expected utility, with  $C^1$  being the contribution that maximizes utility.*

To show how the candidate lottery could be used in a “real-world” dialogue, we consider the following example where one agent needs to choose a good argument to present to the other agent for persuasion.

**Example 30.** *Consider the situation where a couple is looking to buy a new car. Suppose the proponent graph  $G_P$  contains the following three arguments, and one attack from  $\beta$  to  $\alpha$ .*



So by harnessing a probabilistic argument graph, and an argumentation lottery, a participant can optimize its choice of actions in argumentation. This can be for example in monological argumentation (e. g. in a speech or a written article) or dialogical argumentation (e. g. a discussion or debate) when a participant wants to present arguments and/or counterarguments in order to convince the audience that some particular arguments should be accepted and some should be rejected.

In our presentation in this section, we have left the details of the dialogue open so that the approach can be adapted to specific dialogue systems (for instance for persuasion [Pra06]). We have also left open how the utility function would be obtained, as this depends on the nature, functionality and resources of the artificial agent. Nonetheless, if the agent has some form of agenda with regard to argumentation, it should be possible to extract a utility function to represent the priorities of that agenda.

In the basic framework, we used the simple redistribution (Definition 18) to incorporate a contribution into the probabilistic model the proponent has about the audience. The simple redistribution is an intuitive and simple way to redistribute the probability, but there are alternatives. We present two alternatives, namely sticky redistribution (which we will illustrate with an application in supporting value-based argumentation) and rough redistribution (which we will illustrate with an application in handling enthymemes), in the next two subsections.

### 6.2.2 Sticky redistribution

Sticky redistribution is a generalization of the simple redistribution where not all the probability mass is redistributed from  $G'$  to  $G' + C$ . Rather, there is some “stickiness”, and so some probability mass remains with  $G'$ . This models the idea that the audience might not fully believe in the contribution the proponent is making. For this, we introduce a stickiness coefficient which specifies the proportion of the probability mass that is reassigned.

**Definition 20.** *Let  $P$  be a probability distribution over a graph  $G$ , and let  $C$  be an option. We define the **sticky redistribution** of  $P$  wrt.  $C$ , denoted  $P'$ , for all  $G' \sqsubseteq G$  via*

$$P'(G') = \begin{cases} \mu \sum_{G''=G'+C} (P(G'') + (1-\mu)P(G')) & \text{if } C \sqsubseteq G' \\ (1-\mu)P(G') & \text{otherwise} \end{cases} \quad (4)$$

where  $\mu \in [0, 1]$  is the **stickiness coefficient**.

**Example 31.** *Consider Figure 1 and assume  $P$  with  $P(G_{16}) = 0.8$ ,  $P(G_{18}) = 0.2$ , and  $P(G') = 0$  for all other subgraphs. Now consider contributions  $C_1$  where  $\text{Nodes}(C_1) = \{\alpha\}$  and  $\text{Arcs}(C_1) = \{(\alpha, \beta)\}$  and  $C_2$  where  $\text{Nodes}(C_2) = \{\gamma\}$  and  $\text{Arcs}(C_2) = \{(\gamma, \beta)\}$ . So  $G_{16} + C_1$  is  $G_9$ ,  $G_{18} + C_1$  is  $G_{15}$ ,  $G_{16} + C_2$  is  $G_{13}$ , and  $G_{18} + C_2$  is  $G_{17}$ . Let  $\mu = 0.5$  be the stickiness coefficient. Therefore, we get  $P_1$  as sticky redistribution of  $P$  wrt.  $C_1$  and  $P_2$  as sticky redistribution of  $P$  wrt.  $C_2$  with*

$$\begin{array}{cccc} P_1(G_{16}) = 0.4 & P_1(G_{18}) = 0.1 & P_1(G_9) = 0.4 & P_1(G_{15}) = 0.1 \\ P_2(G_{16}) = 0.4 & P_2(G_{18}) = 0.1 & P_2(G_{13}) = 0.4 & P_2(G_{17}) = 0.1 \end{array}$$

Obviously, if we set the stickiness coefficient to  $\mu = 1$ , then the sticky redistribution is the same as the simple redistribution. At the other extreme, if we set the stickiness coefficient to  $\mu = 0$ , then the sticky redistribution returns the original distribution, and hence the augmented subgraph is ignored.

**Proposition 17.** *Let  $P$  be a probability distribution over  $G$ , and let  $C$  be an option. If  $P'$  is the sticky redistribution of  $P$  wrt.  $C$ , then  $P'$  is a probability distribution over  $G$ .*

*Proof.* We have to show that  $\sum_{G':G' \sqsubseteq G} P'(G') = 1$ :

$$\begin{aligned}
\sum_{G':G' \sqsubseteq G} P'(G') &= \sum_{G':C \sqsubseteq G' \sqsubseteq G} P(G') + \sum_{G':C \not\sqsubseteq G' \sqsubseteq G} P(G') \\
&= \sum_{G':C \sqsubseteq G' \sqsubseteq G} \left( \mu \sum_{G'':G''=G'+C} P(G'') + (1-\mu)P(G') \right) + \sum_{G':C \not\sqsubseteq G' \sqsubseteq G} (1-\mu)P(G') \\
&= \sum_{G':C \sqsubseteq G' \sqsubseteq G} \mu \sum_{G'':G''=G'+C} P(G'') + \sum_{G':C \sqsubseteq G' \sqsubseteq G} (1-\mu)P(G') + \sum_{G':C \not\sqsubseteq G' \sqsubseteq G} (1-\mu)P(G') \\
&= \sum_{G'':C \sqsubseteq G''+C \sqsubseteq G} \mu P(G'') + \sum_{G':G' \sqsubseteq G} (1-\mu)P(G') \\
&= \mu \sum_{G'':G'' \sqsubseteq G} P(G'') + (1-\mu) \sum_{G':G' \sqsubseteq G} P(G') \\
&= \mu + (1-\mu) = 1
\end{aligned}$$

□

We now consider an application of the sticky redistribution for enhancing value-based argumentation frameworks [Ben03]. To introduce this extension of abstract argumentation, we use the following examples.

**Example 32.** (Example taken from [Ben03]) Consider the following arguments concerning the theft by Hal of insulin from Carla's House because he has lost his through no fault of his own. We give the argument graph below.

- $\alpha$  = “Hal is justified because a person can use other people's property to save a life” (LIFE)
- $\beta$  = “It is wrong to infringe the property rights of others” (PROPERTY)
- $\gamma$  = “If Hal compensates Carla, then property rights have not been infringed” (PROPERTY)

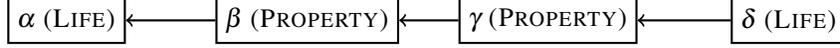
So  $\{\alpha, \gamma\}$  is the preferred extension, but it may appear unjust to accept  $\alpha$  based on the value of (LIFE) using the argument  $\gamma$  which is based on the value of (PROPERTY).



**Example 33.** (Example taken from [Ben03]) We now extend Example 32 by considering the extra argument  $\delta$ , and the resulting argument graph, below.

- $\delta = \text{“If Hal were too poor to compensate, then he should be allowed to take the insulin (LIFE)”}$

So  $\{\beta, \delta\}$  are the acceptable, but it may appear even more unjust for  $\alpha$  to now be defeated taking  $\delta$  into account.



A **value-based argument system** extends an abstract argument system with a set of values  $V$ , a function  $val$  that assigns a value to each argument, a set of audiences  $\Pi$ , and a set of preference relations  $\preceq_\pi \subseteq V \times V$  where  $\pi \in \Pi$  and  $A \preceq_\pi B$  denotes that  $A$  is more or equally preferred for  $\pi$  than  $B$ .

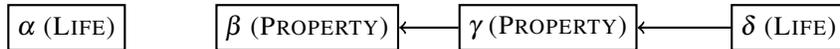
Definitions for abstract argumentation are specialized to those for an audience  $\pi$  and its preferred values. So the usual definitions for abstract argumentation are revised to use the following definitions for an audience.

- $A \text{ attacks}_\pi B$  iff  $A$  attacks  $B$  and not  $B \preceq_\pi A$ .
- A set of arguments  $S$  is conflict free for  $\pi$  iff there are no arguments  $A, B \in S$  such that  $A \text{ attacks}_\pi B$
- A set of arguments  $S$  defends an argument  $A$  for  $\pi$  iff if  $B \text{ attacks}_\pi A$ , then  $S \text{ attacks}_\pi B$

**Example 34.** (Example taken from [Ben03]) Returning to Example 32, let  $V = \{\text{PROPERTY, LIFE}\}$ ,  $val(\alpha) = \text{LIFE}$ ,  $val(\beta) = \text{PROPERTY}$ ,  $val(\gamma) = \text{PROPERTY}$ ,  $\Pi = \{\pi_1\}$ , and  $\text{LIFE} \preceq_{\pi_1} \text{PROPERTY}$ . So the argument graph becomes the following, with the desired effect that  $\alpha$  is in the extension unconditionally.



**Example 35.** (Example taken from [Ben03]) Returning to Example 33, let  $V = \{\text{PROPERTY, LIFE}\}$ ,  $val(\alpha) = \text{LIFE}$ ,  $val(\beta) = \text{PROPERTY}$ ,  $val(\gamma) = \text{PROPERTY}$ ,  $\Pi = \{\pi_1\}$ , and  $\text{LIFE} \preceq_{\pi_1} \text{PROPERTY}$ . So the argument graph becomes the following, with the desired effect that both  $\alpha$  and  $\delta$  are accepted.



So value-based argumentation deletes attacks where the value of the attacking argument is greater than or equal to the value of the attacked argument. So given an argument graph, value-based argumentation makes a binary decision on whether to keep or delete each arc. We can refine this by using probabilistic argumentation so that there is a probability that the arc is kept and a probability that it is deleted. So when the audience is presented an argument  $\beta$  that attacks an argument  $\alpha$ , and the audience is thought to value  $\beta$  higher than  $\alpha$ , then the contribution has a probability of being made. We can use sticky redistribution for this. We illustrate this in the following example.

**Example 36.** We use the arguments in Example 32 to construct the argument graph in a number of steps of the dialogue. Let us suppose one agent is making all the non-empty contributions (the odd moves), and the other agent is just making an empty contribution when it is its turn (the even moves). The contributions are listed below.

- $C_1 = (\{\alpha, \beta, \gamma\}, \emptyset)$
- $C_2 = (\emptyset, \emptyset)$
- $C_3 = (\emptyset, \{\beta, \alpha\})$
- $C_4 = (\emptyset, \emptyset)$
- $C_5 = (\emptyset, \{\gamma, \beta\})$

For the agent making the odd moves, let us suppose at the start it may assume the actual audience graph is the empty graph (thus receiving probability 1 and all other subgraphs receiving probability zero). Now, we consider the moves in sequence, and the effect they have on the audience model. Let us assume that we use the simple redistribution for all moves except for when the attack  $(\alpha, \beta)$  is added (we use a stickiness coefficient of 0.2). This amounts to the idea that  $\text{val}(\alpha) \preceq_\pi \text{val}(\beta)$  very likely holds according to the audience.

- After the move  $C_1$ , a simple redistribution results in the audience model being the graph containing the arguments, and this graph has probability 1.

$$G_1^1 = \alpha \quad \beta \quad \gamma$$

- After the move  $C_2$ , there is no update.

$$G_1^2 = \alpha \quad \beta \quad \gamma$$

- After the move  $C_3$ , it is the case that  $\text{val}(\alpha) \preceq_\pi \text{val}(\beta)$ , and so a sticky redistribution results in the following two argument graphs, where  $P(G_1^3) = 0.2$  and  $P(G_2^3) = 0.8$

$$G_1^3 = \alpha \leftarrow \beta \quad \gamma \qquad G_2^3 = \alpha \quad \beta \quad \gamma$$

- After the move  $C_4$ , there is no update.

$$G_1^4 = \alpha \leftarrow \beta \quad \gamma \qquad G_2^4 = \alpha \quad \beta \quad \gamma$$

- After the move  $C_5$ , it is not the case that  $\text{val}(\beta) \preceq_\pi \text{val}(\gamma)$ , and so a simple redistribution results in the following two additional argument graphs,

$$G_1^5 = \alpha \leftarrow \beta \leftarrow \gamma \qquad G_2^5 = \alpha \quad \beta \leftarrow \gamma$$

The final probabilistic assignment in the dialogue is  $P(G_1^5) = 0.04$ ,  $P(G_2^5) = 0.16$ ,  $P(G_1^4) = 0.16$  and  $P(G_2^4) = 0.64$ . Therefore  $P_X(\alpha) = 1$  (for every semantics  $X \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$ ).

By using sticky redistribution, we get a finer grained way of updating the proponent model. It allows for the modelling of an audience that does not completely accept an update. There may be various reasons for why the audience might be reluctant to accept an update. We give one example that refines value-based argumentation. Whilst we have not considered selecting moves in this subsection, it is clear that using the sticky redistribution as opposed to simple redistribution can have a substantial influence on the expected utility of any option for a contribution. Note, the definition for stickiness is illustrative for a range of possible definitions. The definition applies stickiness to the whole contribution, but it would be possible to do it for part of a contribution according to some criteria. If we did this, we could for instance deal with Example 36 with all the arguments and attacks in the same contribution (instead of splitting them into one contribution for the arguments, and one contribution for each attack).

### 6.2.3 Rough redistribution

In general a contribution is a pair  $(R, S)$  where  $R$  is a set of arguments, and  $S$  is a set of attacks. However, often in real argumentation, the attacks are inferred from the arguments being presented. The arguments are presented explicitly by the proponent, and the audience has to determine whether an attack holds between any pair of arguments. This option treats arguments as enthymemes, and so there is uncertainty as whether the audience decodes them correctly. If the audience decodes them incorrectly, then the attack relationships may be different to those assumed by the proponent.

To address this need, we introduce (in the next definition) a form of redistribution that considers the possible subgraphs that can be formed from the model of the audience. For this, we start with the definition of the function  $\text{Super}(G, G', C)$  which takes a proponent graph  $G$ , a possible audience graph  $G'$ , and the option  $C$ , and returns the subset of arcs that appear in  $G$  that either involve an attack between the arguments in  $G'$  and the arguments in  $C$ , or involve an attack between the arguments in  $C$ . These are the attacks for which there is uncertainty as to whether the audience regards them as attacks. In other words, the audience may only regard some subset (even the empty set) as being attacks involving the arguments in  $C$ .

Then starting with a proponent graph  $G$ , and the option  $C$ , rough distribution assigns mass to the subgraphs  $G' \sqsubseteq G$  by considering each  $G''$ , where  $G'' + C \subseteq G'$  is a possible audience graph. For each such  $G''$ ,  $\text{Super}(G, G'', C)$  is calculated, and each subset  $C'$  of  $\text{Super}(G, G'', C)$ , that when added to  $G'' + C$  gives  $G'$ , is identified. The number of such  $C$  is multiplied by  $P(G')$  and normalized by dividing by the number of possibilities for  $C$  (i. e.  $2^{|\text{Super}(G, G'', C)|}$ ).

**Definition 21.** *Let  $P$  be a probability distribution over  $G$ , and let  $C$  be an option. Define*

$$\begin{aligned} \text{Super}(G, G', C) = & \{(\alpha, \beta) \in \text{Arcs}(G) \mid (\alpha \in \text{Nodes}(G') \text{ and } \beta \in \text{Nodes}(C)) \\ & \text{or } (\alpha \in \text{Nodes}(C) \text{ and } \beta \in \text{Nodes}(G')) \\ & \text{or } (\alpha \in \text{Nodes}(C) \text{ and } \beta \in \text{Nodes}(C))\} \end{aligned}$$

We define the **rough redistribution** wrt.  $P$ , denoted  $P'$ , for all  $G' \sqsubseteq G$  via

$$P'(G') = \begin{cases} \sum_{G'':G''+C \sqsubseteq G'} \frac{1}{2^{|\text{Super}(G,G'',C)|}} \sum_{C' \subseteq \text{Super}(G,G'',C):G''+C+C'=G'} P(G'') & \text{if } C \sqsubseteq G' \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

**Example 37.** Consider Figure 3. Let  $G$  be the proponent graph, and let  $P$  be a probability distribution over  $G$  where  $P(G_4) = 0.8$ ,  $P(G_6) = 0.2$ , and  $P(G') = 0$  for all other subgraphs.

Now suppose that the option  $C$  is such that  $\text{Nodes}(C) = \{\beta\}$  and  $\text{Arcs}(C) = \emptyset$ . With this option, we calculate the super sets as follows.

- $\text{Super}(G, G, C) = \{(\alpha, \beta), (\beta, \alpha)\}$
- $\text{Super}(G, G_1, C) = \{(\alpha, \beta), (\beta, \alpha)\}$
- $\text{Super}(G, G_2, C) = \{(\alpha, \beta), (\beta, \alpha)\}$
- $\text{Super}(G, G_3, C) = \{(\alpha, \beta), (\beta, \alpha)\}$
- $\text{Super}(G, G_4, C) = \{(\alpha, \beta), (\beta, \alpha)\}$
- $\text{Super}(G, G_5, C) = \emptyset$
- $\text{Super}(G, G_6, C) = \emptyset$

Next we calculate  $P'(G')$  for each  $G' \sqsubseteq G$ . For this below, we only explicitly consider  $G'' \sqsubseteq G'$  such that  $P(G'') \neq 0$ . If  $P(G'') = 0$ , then it cannot contribute to  $P'(G')$ .

- For  $G$ , because of case (i) below,  $P'(G) = 1/|\text{Super}(G, G, C)| \times 0.8 = 1/4 \times 0.8 = 0.2$ .
  - i If  $G'' = G_4$ , then  $G'' + C \sqsubseteq G$ , and if  $C' = (\emptyset, \{(\alpha, \beta), (\beta, \alpha)\})$ , then  $G'' + C + C' = G$ .
  - ii If  $G'' = G_6$ , then there is no  $C' \subseteq \text{Super}(G, G_6, C)$  such that  $G'' + C + C' = G$ .
- For  $G_1$ , because of case (i) below,  $P'(G_1) = 1/|\text{Super}(G, G_1, C)| \times 0.8 = 1/4 \times 0.8 = 0.2$ .
  - i If  $G'' = G_4$ , then  $G'' + C \sqsubseteq G_1$ , and if  $C' = (\emptyset, \{(\alpha, \beta)\})$ , then  $G'' + C + C' = G_1$ .
  - ii If  $G'' = G_6$ , then there is no  $C' \subseteq \text{Super}(G, G_6, C)$  such that  $G'' + C + C' = G_1$ .
- For  $G_2$ , because of case (i) below,  $P'(G_2) = 1/|\text{Super}(G, G_2, C)| \times 0.8 = 1/4 \times 0.8 = 0.2$ .
  - i If  $G'' = G_4$ , then  $G'' + C \sqsubseteq G_2$ , and if  $C' = (\emptyset, \{(\beta, \alpha)\})$ , then  $G'' + C + C' = G_2$ .

ii If  $G'' = G_6$ , then there is no  $C' \subseteq \text{Super}(G, G_6, C)$  such that  $G'' + C + C' = G_2$ .

- For  $G_3$ , because of case (i) below,  $P'(G_3) = 1/|\text{Super}(G, G_3, C)| \times 0.8 = 1/4 \times 0.8 = 0.2$ .

i If  $G'' = G_4$ , then  $G'' + C \subseteq G_3$ , and if  $C' = (\emptyset, \emptyset)$ , then  $G'' + C + C' = G_3$ .

ii If  $G'' = G_6$ , then there is no  $C' \subseteq \text{Super}(G, G_6, C)$  such that  $G'' + C + C' = G_3$ .

- For  $G_4$ , since  $C \not\subseteq G_4$ ,  $P(G_4) = 0$ .
- For  $G_5$ , because of case (i) below,  $P'(G_5) = 1/|\text{Super}(G, G_5, C)| \times 0.8 = 1 \times 0.8 = 0.2$ .

i If  $G'' = G_6$ , then  $G'' + C \subseteq G_5$ , and if  $C' = (\emptyset, \emptyset)$ , then  $G'' + C + C' = G_5$ .

- For  $G_6$ , since  $C \not\subseteq G_6$ ,  $P(G_6) = 0$ .

To summarize, the rough redistribution results in the following distribution  $P'(G) = 0.2$ ,  $P'(G_1) = 0.2$ ,  $P'(G_2) = 0.2$ ,  $P'(G_3) = 0.2$ ,  $P'(G_4) = 0$ ,  $P'(G_5) = 0.2$ , and  $P'(G_6) = 0$ .

**Proposition 18.** Let  $P$  be a probability distribution over  $G$ , and let  $C$  be an option. If  $P'$  is the rough redistribution from  $P$  wrt.  $C$ , then  $P'$  is a probability distribution over  $G$ .

*Proof.* We have to show that  $\sum_{G': G' \subseteq G} P'(G') = 1$ . Note that in (5) only those subgraphs  $G' \subseteq G$  receive a positive probability in  $P'$  for which  $C \subseteq G'$  holds. Therefore we have

$$\begin{aligned}
& \sum_{G': G' \subseteq G} P'(G') \\
&= \sum_{G': C \subseteq G' \subseteq G} P'(G') \\
&= \sum_{G': C \subseteq G' \subseteq G} \sum_{G'': G'' + C \subseteq G'} \frac{1}{2^{|\text{Super}(G, G'', C)|}} \sum_{C' \subseteq \text{Super}(G, G'', C): G'' + C + C' = G'} P(G'') \\
&= \sum_{G': C \subseteq G' \subseteq G} \sum_{G'': G'' + C \subseteq G'} \sum_{C' \subseteq \text{Super}(G, G'', C): G'' + C + C' = G'} \underbrace{\frac{1}{2^{|\text{Super}(G, G'', C)|}} P(G'')}_{p(G'')}
\end{aligned}$$

Let us calculate now how often the term  $p(G'')$  appears in the above sum. This term appears whenever we have  $G'$  and  $C'$  with  $G'' + C + C' = G'$ . Note that there are  $2^{|\text{Super}(G, G'', C)|}$  different subsets of  $\text{Super}(G, G'', C)$  and for every subset  $C' \subseteq \text{Super}(G, G'', C)$  we have one appearance of  $p(G'')$  for the uniquely determined  $G' = G'' + C + C'$  (for fixed  $G''$ ). Therefore,  $p(G'')$  appears  $2^{|\text{Super}(G, G'', C)|}$ -times. Furthermore, note that  $p(G'')$  appears for every  $G'' \subseteq G$  this number of times, as for every  $G''$  there is

$G' = G'' + C$  satisfying the condition of the first sum. Therefore, the above equation simplifies to

$$\begin{aligned}
& \sum_{G':G' \sqsubseteq G} P'(G') \\
&= \sum_{G'':G'' \sqsubseteq G} 2^{|\text{Super}(G,G'',C)|} P(G'') \\
&= \sum_{G'':G'' \sqsubseteq G} 2^{|\text{Super}(G,G'',C)|} \frac{1}{2^{|\text{Super}(G,G'',C)|}} P(G'') \\
&= \sum_{G'':G'' \sqsubseteq G} P(G'') \\
&= 1
\end{aligned}$$

□

Real arguments (i. e. those presented by people in general) are normally enthymemes. We consider two types which we will refer to as implicit support enthymemes and implicit claim enthymemes. An implicit support enthymeme does not explicitly represent some of the premises for entailing its claim. So if  $\Delta$  is the set of premises explicitly given for an implicit support enthymeme, and  $\delta$  is the claim, then  $\Delta$  does not entail  $\delta$ , but there are some implicitly assumable premises  $\Delta'$  such that  $\Delta' \cup \Delta$  is a consistent set of formulae that entails  $\delta$ . An implicit claim enthymeme does not explicitly represent all of its claim. In the rest of this section, we will consider implicit support enthymemes.

For instance, for a claim that you need an umbrella today, a husband may give his wife the premise the weather report predicts rain. Clearly, the premise does not entail the claim, but it is easy for the wife to identify the common knowledge used by the husband (i. e. if the weather report predicts rain, then you need an umbrella today) in order to reconstruct the intended argument correctly.

In the following example, we show how we can use rough redistribution to capture aspects of the uncertainty of handling enthymemes.

**Example 38.** *Consider the following arguments in the context of a conversation late in the evening (example adapted from [SW95]). Suppose one agent has said  $\alpha$ , and the other agent has replied with  $\beta$ .*

- $\alpha =$  “You need a coffee”
- $\beta =$  “Coffee would keep me awake”

*There is ambiguity with  $\beta$ . Two possible interpretations of  $\beta$  are  $\beta_1$  and  $\beta_2$  below. Here,  $\beta_1$  appears to be consistent with  $\alpha$ , and whereas  $\beta_2$  appears to be inconsistent with  $\alpha$ .*

- $\beta_1 =$  “Coffee would keep me awake and I need to sleep now. Therefore, I don’t need a coffee.”
- $\beta_2 =$  “Coffee would keep me awake and I would like to stay awake longer. Therefore, I need a coffee.”

Now we consider these arguments in the context of the space of graphs in Figure 3. We consider the proponent argument graph  $G$  as given in the following table, with the audience model being given by  $P$  in the table. Now suppose that the option  $C$  is such that  $\text{Nodes}(C) = \{\beta\}$  and  $\text{Arcs}(C) = \emptyset$ . With this option, we calculate the super sets given in the table.

Subgraph	$G$	$G_1$	$G_2$	$G_3$	$G_4$
Structure	$\beta \rightarrow \alpha$	$\beta \quad \alpha$	$\alpha$	$\beta$	
$P$	0.1	0.5	0.4	0	0
$\text{Super}(G, G, C)$	$\{(\beta, \alpha)\}$	$\{(\beta, \alpha)\}$	$\{(\beta, \alpha)\}$		
$1/2^{ \text{Super}(G, G, C) }$	1/2	1/2	1/2	1	1

Next we calculate  $P'(G')$  for each  $G' \sqsubseteq G$ . For this below, we only explicitly consider  $G'' \sqsubseteq G'$  such that  $P(G'') \neq 0$ . If  $P(G'') = 0$ , then it cannot contribute to  $P'(G')$ .

- For  $G$ , because of (i) to (iv) below,  $P'(G) = (2 \times (1/2^{|\text{Super}(G, G, C)|} \times 0.1)) + (1/2^{|\text{Super}(G, G_1, C)|} \times 0.5) + (1/2^{|\text{Super}(G, G_2, C)|} \times 0.4) = (2 \times (1/2 \times 0.1)) + (1/2 \times 0.5) + (1/2 \times 0.4) = 0.75$ .
  - i If  $G'' = G$ , then  $G'' + C \sqsubseteq G$ , and if  $C' = (\emptyset, \emptyset)$ , then  $G'' + C + C' = G$ .
  - ii If  $G'' = G$ , then  $G'' + C \sqsubseteq G$ , and if  $C' = (\emptyset, \{\beta, \alpha\})$ , then  $G'' + C + C' = G$ .
  - iii If  $G'' = G_1$ , then  $G'' + C \sqsubseteq G$ , and if  $C' = (\emptyset, \{\beta, \alpha\})$ , then  $G'' + C + C' = G_1$ .
  - iv If  $G'' = G_2$ , then  $G'' + C \sqsubseteq G$ , and if  $C' = (\emptyset, \{\beta, \alpha\})$ , then  $G'' + C + C' = G_2$ .
- For  $G_1$ , because of (i) below,  $P'(G) = 1/2^{|\text{Super}(G, G_1, C)|} \times 0.5 = 0.25$ .
  - i If  $G'' = G_1$ , then  $G'' + C \sqsubseteq G_1$ , and if  $C' = (\emptyset, \emptyset)$ , then  $G'' + C + C' = G_1$ .
- For  $G_2$ , since  $C \not\sqsubseteq G_2$ ,  $P'(G_2) = 0$ .
- For  $G_3$ , because of (i) to (iii) below,  $P'(G_3) = 0$ .
  - i If  $G'' = G$ , then  $G'' + C \not\sqsubseteq G_3$ ,
  - ii If  $G'' = G_1$ , then  $G'' + C \not\sqsubseteq G_3$ ,
  - iii If  $G'' = G_2$ , then  $G'' + C \not\sqsubseteq G_3$ ,
- For  $G_4$ , since  $C \not\sqsubseteq G_4$ ,  $P'(G_4) = 0$ .

To summarize, the rough redistribution results in the following distribution  $P'(G) = 0.75$ , and  $P'(G_1) = 0.25$ . This means that after  $\beta$  is said, the audience graph (i. e. the graph for the hearer of  $\beta$ ) is mostly like to be  $G$ , and less likely to be  $G_1$ . Hence, the most likely interpretation is that the speaker of  $\beta$  is that they do not want an coffee.

Using rough redistribution, we can revise the uncertainty over the audience graph resulting from the addition of enthymemes. The revised probability distribution can then be analyzed to determine the uncertainty of attacks by enthymemes. This can then

be used to determine the probability of outcomes, and it can be used by an agent in an argumentation lottery to choose optimal dialogues.

In previous work, we have considered decoding of enthymemes in the context of logical arguments (i. e. arguments of the form  $\langle \Delta, \delta \rangle$  where  $\Delta$  is a set of formulae entailing  $\delta$ ) [Hun07, BH12], and in the context of probabilistic argumentation where we have considered how the possible interpretations of an enthymeme can be used to determine the probability distribution over the subgraphs of the argument graph. Using the proposal in the paper, i.e. rough redistributions, addresses a different problem (i. e. how to update a model of the user by a new dialogue move).

## 7 Experimental Evaluation

We implemented the concepts introduced in this paper and, in particular, the game setting described in the previous section in Tweety<sup>1</sup>, cf. [Thi14b]. Moreover, we conducted some empirical evaluation of the different variants of the lottery approach for move selection in order to validate its feasibility and performance compared to more simple approaches. In this section, we first give an overview on the scenario considered for our empirical evaluation. Afterwards we describe the concrete setup of the experiment and then report on our findings.

### 7.1 Overview

We consider a scenario where a single agent  $Ag$  is placed in front of an audience  $Aud$ . The agent  $Ag$  believes in some argument graph  $G_0$  and assigns utility to some formulae on the arguments in  $G_0$  with a utility function  $U$ . The audience  $Aud$  also believes in some argument graph  $G_1$ . The goal of  $Ag$  is to bring forward a subgraph  $C \sqsubseteq G_0$  such that the grounded extension of  $G_1 + C$  maximizes utility of  $Ag$ , i. e., the agent  $Ag$  could convince the audience  $Aud$  of its own opinion to a maximal degree. In order to accomplish this,  $Ag$  can make use of a probability distribution  $P_0$  on the subgraphs of  $G_0$  such that for every  $G' \sqsubseteq G_0$  the value  $P(G')$  is the belief of  $Ag$  in  $G' = G_1$ . The aim of our evaluation is to compare different approaches on how to select  $C$ .

Note that the dialogue in our scenario implements a direct argumentation protocol [TG10], i. e., the agent  $Ag$  brings forward  $C$  in one single step. Furthermore, the audience is assumed to be passive and will not bring forward arguments itself.

In order to simplify the scenario for our evaluation, we will assume  $G_1 \sqsubseteq G_0$ , i. e., the actual audience graph is a subgraph of the graph of  $Ag$ . This means that there are no arguments and attacks the agent  $Ag$  is unaware of. Note that this is rather a technical restriction than a restriction of the scenario. The probability distribution  $P_0$  of  $Ag$  may assign a low (or zero) probability to specific arguments/attacks which represents the intuition that  $Ag$  believes  $Aud$  is ignorant of these arguments/attacks. So instead of modeling that  $Ag$  is unaware of an argument  $\alpha$ , we model that  $Ag$  is unaware of  $Aud$  knowing  $\alpha$ . Although this is conceptually different, it does not make such a huge difference in our technical treatment.

<sup>1</sup><http://www.tweetyproject.org>

## 7.2 Setup

We will consider the following different variations on selecting the contribution  $C$  in the above described dialogue scenario:

**Random** Choose  $C \sqsubseteq G_0$  with uniform probability from all subgraphs of  $G_0$ .

**Utility-based** Choose  $C \sqsubseteq G_0$  such that  $C$  has maximal utility among all subgraphs of  $G_0$ .

**Lottery-simple** Choose  $C \sqsubseteq G_0$  such that the lottery of the simple redistribution of  $P_0$  with  $C$  has maximal expected utility wrt.  $U$ .

**Lottery-sticky- $\mu$**  This is the same as **Lottery-simple** but we use the sticky redistribution with the stickiness coefficient  $\mu$ .

**Lottery-rough** This is the same as **Lottery-simple** but we use the rough redistribution.

For the **Lottery-sticky- $\mu$**  instantiation we consider variations with  $\mu = 0.1, 0.3, 0.5, 0.7, 0.9$ . In total, we compare nine different approaches for move selection.

For the setup of the dialogue, we generated random connected graphs with up to 7 arguments using the Erdős-Rényi model<sup>2</sup>. We identified  $G_0$  with the selected benchmark graph, randomly generated the probability distribution  $P_0$  on the subgraphs of  $G_0$ , and randomly generated the utility function  $U$  (with all utilities in the range  $[0, 1]$ ). Then we sampled (wrt. to the probability distribution  $P_0$ ) some subgraph of  $G_0$  to be the actual subgraph  $G_1$  assigned to the audience. For each of the nine different approaches for move selection we determined the best contribution  $C$  and determined the utility of the grounded extension of  $G_1 + C$ .

We repeated the experiment with 45 times (15 different graphs with 3 repetitions each) and computed the average utility and standard deviation for each approach over these graphs. The implementation of the experiment and the used graphs can be found online<sup>3</sup>.

## 7.3 Results

The final results can be seen in Figure 6 which shows for each move selection strategy the average utility and standard deviation after executing the dialogue (larger numbers mean more successful dialogues). As expected, all our more sophisticated models outperform the base line approaches **random** and **utility-based**. Moreover, all lottery-based approaches also have a much smaller standard deviation, which is an indicator for their robustness wrt. the randomly generated instances, while the performance of the base line approaches **random** and **utility-based** strongly depend on the actual instance—which is evidenced by the large standard deviation—and thus are not generally applicable strategies. Interestingly, the strategy **lottery-rough**, although being the most complex of the lottery-based strategies, performs worst among those. For all

<sup>2</sup>Due to the complexity of some of our approaches we could not include larger graphs as this would have been infeasible. Ongoing research is about more effective algorithms for these approaches.

<sup>3</sup><http://mthimm.de/r/?r=tweety-lotteries>

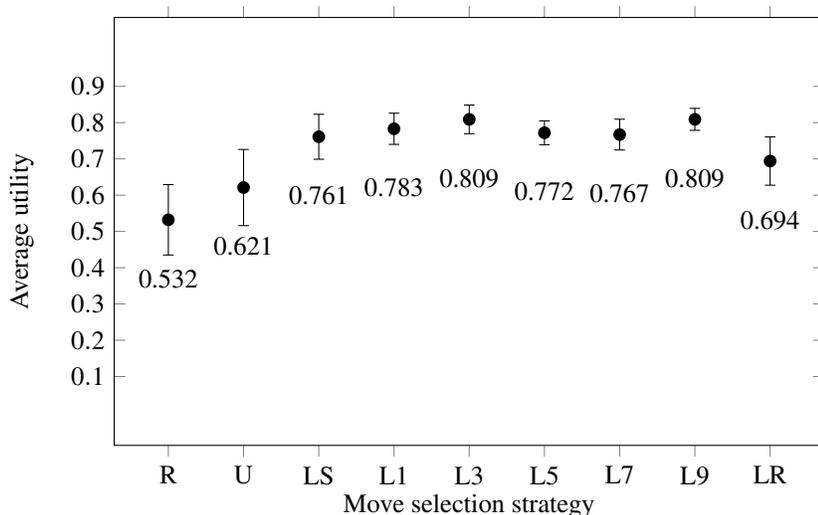


Figure 6: Performance of the random (R), utility-based (U), lottery-simple (LS), lottery-sticky-0.1 (L1), lottery-sticky-0.3 (L3), lottery-sticky-0.5 (L5), lottery-sticky-0.7 (L7), lottery-sticky-0.9 (L9), and lottery-rough (LR) strategies for move selection in comparison (with standard deviation)

tested values of  $\mu$  the strategy **lottery-sticky- $\mu$**  performed best with only minor differences between the different values of  $\mu$ . Also a little bit surprising, is that there is no linear or other obvious relationship between the value of  $\mu$  and the average utility. For  $\mu = 0.3$  and  $\mu = 0.9$  the strategies performed best in our experiments. It is, however, imaginable that these observations are due to the relatively small scope of this experimental evaluation.

## 8 Comparison with the literature

Logical encodings of argument graphs were proposed by Besnard and Doutre [BD04], and subsequently developed in a range of encodings for specific semantics (for example, using an ASP encoding [EGW10], using a CSP encoding [AD11], and using Łukasiewicz’s three valued logic [Dyr13]). In a generalization of the idea of encodings, Besnard et al. [BDH14] provide a model checking approach to encoding where a formula  $\theta$  is constructed to characterize the extensions of a graph  $G$  with respect to a semantics  $X$  such that the models of  $\theta$  (for example using classical logic) are isomorphic to the extensions of the graph  $G$  with respect to the semantics  $X$ .

Modal logics have also been harnessed for logical reasoning with argument graphs. In [CG09], modal logic is used to encode the constraints on labellings that can be obtained from an argument graph according to particular semantics, and this gives for instance a correspondence of the labellings of the argument graph and the possible worlds model of the argument graph. And in Grossi [Gro10], an argument graph is treated as

a Kripke frame, and modal operators are introduced to capture “there exists an attacking argument such that”, and “there is an argument such that”. Using this operators, notions of abstract argumentation such as acceptability, conflictfreeness, completeness, and various semantics can be axiomatized.

In an approach to reason about the specification of an argument graph, Villata et al. [VBG<sup>+</sup>13] introduce a satisfaction relation with a model being a pair  $(R, S)$  where  $R$  is a binary relation, and  $S$  is a subset of the set of arguments in the argument graph. Operators are introduced such as  $p \triangleright q$  to denote that the arguments in  $p$  attack the arguments in  $q$ ,  $p \oslash q$  to denote the arguments in  $p$  defend the arguments in  $q$ ,  $F(q)$  to denote the arguments in  $q$  are conflictfree,  $C(q)$  to denote the arguments in  $q$  are a complete extension, and  $P(q)$  to denote the arguments in  $q$  are a preferred extension. Using this language, axioms can be written concerning the structure and extensions of an argument graph that can be checked as to their satisfaction by a specific argument graph.

When comparing our proposal for a logic of dialectical outcomes with the logical encodings considered above, we see that our language involves constructs not considered in their encodings. In our language, we are able to write formulae that concern the membership of some or all extensions with respect to any semantics and with respect to some or all the subgraphs of the argument graph. This expressibility is needed for representing and reasoning with dialectical outcomes where there is uncertainty about the structure of the argument graph and the need to bring probability theory into the argumentation.

Our logic of dialectical outcomes was inspired by the proposal by Booth et al. [BKRvdT13] for a logical theory about dynamics in abstract argumentation. They assume a model is a labelling (as defined by Caminada [Cam06]), and then a satisfaction relation is  $\models_G \subseteq \mathcal{L}_G \times \mathcal{F}_G$  where  $\mathcal{L}_G$  is the set of labellings for the argument graph  $G$  and  $\mathcal{F}_G$  is the set of formulae of the language. The language has atoms for an argument being in, out, or undecided, and Boolean combinations can be obtained using disjunction and negation. Whilst our logic of dialectical outcomes does not have the distinction between out and undecided, we go beyond this proposal by having formulae that concern membership of some/all extensions, and with respect to some/all subgraphs.

There are a number of proposals for using probability theory in argumentation including the epistemic approach (e. g. [Thi12, Hun13b, HT14b]) and the constellations approach (e. g. [LON11, Hun12]). The epistemic approach is concerned with representing and reasoning with the uncertainty about the belief in individual arguments using a probability distribution over the subsets of the arguments in the graph (i. e.  $P : \wp(\text{Nodes}(G)) \rightarrow [0, 1]$  and so for an argument  $\alpha \in \text{Nodes}(G)$  the belief  $P(\alpha)$  in the argument is  $\sum_{X \in \wp(\text{Nodes}(G)) \text{ s. t. } \alpha \in X} P(X)$ ). In contrast, the constellations approach, which we also followed in our work here, is concerned with representing and reasoning with the uncertainty about the structure of the graph using a probability distribution over the subgraphs of the graph (i. e.  $P : \text{Sub}(G) \rightarrow [0, 1]$ ).

Most dialogue systems are aimed at providing protocols for dialogues (e. g. [Pra05, Pra06, FT11, CP12]), but strategies, in particular taking into account beliefs of the opponent are under-developed. Some proposals for strategies include [BH09, TG10, BA11, FT12], and see [Thi14a] for a review of strategies in multi-agent argumentation.

In order to operate strategically in an argumentation dialogue, it is desirable for the agent to handle the uncertainty concerning the argumentation including the knowledge, aims, and behaviour of the other agent(s). We can consider uncertainty from either the audience’s perspective or the participant’s perspective. In previous work, we have modelled audiences in terms of the beliefs and desires to assist a participant in choosing the most believable or the most desirable arguments to make [Hun04b, Hun04a]. However, these proposals do not consider the uncertainty associated with modelling the audience, and they do not provide a utility-theoretic framework.

Besides probability theory, other formalisms for representing uncertainty—such as possibility theory and fuzzy logic—have also been applied to computational models of argumentation [AP04, AP06, ACGS08, JCV08]. For example, in [AP04, ACGS08] classic-logical knowledge bases are augmented with uncertainty values to sentences and those are used to derive uncertainty values for arguments constructed from them. In order to determine defeat between arguments, these uncertainty values are taken into account and thus provide a more realistic approach for comparing arguments. In [AP04] this framework has been applied to a negotiation setting in multi-agent systems. In [JCV08] the attack relation in argument graphs is a fuzzy relation and, similarly to our probabilistic setting, this relation can be used to define uncertain notions of extensions and acceptability of arguments. We believe that the choice of the actual formalism for representing uncertainty (probability theory, possibility theory, fuzzy logic) is orthogonal to the work presented in this paper and that our notions of argumentation lottery and their use in dialogical argumentation could also be phrased using a different approach than probability theory.

Another approach to incorporate a “graded” assessment of the acceptability of arguments in argument graphs is provided by *ranking semantics*, see e. g. [BDKM16, GM15, AB13, TG14]. In this setting, the topology of the argument graph is exploited in order to obtain a more fine-grained ranking of arguments, from the most acceptable to the least acceptable ones. For example, in [GM15] an argument is more acceptable than another if the former is defended by more arguments than the latter (everything else being the same). These works are therefore concerned with a kind of *intrinsic* uncertainty in argumentation graphs, similar to the *epistemic* approach of probabilistic argumentation [Thi12, Hun13b, HT14b]. Acceptability rankings of arguments are derived from the topology of the plain argumentation graph. As we follow the constellations approach of probabilistic argumentation [LON11, Hun12] we deal with *extrinsic* uncertainty. The uncertainty on the acceptability of arguments stems from the uncertainty we may have in the topology of the argument graph. It could be worthwhile to consider combinations of both forms of uncertainty for future work.

Persuasion has also been considered through uncertainty modelling of the audience [OAL12]. This uncertainty is with respect to the structure of the graph, but there is no consideration of dialogues or strategies. In a proposal by Rienstra et al., a probabilistic model of the opponent has been used in a dialogue strategy allowing the selection of moves for an agent based on what it believes the other agent believes [RTO13]. This uncertainty concerns what the opposing agent is aware of rather than what it believes. In another approach to a probabilistic opponent model, the history of previous dialogues is used to estimate the arguments that an agent might put forward [HSM<sup>+</sup>13]. The method for updating the opponent model is of exponential complexity, and there

is no consideration of how utility theory could be employed.

Utility theory has been considered previously in argumentation to consider issues of manipulation (for example [RL08, RL09, RPRS08, ON10]) and to analyze an argument graph to quantify the degree of conflict [MT08]). Utility theory has also been used to determine the best move for an agent to make based in an argumentation dialogue using either minimax reasoning with a finite state machine [Hun13a] or analyzing probabilistic finite state machine [Hun14]. However, the proposal in [HT14a], and extended in this paper, is the first framework for using lotteries in argumentation.

## 9 Discussion

In this paper, we have investigated how probabilistic argumentation can be harnessed to formalize the notion of a lottery for argumentation. An argumentation lottery can be used to judge the expected utility of the outcomes in an argument graph. Furthermore, an argumentation lottery can be constructed for each of a number of possible contributions that can be made to a discussions, debate, etc. These lotteries can then be used to determine the contribution that maximizes the expected utility.

Therefore, an agent making a decision on what contribution (if any) to make in argumentation now has a formal tool to make the best choice. Primarily, we are concerned with using argumentation lotteries with artificial agents who have the computational capacity to represent and reason with probability distributions over the subgraphs and to undertake the expected utility calculations for each possible move. Given recent developments in algorithms and systems for computing with argument graphs (e. g. [CDG<sup>+</sup>15, NAD14]), we believe it is viable for artificial agents to undertake these calculations. Our empirical evaluations (reported in Section 7) with naive algorithms for identifying extensions supports this claim. Furthermore, specialised techniques for estimating the probability of extensions using Monte-Carlo techniques have been developed [FFP13].

The notion of a dialogue used in this paper (Definition 15) is general and covers the vast majority of proposals for dialogical argumentation in the literature on computational models of argument. We therefore believe that our proposal for argumentation lotteries and the use of redistribution is applicable for a wide variety of situations in computational argumentation.

In order to use our framework, it is necessary to obtain a probability distribution over the subgraphs of the argument graph. Such a distribution can be obtained empirically by learning from previous dialogues together with studies of classes of user. Some recent studies indicate the potential viability of an empirical approach for developing computational models of argument [CTO14, RK15]. An empirical approach to obtaining the probability distribution fits our aim of primarily supporting artificial agents to optimize their dialogue argumentation. Nonetheless, our framework is sufficiently general as to be used (in principle) by any agent with objective or subjective probabilities concerning dialectical outcomes.

In future work, we would like to further investigate how we can optimize strategies for dialogues, including consideration of how lotteries can be used for diverse types of dialogue, and further investigate possible definitions and properties of redistributions.

To support this, we would like to explore how our approach to using lotteries could draw further on established results in game theory. We would also like to investigate how our approach could be integrated with techniques for updating argument graphs to enforce particular outcomes (see for example [CFL10, Bau12]) since our notion of a contribution can be regarded as an update.

## Acknowledgements

The authors would like to thank the reviewers of our COMMA'14 submission and the reviewers of this version for raising some valuable research questions that we have attempted to address in this paper. This research was partly funded by EPSRC grant EP/N008294/1.

## References

- [ACGS08] T. Alsinet, C. I. Chesñevar, L. Godo, and G. R. Simari. A Logic Programming Framework for Possibilistic Argumentation: Formalization and Logical Properties. In *Fuzzy Sets and Systems*, 159(10):1208–1228, 2008.
- [AB13] L. Amgoud and J. Ben-Naim. Ranking-based Semantics for Argumentation Frameworks. In *Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM 2013)*, volume 8078 of *LNCS*, Springer, 2013.
- [AD11] L. Amgoud and C. Devred. Argumentation frameworks as constraint satisfaction problems. In *Proceedings of the International Conference on Scalable Approaches to Uncertainty Management*, volume 6929 of *LNCS*, pages 110–122. Springer, 2011.
- [AP04] L. Amgoud and H. Prade. Reaching agreement through argumentation: A possibilistic approach. In *Proceedings of the 9th International Conference on Principles of Knowledge Representation and Reasoning (KR'2004)*, pages 175–182.
- [AP06] L. Amgoud and H. Prade. Explaining qualitative decision under uncertainty by argumentation. In *Proceeding of the 21st AAAI Conference on Artificial Intelligence (AAAI'06)*.
- [BA09] E. Black and K. Atkinson. Dialogues that account for different perspectives in collaborative argumentation. In *Proceedings of the Eighth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'09)*, pages 867–874, Budapest, Hungary, 2009.
- [BA11] E. Black and K. Atkinson. Choosing persuasive arguments for action. In *In Proceedings of the 10th International Conference on Autonomous*

- Agents and Multi-Agent Systems (AAMAS11)*, pages 905–912. IFAA-MAS., 2011.
- [Bau12] R. Baumann. What does it take to enforce an argument? minimal change in abstract argumentation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, pages 127–132. IOS Press, 2012.
- [BCD07] T. Bench-Capon and P. Dunne. Argumentation in artificial intelligence. *Artificial Intelligence*, 171(10-15):619–641, 2007.
- [BD04] Ph. Besnard and S. Doutre. Checking the acceptability of a set of arguments. In *Proceedings of the International Workshop on Non-monotonic Reasoning*, pages 59–64, 2004.
- [Ben03] T. Bench-Capon. Persuasion in practical argument using value based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.
- [BGW05] H. Barringer, D. Gabbay, and J. Woods. Temporal dynamics of support and attack networks: From argumentation to zoology. In *Mechanizing Mathematical Reasoning*, pages 59–98. Springer, 2005.
- [BH08] Ph. Besnard and A. Hunter. *Elements of Argumentation*. MIT Press, 2008.
- [BH09] E. Black and A. Hunter. An inquiry dialogue system. *Autonomous Agents and Multi-Agent Systems*, 19(2):173–209, 2009.
- [BH12] E. Black and A. Hunter. A relevance-theoretic framework for constructing and deconstructing enthymemes. *Journal of Logic and Computation*, 22(1):55–78, 2012.
- [BDKM16] E. Bonzon, J. Delobelle, S. Konieczny, and N. Maudet. A Comparative Study of Ranking-based Semantics for Abstract Argumentation. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI'16)*, 2016.
- [BKRvdT13] R. Booth, S. Kaci, T. Rienstra, and L. van der Torre. A logical theory about dynamics in abstract argumentation. In *Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM 2013)*, 2013.
- [BDH14] Ph. Besnard, S. Doutre, and A. Herzig. Encoding argument graphs in logic. In *Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'14)*, pages 345–354. Springer, 2014.

- [Cam06] M. Caminada. On the issue of reinstatement in argumentation. In *Proceedings of 10th European Conference on Logics in Artificial Intelligence (JELIA'06)*, volume 4160 of *LNCS*, pages 111–123. Springer, 2006.
- [CDG<sup>+</sup>15] G. Charwat, W. Dvorák, S. Gaggl, J. Wallner, and S. Woltran. Methods for solving reasoning problems in abstract argumentation - A survey. *Artificial Intelligence*, 220:28–63, 2015.
- [CFL10] C. Cayrol, Dupin de Saint-Cyr F, and M. Lagasquie-Schiex. Change in abstract argumentation frameworks: Adding an argument. *Journal of Artificial Intelligence Research*, 38:49–84, 2010.
- [CG09] M Caminada and D Gabbay. A logical account of formal argumentation. *Studia Logica*, 93:109–145, 2009.
- [CP12] M. Caminada and M. Podlaskowski. Grounded semantics as persuasion dialogue. In *Computational Models of Argument (COMMA'12)*, pages 478–485, 2012.
- [CS10] M. Capobianco and G. R. Simari. An argument-based multi-agent system for information integration. In Peter McBurney, Iyad Rahwan, and Simon Parsons, editors, *Proceedings of the Seventh International Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2010)*, 2010.
- [CTO14] F. Cerutti, N. Tintarev, and N. Oren. Formal arguments, preferences, and natural language interfaces to humans: An empirical evaluation. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pages 207–212, 2014.
- [Dun95] P. 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.
- [Dyr13] S. Dyrkolbotn. The same, similar, or just completely different? equivalence for argumentation in light of logic. In *Proceedings of Wollic 2013*, volume 8071, pages 96–110. Springer, 2013.
- [EGW10] U. Egly, S. Gaggl, and S. Woltran. Answer-set programming encodings for argumentation frameworks. *Argument and Computation*, 1(2):147–177, 2010.
- [FFP13] B. Fazzinga, S. Flesca, and F. Parisi. Efficiently estimating the probability of extensions in abstract argumentation. In *Scalable Uncertainty Management*, volume 8078 of *LNCS*, pages 106–119. Springer, 2013.
- [FT11] X. Fan and F. Toni. Assumption-based argumentation dialogues. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'11)*, pages 198–203, 2011.

- [FT12] X. Fan and F. Toni. Mechanism design for argumentation-based persuasion. In *Computational Models of Argument (COMMA'12)*, pages 322–333, 2012.
- [Gro10] D. Grossi. On the logic of argumentation theory. In *Proceedings of 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, pages 409–416. IFAAMAS, 2010.
- [GM15] D. Grossi and S. Modgil. On the Graded Acceptability of Arguments. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, pages 868–874, 2015.
- [Han99] S. O. Hansson. A survey of non-prioritized belief revision. *Erkenntnis*, 50(2–3):413–427, 1999.
- [Han01] S. O. Hansson. *A Textbook of Belief Dynamics*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.
- [HSM<sup>+</sup>13] C. Hadjinikolis, Y. Siantos, S. Modgil, E. Black, and P. McBurney. Opponent modelling in persuasion dialogues. In *Proceedings of IJCAI*, pages 164–170, 2013.
- [HT14a] A. Hunter and M. Thimm. Probabilistic argument graphs for argumentation lotteries. In *Computational Models of Argument (COMMA'14)*, pages 1033–1034, 2014.
- [HT14b] A. Hunter and M. Thimm. Probabilistic argumentation with incomplete information. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pages 313 — 324, 2014.
- [Hun04a] A. Hunter. Making argumentation more believable. In *Proceedings of the 19th National Conference on Artificial Intelligence (AAAI'04)*, pages 269–274. MIT Press, 2004.
- [Hun04b] A. Hunter. Towards higher impact argumentation. In *Proceedings of the 19th National Conference on Artificial Intelligence (AAAI'04)*, pages 275–280. MIT Press, 2004.
- [Hun07] A. Hunter. Real arguments are approximate arguments. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence (AAAI'07)*, pages 66–71. MIT Press, 2007.
- [Hun12] A. Hunter. Some foundations for probabilistic argumentation. In *Proceedings of the International Conference on Computational Models of Argument (COMMA'12)*, pages 117–128, 2012.
- [Hun13a] A. Hunter. Analysis of dialogical argumentation via finite state machines. In *Proceedings of the International Conference on Scalable Uncertainty Methods*, volume 8078 of *LNCS*, pages 1–14. Springer, 2013.

- [Hun13b] A. Hunter. A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning*, 54(1):47–81, 2013.
- [Hun13c] A. Hunter. Probabilistic qualification of attack in abstract argumentation. *International Journal of Approximate Reasoning*, 55:607–638, 2013.
- [Hun14] A. Hunter. Probabilistic strategies in dialogical argumentation. In *Proceedings of the International Conference on Scalable Uncertainty Management (SUM'14)*, volume 8720 of *LNCS*, pages 190–202. Springer, 2014.
- [JCV08] J. Janssen, M. D. Cock, and D. Vermeir. Fuzzy argumentation frameworks. In *Proceedings of the 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'08)*, 513–520, 2008.
- [KJRM09] N. C. Karunatillake, N. R. Jennings, I. Rahwan, and P. McBurney. Dialogue games that agents play within a society. *Artificial Intelligence*, 173(9-10):935–981, April 2009.
- [LON11] H. Li, N. Oren, and T. Norman. Probabilistic argumentation frameworks. In *Proceedings of the First International Workshop on the Theory and Applications of Formal Argumentation (TFAFA'11)*, 2011.
- [MT08] P. Matt and F. Toni. A game-theoretic measure of argument strength for abstract argumentation. In *Proceedings on the 11th European Conference on Logics in Artificial Intelligence (JELIA'08)*, pages 285–297. Springer, 2008.
- [NAD14] S. Nofal, K. Atkinson, and P. Dunne. Algorithms for decision problems in argument systems under preferred semantics. *Artificial Intelligence*, 207:23–51, 2014.
- [OAL12] N. Oren, K. Atkinson, and H. Li. Group persuasion through uncertain audience modelling. In *Computational Models of Argument (COMMA'12)*, pages 350–357, 2012.
- [ON10] N. Oren and T. J. Norman. Arguing using opponent models. In *Proceedings of the Sixth International Workshop on Argumentation in Multi-Agent Systems (ArgMAS'09)*, pages 160–174. Springer, 2010.
- [Pea88] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Springer-Verlag, 1988.
- [Pet09] M. Peterson. *An Introduction to Decision Theory*. Cambridge University Press, 2009.

- [Pra05] H. Prakken. Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation*, 15(6):1009–1040, 2005.
- [Pra06] H. Prakken. Formal systems for persuasion dialogue. *Knowledge Engineering Review*, 21(2):163–188, 2006.
- [RK15] A. Rosenfeld and S. Kraus. Providing arguments in discussions based on the prediction of human argumentative behavior. In *Proceedings of the 29th AAAI Conference (AAAI’15)*, pages 1320—1327, 2015.
- [RL08] I. Rahwan and K. Larson. Pareto optimality in abstract argumentation. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI’08)*, pages 150—155, 2008.
- [RL09] I. Rahwan and K. Larson. Argumentation and game theory. In Iyad Rahwan and Guillermo R. Simari, editors, *Argumentation in Artificial Intelligence*, pages 321–339. Springer, 2009.
- [RPRS08] R. Riveret, H. Prakken, A. Rotolo, and G. Sartor. Heuristics in argumentation: A game theory investigation. In *Computational Models of Argument (COMMA 2008)*, pages 324–335. IOS Press, 2008.
- [RTO13] T. Rienstra, M. Thimm, and N. Oren. Opponent models with uncertainty for strategic argumentation. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI’13)*, pages 332–338. IJCAI/AAAI, 2013.
- [SW95] D. Sperber and D. Wilson. *Relevance: Communication and Cognition*. Blackwell, second edition, 1995.
- [TG10] M. Thimm and A. J. Garcia. Classification and strategical issues of argumentation games on structured argumentation frameworks. In *Proceedings of the Ninth International Joint Conference on Autonomous Agents and Multi-Agent Systems 2010 (AAMAS’10)*, May 2010.
- [Thi12] M. Thimm. A probabilistic semantics for abstract argumentation. In *Proceedings of the European Conference on Artificial Intelligence (ECAI’12)*, pages 750–755, 2012.
- [TG14] M. Thimm and G. Kern-Isberner. On Controversiality of Arguments and Stratified Labelings. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA’14)*, 2014.
- [Thi14a] M. Thimm. Strategic argumentation in multi-agent systems. *Kunstliche Intelligenz*, 2014.
- [Thi14b] M. Thimm. Tweety - A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR’14)*, 2014.

- [VBG<sup>+</sup>13] S. Villata, G. Boella, D. Gabbay, L. van der Torre, and J. Hulstijn. A logic of argumentation for specification and verification of abstract argumentation frameworks. *Annals of Mathematics and Artificial Intelligence*, 66:199–230, 2013.
- [WK95] D. Walton and E Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. SUNY Press, 1995.