Probabilistic Qualification of Attack in Abstract Argumentation

Anthony Hunter
Department of Computer Science
University College London
Gower Street
London, WC1E 6BT, UK
(a.hunter@cs.ucl.ac.uk)

July 30, 2013

Abstract

An argument graph is a graph where each node denotes an argument, and each arc denotes an attack by one argument on another. It offers a valuable starting point for theoretical analysis of argumentation following the proposals by Dung. However, the definition of an argument graph does not take into account the belief in the attacks. In particular, when constructing an argument graph from informal arguments, where each argument is described in free text, it is often evident that there is uncertainty about whether some of the attacks hold. This might be because there is some expressed doubt that an attack holds or because there is some imprecision in the language used in the arguments. In this paper, we use the set of spanning subgraphs of an argument graph as a sample space. A spanning subgraph contains all the arguments, and a subset of the attacks, of the argument graph. We assign a probability value to each spanning subgraph such that the sum of the assignments is 1. This means we can reflect the uncertainty over which is the actual subgraph using this probability distribution. Using the probability distribution over subgraphs, we can then determine the probability that a set of arguments is admissible or an extension. We can also obtain the probability of an attack relationship in the original argument graph as a marginal distribution (i.e. it is the sum of the probability assigned to each subgraph containing that attack relationship). We investigate some of the features of this proposal, and we consider the utility of our framework for capturing some practical argumentation scenarios.

Keywords: Computational models of argument; Argument systems; Abstract argumentation; Probabilistic argumentation; Uncertain argumentation; Enthymemes

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, RS09]).

In abstract argumentation, a graph is used to represent a set of arguments and counterarguments. Each node is an argument and each arc from α to β denotes an attack by α on β. 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, abstract argumentation does not explicitly consider whether an attack by an argument is believed or not. It only represents the existence of arguments and counterarguments. Yet often there is uncertainty with regard to the attacks.

In this paper, we will consider uncertainty of attacks. So given an argument graph $(\mathcal{A}, \mathcal{R})$, we may wish to assess the probability of each attack in $\mathcal{R}$ holding. Hence some attacks might be believed, some might be disbelieved, and some might be unknown. To illustrate, consider the following example.

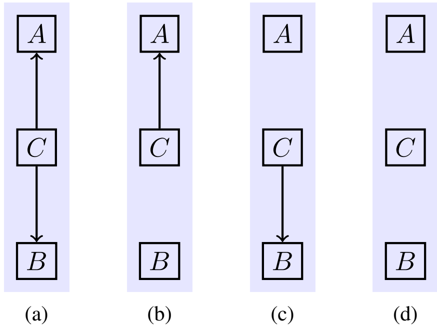


Figure 1: For argument graph G_1, the subgraphs are (a) G_1, (b) G_2, (c) G_3 and (d) G_4.

Example 1. *Consider the following arguments. The arguments A and B are arguments that each claim that the speaker is not involved in the robbery, and the argument C is by a potential witness casting doubt on the premises of arguments A and B by undercutting their premises. Also, we see that argument C attacks A with explicit certainty and C attacks B with explicit uncertainty.*

- *A = "John says he was not in town when the robbery took place, and therefore denies being involved in the robbery."*
- *B = "Peter says he was at home watching TV when the robbery took place, and therefore denies being involved in the robbery."*
- *C = "Harry says that he is certain that he saw John outside the bank just before the robbery took place, and he also thinks that possibly he saw Peter there too."*

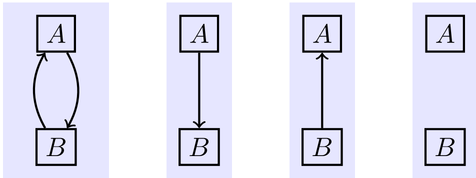
If we consider both attacks made by argument C, then we get the argument graph given in Figure 1a. However, if we also take into account the doubt in the attack by C on B, then we get the argument graph given in Figure 1b. This means that there is uncertainty over whether the actual argument graph should be Figure 1a or Figure 1b. We can deal with this uncertainty by regarding the set of spanning subgraphs of Figure 1a (i.e. the four subgraphs given in Figure 1) as a sample space, and assigning a probability to each of them such that the sum is 1. For instance, if Harry has only weak confidence in C attacking B, then the probabilities might be 0.2 for Figure 1a and 0.8 for Figure 1b (i.e. $P(G_1) = 0.2$ and $P(G_2) = 0.8$).

In the example given above, there is explicit uncertainty expressed qualitatively in the attacks made by the arguments. Other situations where uncertainty arises is when there is ambiguity, a form of imprecision in the language used in the arguments, as illustrated in the following example.

Example 2. *Suppose there are two witnesses to a criminal escaping in a car. Also, suppose the first witness says that the getaway car is red, and the other witness says that the getaway car is orange. If we take a strict interpretation of the colours, then we have two arguments A and B below, where each argument attacks the other.*

- *A = "the getaway car is red"*
- *B = "the getaway car is orange"*

For these arguments, it may be inappropriate to treat "red" and "orange" as contradictory. There is some ambiguity, and hence some imprecision, in the use of these terms. And so, it may be possible to regard these two terms as consistent together. So if we consider the argument graph, there is some uncertainty as to whether A attacks B and vice versa. This means we have the four spanning subgraphs in Figure 2. We could then for instance consider either Figure 2a or Figure 2d to be the actual argument graph, and so the sum of the probability assigned to these two graphs is 1. Furthermore, the more we consider them to be inconsistent together, the more we assign the probability to Figure 2a, and the more we consider them to be consistent together, the more we assign the probability to Figure 2d. For instance, if we believe with high probability that they are not inconsistent together, then we could let $P(G_1) = 0.1$ and $P(G_4) = 0.9$.



(a) (b) (c) (d)

Figure 2: For argument graph G_1, the subgraphs are (a) G_1, (b) G_2, (c) G_3 and (d) G_4.

Another important reason for uncertainty in attacks is that real-world arguments presented in natural language are normally enthymemes [Wal89]. An enthymeme is an argument with a support that is insufficient for the claim to be entailed and/or a claim that is incomplete. We consider this in the next example (though Examples 1 and 2 also contain enthymemes).

Example 3. *Consider the following arguments. Here, A is an argument that has incomplete premises for obtaining the claim (i.e. the premise "the sun is shining now" is insufficient for entailing the claim "we should organize a BBQ for this evening"). And B is an argument that has the premise "the weather report predicts rain this evening", but lacks a claim. Implicitly, the claim of B should negate the premises or the claim of argument A.*

- A = *"the sun is shining now, we should organize a BBQ for this evening"*
- B = *"the weather report predicts rain this evening"*

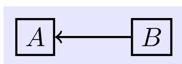
When a counterargument is an enthymeme, there may be uncertainty as to whether the argument being attacked is attacked because a premise is being contradicted or because the claim is being contradicted. In the above example, it could be that the implicit claim of B is negating the premise "the sun is shining now", or negating the claim "we should organize a BBQ for this evening". Also, because A is an enthymeme with incomplete premises, B could have a claim that contradicts a missing premise of A. For instance, suppose we make the premises of argument A explicit as follows:

- A_1 = *"the sun is shining now"*
- A_2 = *"if the sun is shining now, it will be warm and dry this evening"*
- A_3 = *"if it is warm and dry this evening, then we should organize a BBQ for this evening"*

So if the argument is explicit, it will have premises A_1, A_2 and A_3, and the claim "we should organize a BBQ for this evening". Then the claim of B can negate some combination of the explicit premises and/or the claim of A. Similarly, we can make the argument B explicit.

- B_1 = *"if the weather report predicts rain this evening, then it will not be warm and dry this evening"*
- B_2 = *"it will not be warm and dry this evening"*

Here B_1 is used to make the premises explicit, and B_2 is used to make the claim explicit. This claim then explicitly contradicts the premises A_1 and A_2. In other words, if we regard the premises of A as being represented by A_1, A_2 and A_3, and the premises of B being represented by B and B_1, and the claim of B being B_2, then we get the following argument graph.



However, there may be other interpretations of A and/or B, such that the interpretation of B does not attack the interpretation of A, then we get the following argument graph.

In this way, there is doubt about whether A does indeed attack B. We can treat the sample space as being composed of the above two argument graphs, and we define a probability distribution over this sample space (i.e. assign a probability to each of these graphs such that the sum is 1).

So to summarize, we see at least three kinds of uncertainty that arise in argumentation that we want to capture by quantifying the probability of attack.

Explicit uncertainty of attack Some arguments include explicit qualification of the attacks made on other arguments. This explicit qualification is normally qualitative (as in Example 1), but sometimes it can involve quantitative qualification (such as "I am 99% sure that what John said is a lie").

Implicit imprecision of argument Many arguments have a degree of imprecision in the terminology used (as in Example 2). Unless all the language is formally defined, and all participants use the same definitions, it is difficult to avoid some imprecision. This means that when considering two arguments it is not always certain whether or not one attacks the other. For instance, it is possible that "red" and "orange" are consistent together as the description of the same object.

Incomplete premises and/or claims Most arguments in natural language are enthymemes, which means that they do not explicitly present all their premises and/or claims (as in Example 3). With this incompleteness, it is difficult to be certain whether one argument attacks another. If a counterargument has an explicit claim, there may be uncertainty as to whether the attacked argument has the premise that the attacker has contradicted. And if a counterargument has an implicit claim, there may be further uncertainty as to what is being contradicted.

So to address these kinds of uncertainty in this paper, we investigate the use of a probability distribution over the spanning subgraphs of an argument graph. From this distribution, we can then determine the probability that a set of arguments is admissible, or an extension.

The idea to consider uncertain attacks was first proposed by Barringer *et al* [BGW05]. They suggested that an attack could be a conditional probability statement, but they did not investigate how this could be obtained, what it meant, or how it could be combined with Dung's dialectical semantics. The idea to assign a probability value to an attack and then use this in a development of dialectical semantics was first proposed by Li *et al* [LON11]. They used the probability assigned to each attack to obtain a probability distribution over the spanning subgraphs of the original argument graph, and they showed how extensions could be qualified by using this probability distribution. Unfortunately, the proposal by Li *et al* used an independence assumption between attacks. Whilst the independence assumption is useful and appropriate in some situations, it is not always appropriate for dealing with attacks.

So whilst the proposal by Li *et al* is very interesting, they left a number of interesting issues unexplored. We go beyond their proposal, by avoiding the need for an independence assumption, since we start with a probability distribution over the subgraphs. We also provide a number of further definitions and we provide numerous results. We see that the approach subsumes preference-based argumentation frameworks [AC02], weighted argument graphs [DHM$^+$09, DHM$^+$11], value-based argumentation frameworks [Ben03], resolution-based argumentation frameworks [BG08], and extended argumentation frameworks [Mod09]. We explore the question of independence of attacks, and see how the approach can be used when there are dependences between attacks. We also consider numerous examples to illustrate the utility of the approach, and in particular, we investigate the applicability of the approach to modelling enthymemes. This raises a number of interesting opportunities for generalizing abstract argumentation, and more importantly, opens the way for better modelling of important phenomena in argumentation such as the kinds of uncertainty identified above (i.e. explicit uncertainty of attack, implicit imprecision of argument, and incomplete premises and/or claims).

The rest of this paper is structured as follows: Section 2 is a review of Dung's definitions for abstract argumentation; Section 3 introduces the notion of a probabilistic attack graph; Section 4 investigates ways to analyse probabilistic attack graphs; Section 5 introduces a methodology for modelling enthymemes in probabilistic attack graphs; Section 6 relates the proposal in this paper to other developments of abstract argumentation; and Section 7 summarizes the contributions in this paper, and outlines future work.

2 Preliminaries

In this section, we review the proposal for abstract argumentation by Dung [Dun95]. Essentially, a collection of arguments can be formalized as a directed graph.

Definition 1. *An* **argument graph** *is a pair* $(\mathcal{A}, \mathcal{R})$ *where* $\mathcal{A}$ *is a finite set and* $\mathcal{R}$ *is a binary relation over* $\mathcal{A}$ *(i.e.,* $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$*).*

Each element $\alpha \in \mathcal{A}$ is called an **argument** and $(\alpha, \beta) \in \mathcal{R}$ means that α **attacks** β (accordingly, α is said to be an **attacker** of β). So α is a **counterargument** for β when $(\alpha, \beta) \in \mathcal{R}$ holds. For an argument α, we say that: α is **self-attacking** when α is an attacker of α; α is **unattacked** when there is no attacker of α; and α is **undefended** when there is an attacker β of α but there is no attacker γ of β.

Later we require the following subsidiary functions: For an argument graph $G = (\mathcal{A}, \mathcal{R})$, let $\mathsf{Nodes}(G) = \mathcal{A}$ and $\mathsf{Arcs}(G) = \mathcal{R}$.

Arguments can work together as a coalition by attacking other arguments and by defending their members from attack as follows.

Definition 2. *Let* $\Gamma \subseteq \mathcal{A}$ *be a set of arguments.* Γ **attacks** $\beta \in \mathcal{A}$ *iff there is an argument* $\alpha \in \Gamma$ *such that* α *attacks* β*.* Γ **defends** $\alpha \in \Gamma$ *iff for each argument* $\beta \in \mathcal{A}$*, if* β *attacks* α *then* Γ *attacks* β*.*

The following gives a requirement that should hold for a coalition of arguments to make sense. If it holds, it means that the arguments in the set offer a consistent view on the topic of the argument graph.

Definition 3. *A set* $\Gamma \subseteq \mathcal{A}$ *of arguments is* **conflictfree** *iff there are no* α *and* β *in* Γ *such that* α *attacks* β*.*

Now, we consider how we can find an acceptable set of arguments from an abstract argument graph. The simplest case of arguments that can be accepted is as follows.

Definition 4. *A set* $\Gamma \subseteq \mathcal{A}$ *of arguments is* **admissible** *iff* Γ *is conflictfree and defends all its elements.*

The intuition here is that for a set of arguments to be accepted, we require that, if any one of them is challenged by a counterargument, then they offer grounds to challenge, in turn, the counterargument. There always exists at least one admissible set: The empty set is always admissible.

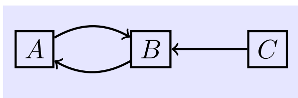
Clearly, the notion of admissible sets of arguments is the minimum requirement for a set of arguments to be accepted. In this paper, we will focus on the following classes of acceptable arguments proposed by Dung [Dun95].

Definition 5. *Let* Γ *be a conflictfree set of arguments, and let* $\mathsf{Defended} : \wp(\mathcal{A}) \to \wp(\mathcal{A})$ *be a function such that* $\mathsf{Defended}(\Gamma) = \{\alpha \mid \Gamma \textit{ defends } \alpha\}$*.*

1. Γ *is a* **complete extension** *iff* $\Gamma = \mathsf{Defended}(\Gamma)$
2. Γ *is a* **grounded extension** *iff it is the minimal (w.r.t. set inclusion) complete extension.*
3. Γ *is a* **preferred extension** *iff it is a maximal (w.r.t. set inclusion) complete extension.*
4. Γ *is a* **stable extension** *iff it is a preferred extension that attacks all arguments in* $\mathcal{A} \setminus \Gamma$*.*

In general, the grounded extension provides a skeptical view on which arguments can be accepted, whereas each preferred extension takes a credulous view on which arguments can be accepted. We illustrate these definitions with the following examples.

Example 4. *Consider the argument graph below. The conflictfree sets are* $\{\}$, $\{A\}$, $\{B\}$, $\{C\}$, *and* $\{A, C\}$*; The admissible sets are* $\{\}$, $\{A\}$, $\{C\}$, *and* $\{A, C\}$*; And the only complete set is* $\{A, C\}$*, and so this set is grounded, preferred and stable.*



Whilst the focus of this paper is on Dung's definitions for extensions, it would appear that the ideas would generalize to other definitions for extensions such as semi-stable semantics [Cam06] and ideal semantics [DMT07].

3 Representing uncertainty of attacks

In this section, we introduce probabilistic attack graphs as a way of representing the uncertainty of attacks, we then consider the nature of dependencies between attacks, including considering when independence can be assumed, and we conclude with a discussion of how we can use the independence assumption and how this relates to the proposal by Li *et al* [LON11].

3.1 Probabilistic attack graphs

In order to capture the uncertainty of attacks in an argument graph, we use the set of spanning subgraphs of an argument graph as a sample space. We define a spanning subgraph as follows.

Definition 6. *If $G = (\mathcal{A}, \mathcal{R})$ and $G' = (\mathcal{A}, \mathcal{R}')$ are argument graphs, then G' is a* **spanning subgraph** *of G, denoted $G' \sqsubseteq G$, iff $\mathcal{R}' \subseteq \mathcal{R}$. Also, let $G' \sqsubset G$ denote $G' \sqsubseteq G$ and $G \not\sqsubseteq G'$.*

So a spanning subgraph has the same set of arguments but a subset of attacks. Since spanning subgraphs are the only kind of subgraph we consider in this paper, we will simplify the terminology by referring to a spanning subgraph as a **subgraph**.

We assign a probability value to each subgraph such that the sum of the assignment is 1. This means we can reflect the uncertainty over which is the actual subgraph using this probability distribution.

Definition 7. *A* **probabilistic attack graph** *is a tuple $(\mathcal{A}, \mathcal{R}, P)$ where $G = (\mathcal{A}, \mathcal{R})$ is an argument graph, $\mathcal{G} = \{G' \mid G' \sqsubseteq G\}$ is the set of subgraphs of G, and $P : \mathcal{G} \to [0, 1]$, is such that*

$$\sum_{G' \in \mathcal{G}} P(G') = 1$$

We call P a **probability distribution over subgraphs** *in $\mathcal{G}$.*

Example 5. *Consider the sample space given in Figure 2. Suppose the probability distribution over subgraphs is $P(G_1) = 0.5$, $P(G_2) = 0.1$, $P(G_3) = 0.3$, and $P(G_4) = 0.1$.*

Given a probabilistic attack graph $(\mathcal{A}, \mathcal{R}, P)$, we can then obtain the probability of any arc in $(\mathcal{A}, \mathcal{R})$ being an attack by taking the marginal probability as defined next.

Definition 8. *For a probabilistic attack graph $(\mathcal{A}, \mathcal{R}, P)$, and an attack $(\alpha, \beta) \in \mathcal{R}$, the* **probability of attack***, denoted $P((\alpha, \beta))$, is*

$$P((\alpha, \beta)) = \sum_{G' \in \mathcal{G}(\alpha, \beta)} P(G')$$

where $\mathcal{G} = \{G' \mid G' \sqsubseteq G\}$ is the set of subgraphs of G, and $\mathcal{G}(\alpha, \beta) = \{G' \in \mathcal{G} \mid G' = (\mathcal{A}, \mathcal{R}')$ and $(\alpha, \beta) \in \mathcal{R}'\}$ is the set of subgraphs containing the arc.

So the probability of attack is the marginal probability obtained from the probability distribution over the subgraphs. Using this, we can qualify each attack in an argument graph by a probability value that indicates the belief that the attacks holds. So each attack $(\alpha, \beta) \in \mathcal{R}$ is assigned a value $P((\alpha, \beta))$ in the unit interval. Note, in order to simplify notation, when we will use $P(\alpha, \beta)$ to denote $P((\alpha, \beta))$ where there is no risk of confusion.

For a diagrammatic representation, we will use a **labelled graph** which is an argument graph where each arc is labelled with the probability of attack (obtained as the marginal probability from the probability distribution over subgraphs).

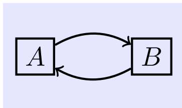
In general, we do not make any further constraints on the probability distributions over subgraphs beyond Definition 7. So for example, it is possible for every attack in a graph to have a probability of 1 (i.e. there original graph has probability of 1, and all strict subgraphs have probability 0). Similarly, it is possible for any or every attack in a graph to have a probability of 0 (i.e. the subgraph with no arcs has probability 1, and all the other subgraphs have probability 0). In the rest of this paper, we will explore some of the options we have for choosing and using probability distributions over subgraphs.

3.2 Dependencies between attacks

In this paper, we start with a probability distribution over subgraphs of an argument graph. Given a probabilistic attack graph $(\mathcal{A}, \mathcal{R}, P)$, we assume one subgraph $G' = (\mathcal{A}, \mathcal{R}')$ is the actual argument graph, where $G' \sqsubseteq G$ and $G = (\mathcal{A}, \mathcal{R})$. And since the sum of the probabilities of the subgraphs is 1, we can think of this in terms of a joint distribution by creating a conjunction of literals for each subgraph $G' \sqsubseteq G$, and for each arc ρ_i in the original graph G, the positive literal ρ_i in the conjunction denotes the arc is in the subgraph, and for each arc ρ_j in the original graph G, the negative literal $\overline{\rho_j}$ in the conjunction denotes the arc is not in the subgraph, as follows.

$$\left(\bigwedge_{\rho_i \in \mathcal{R}'} \rho_i \right) \wedge \left(\bigwedge_{\rho_j \in \mathcal{R} \setminus \mathcal{R}'} \overline{\rho_j} \right)$$

So for the joint distribution, we need a probability assignment for each of these conjuncts. Obviously, if we have n arcs in G, then there will be 2^n conjunctions to consider. For instance, consider the following graph, where ρ_1 denotes the attack by A on B and ρ_2 denotes the attack by B on A. For this, we can represent the joint distribution by $P(\rho_1 \wedge \rho_2)$, $P(\rho_1 \wedge \overline{\rho_2})$, $P(\overline{\rho_1} \wedge \rho_2)$, and $P(\overline{\rho_1} \wedge \overline{\rho_2})$.



By marginalization, we can obtain the probability of each attack. So for the above example, we have the probability of A attacking B given by $P(\rho_1)$ and the probability of B attacking A given by $P(\rho_2)$. These can be calculated as follows.

$$\begin{aligned} P(\rho_1) &= P(\rho_1 \wedge \rho_2) + P(\rho_1 \wedge \overline{\rho_2}) \\ P(\overline{\rho_1}) &= P(\overline{\rho_1} \wedge \rho_2) + P(\overline{\rho_1} \wedge \overline{\rho_2}) \\ P(\rho_2) &= P(\rho_1 \wedge \rho_2) + P(\overline{\rho_1} \wedge \rho_2) \\ P(\overline{\rho_2}) &= P(\rho_1 \wedge \overline{\rho_2}) + P(\overline{\rho_1} \wedge \overline{\rho_2}) \end{aligned}$$

For any probability distribution P, and variables X and Y, by definition $P(X \mid Y) = P(X \wedge Y)/P(Y)$. Therefore, $P(X \wedge Y) = P(X \mid Y)P(Y)$. Furthermore, X and Y are independent when $P(X \mid Y)P(Y) = P(X) \cdot P(Y)$. Hence, $P(X \wedge Y) = P(X) \cdot P(Y)$ when they are independent.

So returning to our running example, if ρ_1 and ρ_2 are independent, we obtain the joint distribution we started with, as follows.

$$\begin{aligned} P(\rho_1 \wedge \rho_2) &= P(\rho_1) \cdot P(\rho_2) \\ P(\rho_1 \wedge \overline{\rho_2}) &= P(\rho_1) \cdot P(\overline{\rho_2}) \\ P(\overline{\rho_1} \wedge \rho_2) &= P(\overline{\rho_1}) \cdot P(\rho_2) \\ P(\overline{\rho_1} \wedge \overline{\rho_2}) &= P(\overline{\rho_1}) \cdot P(\overline{\rho_2}) \end{aligned}$$

In probability theory, the assumption of independence between two variables X and Y means that knowing the outcome of X does not influence the probability of Y and vice versa. For example, if we toss an ordinary coin twice, and let X be the event that heads comes up for the first toss, $\overline{X}$ be the event that tails comes up for the first toss, Y be the event that heads comes up for the second toss, and $\overline{Y}$ be the event that tails comes up for the second toss. Then of course knowing the outcome of the first toss does not influence the probability of whether heads comes up for the second toss. In other words, X is independent of Y (and therefore of $\overline{Y}$), and vice versa, and $\overline{X}$ is independent of Y (and therefore of $\overline{Y}$). So if we start with a joint probability distribution, in our case a probability distribution over subgraphs, we have all the information we require for using probabilistic reasoning. Therefore, we can determine whether independence holds for a given probability distribution over subgraphs.

Returning to the examples in the introduction, in the first example (Example 1), it is reasonable to assume independence, and we can ensure that independence holds by selecting an appropriate probability distribution over the subgraphs. Take the attack by argument C on argument A. Knowing that this attack holds (i.e. that the argument A has been undermined) does not necessarily influence the probability that the attack by C on B holds. For this, we are taking the arguments at "face value". In other words, we are

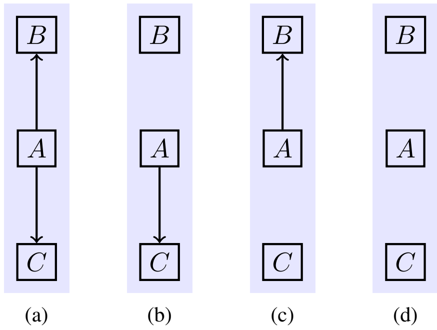


Figure 3: For argument graph G_1, the subgraphs are (a) G_1, (b) G_2, (c) G_3 and (d) G_4.

not adding any background information that may be useful. For instance, if we have some doubts about the reliability of Harry as a witness (perhaps someone has questioned his ability to recognize faces from a distance), then knowing that the attack by C on A is true may influence the probability of the attack by C on B. But since the example does not mention doubts about Harry, we can assume that Harry is reliable. So if he is correct in one of the attacks that neither increases or decreases the probability of the other attack. In other words, if he made a mistake in recognizing one of the characters, the mistake is not assumed to be a systematic problem but rather a chance occurrence.

Whilst assuming independence brings advantages, there are situations where it is not appropriate. To illustrate this, we consider the following example.

Example 6. *Consider argument graph G_1 in Figure 3a where the meaning for the arguments is as follows.*

- A = *CheapAir is going bust.*
- B = *The CheapAir tickets to Paris are a bargain; We should buy them for our holiday.*
- C = *The CheapAir tickets to New York are a bargain; We should buy them for our conference trip.*

Here the attack by A on B and the attack by A on C are not independent. For instance, if the attack by A on B is shown to be true, then there is a raised probability that the attack by A on C is true.

We can handle this example, and indeed any example where there are dependencies between attacks. This is because we start with a probability distribution over the subgraphs rather than a probability distribution on attacks. So we choose a probability distribution over subgraphs where there is dependence between the attacks. We illustrate this next.

Example 7. *Continuing Example 6, if we start with graph G_1 in Figure 3a, then there are four subgraphs to consider. Suppose we let the probability function over subgraphs be the following.*

$$P(G_1) = 0.7 \qquad P(G_2) = 0 \qquad P(G_3) = 0 \qquad P(G_4) = 0.3$$

The marginals for the attacks are as follows.

$$P(A, B) = P(G_1) + P(G_3) = 0.7$$
$$P(A, C) = P(G_1) + P(G_2) = 0.7$$

The main disadvantage of starting with a probability distribution over the subgraphs (rather than starting with a probability distribution over attacks and then using the independence assumption to get the probability distribution over subgraphs) is that we need to explicitly determine the probability value for each subgraph and, of course, if the probabilistic attack graph has n attacks, we need to consider 2^n subgraphs in the worst case. Often however, it will only be that some attacks are dependent, and the problem can be simplified in terms of simpler conditional dependencies. For instance, suppose we have three attacks ρ_1,

	Subgraph	Probability of subgraph	Grounded extension	Preferred extensions
G_1	$A \leftrightarrow B$	$P(A,B) \cdot P(B,A)$	$\{\}$	$\{A\}, \{B\}$
G_2	$A \rightarrow B$	$P(A,B) \cdot (1 - P(B,A))$	$\{A\}$	$\{A\}$
G_3	$A \leftarrow B$	$(1 - P(A,B)) \cdot P(B,A)$	$\{B\}$	$\{B\}$
G_4	$A \quad B$	$(1 - P(A,B)) \cdot (1 - P(B,A))$	$\{A,B\}$	$\{A,B\}$

Table 1: Graph G_1 and its subgraphs together with the probability of each subgraph and extensions. In this example, it is assumed that the probability distribution over subgraphs is such that the attacks are independent, and so the probability of each subgraph can be calculated from the probability of each attack.

ρ_2, and ρ_3 to consider in a graph G, and ρ_1 and ρ_2 are independent, but ρ_3 is dependent on ρ_1 and ρ_2, then we could calculate the conjunction as follows.

$$\begin{aligned} P(\rho_1 \wedge \rho_2 \wedge \rho_3) &= P(\rho_3 \mid \rho_1 \wedge \rho_2) \cdot P(\rho_1 \wedge \rho_2) \\ &= P(\rho_3 \mid \rho_1 \wedge \rho_2) \cdot P(\rho_1) \cdot P(\rho_2) \end{aligned}$$

So by determining which are the dependent attacks and which are the independent attacks, the problem of finding the joint distribution can be broken down into smaller joint distributions which can then be combined using independence assumptions.

3.3 Using the independence assumption

As we discussed in the previous section, we may assume independence for some examples of probabilistic attack graphs. If so, it can be more efficient to use the probability assignment to attacks (i.e. the marginals) to generate the probability distribution over subgraphs. For an attack (α, β) in a graph G, with a probability assignment P, there is the probability $P(\alpha, \beta)$ that (α, β) holds in the graph, and there is the probability $1 - P(\alpha, \beta)$ that (α, β) does not hold in the graph.

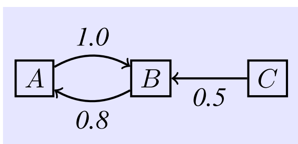
Let $G = (\mathcal{A}, \mathcal{R}, P)$ and $G' = (\mathcal{A}, \mathcal{R}', P)$ be probabilistic attack graphs such that $R' \subseteq R$ (i.e. $(\mathcal{A}, \mathcal{R}')$ is a subgraph of $(\mathcal{A}, \mathcal{R})$). If the independence assumption holds for P (i.e. the attacks are independent according to the probability distribution over the subgraphs), then the probability of subgraph G', i.e. $P(G')$, is obtained as follows:

- If $\mathcal{R} = \emptyset$, then $P(G) = 1$, and for all $G' \sqsubset G$, $P(G') = 0$.
- If $\mathcal{R} \neq \emptyset$, then for all $G' \sqsubseteq G$, $P(G') = (\prod_{(\alpha,\beta)\in\mathcal{R}'} P(\alpha, \beta)) \times (\prod_{(\alpha,\beta)\in\mathcal{R}\setminus\mathcal{R}'} (1 - P(\alpha, \beta)))$

To use this, we construct all subgraphs of G, and calculate the probability of each subgraph based on the probabilities of the attacks holding in the graph. So the probability of a subgraph captures the degree of certainty that the subgraph contains exactly the attacks that hold. Note, a similar definition to the above definition has been presented in [LON11].

In Table 1 and Table 2, the structure for an argument graph is provided. In each case it is given as G_1. For each of these graphs, the subgraphs are listed. For each subgraph, the probability of it is calculated (in terms of the probability of the attacks), and the grounded and preferred extensions identified. A specific probabilistic attack graph is presented in the next example.

Example 8. *Consider the following argument graph where each arc is labelled with the probability of attack. The probability distribution of the eight subgraphs of the graph as listed in Table 2 is* $P(G_1) = 0.4$, $P(G_2) = 0.4$, $P(G_3) = 0.1$, $P(G_4) = 0$, $P(G_5) = 0.1$, $P(G_6) = 0$, $P(G_7) = 0$, *and* $P(G_8) = 0$.



	Subgraph	Probability of subgraph	Grounded extension	Preferred extensions
G_1	$A \leftrightarrow B \leftarrow C$	$P(A,B) \cdot P(B,A) \cdot P(C,B)$	$\{A,C\}$	$\{A,C\}$
G_2	$A \leftrightarrow B \quad C$	$P(A,B) \cdot P(B,A) \cdot (1 - P(C,B))$	$\{C\}$	$\{A,C\}, \{B,C\}$
G_3	$A \rightarrow B \leftarrow C$	$P(A,B) \cdot (1 - P(B,A)) \cdot P(C,B)$	$\{A,C\}$	$\{A,C\}$
G_4	$A \leftarrow B \leftarrow C$	$(1 - P(A,B)) \cdot P(B,A) \cdot P(C,B)$	$\{A,C\}$	$\{A,C\}$
G_5	$A \rightarrow B \quad C$	$P(A,B) \cdot (1 - P(B,A)) \cdot (1 - P(C,B))$	$\{A,C\}$	$\{A,C\}$
G_6	$A \leftarrow B \quad C$	$(1 - P(A,B)) \cdot P(B,A) \cdot (1 - P(C,B))$	$\{B,C\}$	$\{B,C\}$
G_7	$A \quad B \leftarrow C$	$(1 - P(A,B)) \cdot (1 - P(B,A)) \cdot P(C,B)$	$\{A,C\}$	$\{A,C\}$
G_8	$A \quad B \quad C$	$(1 - P(A,B)) \cdot (1 - P(B,A)) \cdot (1 - P(C,B))$	$\{A,B,C\}$	$\{A,B,C\}$

Table 2: Graph G_1 and its subgraphs together with the probability of each subgraph and extensions. In this example, it is assumed that probability distribution over subgraphs is such that the attacks are independent, and so the probability of each subgraph can be calculated from the probability of each attack.

If all the attacks in a probabilistic attack graph G have probability of 1, then the only subgraph of G to have non-zero probability is G, and so it has probability 1. Hence, we will see that we recover Dung's original definitions and results by assuming all attacks have probability 1. At the other extreme, if all the attacks in a probabilistic attack graph G have probability of 0, there is one graph with probability 1 and it contains all the arguments but no attacks. Note, we do not need to consider the independence assumption for either of these observations.

If all the attacks in a probabilistic attack graph G have probability of 0.5, and the independence assumption holds, then each subgraph of G has the same probability (as shown by the following result), and so there is a uniform distribution over the sample space.

Proposition 1. *If $G = (\mathcal{A}, \mathcal{R}, P)$ is a probabilistic attack graph such that P for all $(\alpha, \beta) \in \mathcal{R}$, $P(\alpha, \beta) = 0.5$, and the independence assumption holds for P, then for each $G' \sqsubseteq G$, $P(G') = (0.5)^n$, where $n = |\mathcal{R}|$.*

Proof. By definition, for each $G' \sqsubseteq G$, $p(G) = \prod_{(\alpha,\beta)\in\mathcal{R}'} P(\alpha, \beta) \times \prod_{(\alpha,\beta)\in\mathcal{R}\setminus\mathcal{R}'}(1 - P(\alpha, \beta))$. Since for all $(\alpha, \beta) \in \mathcal{R}$, $P(\alpha, \beta) = 0.5$, the above is equal to $\prod_{(\alpha,\beta)\in\mathcal{R}} P(\alpha, \beta)$, which is equal to $(0.5)^n$, where $n = |\mathcal{R}|$. □

We conclude this subsection with a couple of subsidiary definitions that we will use later. Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, we call the marginal distribution P: **maximal** iff for all $\alpha \in \mathcal{R}$, $P(\alpha) = 1$; and **minimal** iff for all $\alpha \in \mathcal{R}$, $P(\alpha) = 0$;

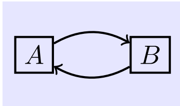
4 Analysing probabilistic attack graphs

In this section, we show how we can use probabilistic attack graphs to give a probabilistic qualification to admissible sets, extensions, and inferences (i.e. arguments that appear in one or more extensions).

In order to harness the notion of a probabilistic attack graph, we require some further subsidiary definitions and notation that we will use in the rest of this section.

For an argument graph $G = (\mathcal{A}, \mathcal{R})$, and a set of arguments $\Gamma \subseteq \mathcal{A}$, $G \Vdash_{\mathsf{ad}} \Gamma$ denotes that Γ is an admissible set in $(\mathcal{A}, \mathcal{R})$ and $G \Vdash_X \Gamma$ denotes that Γ is an X extension of $(\mathcal{A}, \mathcal{R})$ for $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, where co denotes complete semantics, pr denotes preferred semantics, st denotes stable semantics, and gr denotes grounded semantics. When $G \Vdash_{\mathsf{ad}} \Gamma$ holds, we say that G **entails** Γ. The set of subgraphs that entails a set of arguments Γ, denoted $Q_{\mathsf{ad}}(\Gamma)$, is $\{G' \sqsubseteq G \mid G' \Vdash_{\mathsf{ad}} \Gamma\}$. Similarly, $Q_X(\Gamma) = \{G' \sqsubseteq G \mid G' \Vdash_X \Gamma\}$, for $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$. For an argument $\alpha \in \mathcal{A}$, the set of subgraphs that imply an argument is admissible, denoted $I(\alpha)$, is $I_{\mathsf{ad}}(\alpha) = \{G' \sqsubseteq G \mid G' \Vdash_{\mathsf{ad}} \Gamma \text{ and } \alpha \in \Gamma\}$. Similarly, $I_X(\alpha) = \{G' \sqsubseteq G \mid G' \Vdash_X \Gamma \text{ and } \alpha \in \Gamma\}$.

Example 9. *Consider the following argument graph with the probability space given in Figure 2.*



For this argument graph G_1, we get the following for the $\Vdash_{\mathsf{ad}}$, $\Vdash_{\mathsf{pr}}$ and $\Vdash_{\mathsf{gr}}$ relations.

$$\begin{array}{lll} G_1 \Vdash_{\mathsf{ad}} \emptyset & G_1 \nVdash_{\mathsf{pr}} \emptyset & G_1 \Vdash_{\mathsf{gr}} \emptyset \\ G_1 \Vdash_{\mathsf{ad}} \{A\} & G_1 \Vdash_{\mathsf{pr}} \{A\} & G_1 \nVdash_{\mathsf{gr}} \{A\} \\ G_1 \Vdash_{\mathsf{ad}} \{B\} & G_1 \Vdash_{\mathsf{pr}} \{B\} & G_1 \nVdash_{\mathsf{gr}} \{B\} \\ G_1 \nVdash_{\mathsf{ad}} \{A,B\} & G_1 \nVdash_{\mathsf{pr}} \{A,B\} & G_1 \nVdash_{\mathsf{gr}} \{A,B\} \end{array}$$

For this argument graph G_1, we get the following for the Q_{ad}, Q_{pr} and Q_{gr} functions.

$$\begin{array}{lll} Q_{\mathsf{ad}}(\emptyset) = \{G_1, G_2, G_3, G_4\} & Q_{\mathsf{pr}}(\emptyset) = \{\} & Q_{\mathsf{gr}}(\emptyset) = \{G_1\} \\ Q_{\mathsf{ad}}(\{A\}) = \{G_1, G_2, G_4\} & Q_{\mathsf{pr}}(\{A\}) = \{G_1, G_2\} & Q_{\mathsf{gr}}(\{A\}) = \{G_2\} \\ Q_{\mathsf{ad}}(\{B\}) = \{G_1, G_3, G_4\} & Q_{\mathsf{pr}}(\{B\}) = \{G_1, G_3\} & Q_{\mathsf{gr}}(\{B\}) = \{G_3\} \\ Q_{\mathsf{ad}}(\{A,B\}) = \{G_4\} & Q_{\mathsf{pr}}(\{A,B\}) = \{G_4\} & Q_{\mathsf{gr}}(\{A,B\}) = \{G_4\} \end{array}$$

For this argument graph G_1, we get the following for the I_{ad}, I_{pr} and I_{gr} functions.

$$\begin{array}{lll} I_{\mathsf{ad}}(A) = \{G_1, G_2, G_4\} & I_{\mathsf{pr}}(A) = \{G_1, G_2, G_4\} & I_{\mathsf{gr}}(A) = \{G_2, G_4\} \\ I_{\mathsf{ad}}(B) = \{G_1, G_3, G_4\} & I_{\mathsf{pr}}(B) = \{G_1, G_3, G_4\} & I_{\mathsf{gr}}(B) = \{G_3, G_4\} \end{array}$$

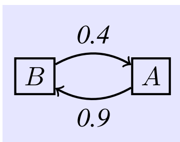
4.1 Probability functions over admissible sets

The probabilistic information in probabilistic attack graphs can be harnessed to qualify the uncertainty that a given set of arguments is an admissible set. Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, and a set of arguments $\Gamma \subseteq \mathcal{A}$, we show in this section how to calculate the probability that Γ is admissible, which we denote by $P_{\mathsf{ad}}(\Gamma)$. For this, we define $P_{\mathsf{ad}}(\Gamma)$ as the sum of the probability of the subgraphs for which Γ is an admissible set.

Definition 9. *Let $G = (\mathcal{A}, \mathcal{R}, P)$ be a probabilistic attack graph and let $\Gamma \subseteq \mathcal{A}$. The probability that Γ is in an admissible set is*

$$P_{\mathsf{ad}}(\Gamma) = \sum_{G' \in Q_{\mathsf{ad}}(\Gamma)} P(G')$$

Example 10. *Consider the following labelled argument graph. With the independence assumption holding, the probability distribution over the four subgraphs of the graph (as listed in Table 1) is $P(G_1) = 0.36$, $P(G_2) = 0.54$, $P(G_3) = 0.04$, and $P(G_4) = 0.06$. Therefore, $P_{\mathsf{ad}}(\emptyset) = 0.36 + 0.54 + 0.04 + 0.06 = 1$, $P_{\mathsf{ad}}(\{A\}) = 0.36+0.54+0.06 = 0.96$, $P_{\mathsf{ad}}(\{B\}) = 0.36+0.04+0.06 = 0.46$, and $P_{\mathsf{ad}}(\{A,B\}) = 0.06$.*



So in this example, we see that the empty set is an admissible set with probability 1, the set $\{A\}$ is an admissible set with the high probability of 0.96, the set $\{B\}$ is an admissible set with the substantially lower probability of 0.46, and the set $\{A,B\}$ is an admissible set with the very low probability of 0.06. Therefore the high probability of attack by A on B, and the lower probability by B on A, means that the attack by A on B dominates in determining the admissible sets.

The next example is a light sensor fusion problem where we consider three sources.

Example 11. *Consider a network of three light sensors that all sense the same light source with the aim of describing the colour seen. Suppose sensor s1 gives argument A = "colour is grey", sensor s2 gives*

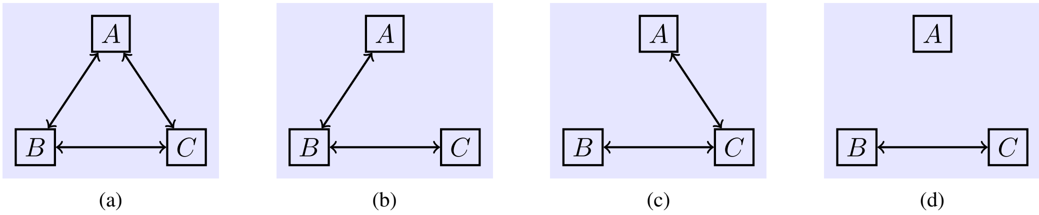


Figure 4: For argument graph G_1 in Example 11, four of the subgraphs of G_1 are (a) G_1, (b) G_2, (c) G_3 and (d) G_4.

argument B = "colour is white", and sensor s3 gives argument C = "colour is black". Suppose this is represented by the argument graph G_1 in Figure 4a. Suppose there is some uncertainty about whether "grey" and "white" contradict, and similarly there is some uncertainty about whether "grey" and "black" contradict. We can represent this using a probabilistic attack. So there are 64 subgraphs to consider in the sample space. Let us suppose that we only assign probability to four of the subgraphs given in Figure 4. Now assume a probability distribution over the subgraphs so that $P(G_1) = 0.2$, $P(G_2) = 0.2$, $P(G_3) = 0.3$ and $P(G_4) = 0.3$. So the attacks between B and C are certain, whereas the other attacks are uncertain. Given the probability distribution over subgraphs, we can obtain, for each set of arguments, the probability of being an admissible set as follows. Each of the singleton sets is an admissible set in each of G_1 to G_4. Hence, the probability for each of these singleton sets being admissible is 1. The probability of $\{A, B\}$ being admissible is 0.6, and so there is reasonably high probability that A and B are consistent together. Similarly, the probability of $\{A, C\}$ being admissible is 0.5, and so there is reasonably high probability that A and C are consistent together. The probability of $\{B, C\}$ being admissible is 0, and so there is zero probability that B and C are consistent together.

Admissible set	$\{A\}$	$\{B\}$	$\{C\}$	$\{A, B\}$	$\{A, C\}$	$\{B, C\}$	$\{A, B, C\}$
Probability	*1*	*1*	*1*	*0.6*	*0.5*	*0*	*0*

In an argument graph, the empty set is always an admissible set. For probabilistic attack graphs, we have the following result that shows that the empty set is admissible set with probability 1.

Proposition 2. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, the probability that the empty set is admissible is 1 (i.e. $P_{\mathsf{ad}}(\emptyset) = 1$).*

Proof. For each $G' \sqsubseteq G$, the definition of admissible set implies that $G' \Vdash_{\mathsf{ad}} \emptyset$ holds. Therefore, we have $P_{\mathsf{ad}}(\emptyset) = 1$. □

For an argument graph $(\mathcal{A}, \mathcal{R})$, if $\mathcal{R} = \emptyset$, then $\mathcal{A}$ is conflictfree and hence admissible. So for a probabilistic attack graph, if $\mathcal{A}$ is conflictfree, then $\mathcal{A}$ is admissible with probability 1.

Proposition 3. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, if $\mathcal{A}$ is conflictfree, and $\Gamma \subseteq \mathcal{A}$, then $P_{\mathsf{ad}}(\Gamma) = 1$.*

Proof. Assume $\mathcal{A}$ is conflictfree. Therefore $\mathcal{R} = \emptyset$, and so by definition $P(G) = 1$. Also any $\Gamma \subseteq \mathcal{A}$ is conflictfree, and therefore $G \Vdash_{\mathsf{ad}} \Gamma$. Hence, $P_{\mathsf{ad}}(\Gamma) = 1$. □

When P is maximal, each attack has a probability of 1. Therefore, there is only one subgraph with non-zero probability, and this is the graph itself. So when P is maximal, a probabilistic attack graph is reduced to an argument graph. Hence, any set of arguments is either admissible in the argument graph (and therefore with probability 1 in the probabilistic attack graph) or not admissible in the argument graph (and therefore with probability 0 in the probabilistic attack graph).

Proposition 4. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, if P is maximal (i.e. $P(\alpha) = 1$ for all $\alpha \in \mathcal{R}$), then for all $\Gamma \subseteq \mathcal{A}$, $P_{\mathsf{ad}}(\Gamma) = 1$ or $P_{\mathsf{ad}}(\Gamma) = 0$.*

Proof. Assume the probability of each attack in $\mathcal{R}$ is 1. Therefore, $P_{\mathsf{ad}}(G) = 1$, and for all $G' \sqsubset G$, $P(G') = 0$. Therefore, $P_{\mathsf{ad}}(\Gamma) = 1$ iff $G \Vdash_{\mathsf{ad}} \Gamma$, and $p(\Gamma) = 0$ iff $G \nVdash_{\mathsf{ad}} \Gamma$. Therefore, for all $\Gamma \subseteq \mathcal{A}$, $P_{\mathsf{ad}}(\Gamma) = 1$ or $P_{\mathsf{ad}}(\Gamma) = 0$. □

Now it is straightforward to show how probabilistic attack graphs subsume Dung's original notion of admissibility. Essentially, if we choose to assign a probability of 1 to all attacks, then we get exactly Dung's admissible sets by selecting the sets with probability of 1.

Proposition 5. *Let $G = (\mathcal{A}, \mathcal{R}, P)$ be a probabilistic attack graph such that P is maximal (i.e. $P(\alpha) = 1$ for all $\alpha \in \mathcal{R}$). For $\Gamma \subseteq \mathcal{A}$, Γ is admissible with respect to $(\mathcal{A}, \mathcal{R})$ iff $P_{\mathsf{ad}}(\Gamma) = 1$.*

Proof. From the assumption that P is maximal, we have $P(G) = 1$. ($\Rightarrow$) Assume Γ is admissible with respect to $(\mathcal{A}, \mathcal{R})$. Therefore, by definition, $G \in Q_{\mathsf{ad}}(\Gamma)$. Hence, $\sum_{G' \in Q_{\mathsf{ad}}(\Gamma)} P(G') = 1$, since $P(G) = 1$. Therefore, by definition, $P_{\mathsf{ad}}(\Gamma) = 1$. ($\Leftarrow$) Assume $P_{\mathsf{ad}}(\Gamma) = 1$. Therefore, by definition, $\sum_{G' \in Q_{\mathsf{ad}}(\Gamma)} P(G') = 1$. Hence, $G \in Q_{\mathsf{ad}}(\Gamma)$, since $P(G) = 1$. Therefore, by definition, Γ is admissible with respect to $(\mathcal{A}, \mathcal{R})$. □

When all attacks are uncertain (i.e. for each attack (α, β), $P(\alpha, \beta) < 1$), and the independence assumption holds, then each subgraph has non-zero probability. So the subgraph with no attacks has non-zero probability. Since the set of all arguments is an admissible set for a graph with no attacks, the set of all arguments is an admissible set with non-zero probability.

Proposition 6. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, where the independence assumption holds for P, if for all $(\alpha, \beta) \in \mathcal{R}$, $P(\alpha, \beta) < 1$, then for each $\Gamma \subseteq \mathcal{A}$, $P_{\mathsf{ad}}(\Gamma) > 0$.*

Proof. Assume that for all $(\alpha, \beta) \in \mathcal{R}$, $P(\alpha, \beta) < 1$. Consider $G' = (\mathcal{A}, \mathcal{R}', P)$ where $\mathcal{R}' = \emptyset$. Clearly $G' \sqsubseteq G$. Since $\mathcal{R}' = \emptyset$, we have $\mathcal{R} \setminus \mathcal{R}' = \mathcal{R}$. Therefore, the independence assumption implies $P(G') = \prod_{(\alpha,\beta)\in\mathcal{R}}(1 - P(\alpha, \beta))$. Therefore, we have $P(G') > 0$ (since we have assumed that for all $(\alpha, \beta) \in \mathcal{R}$, $P(\alpha, \beta) < 1$). Also, because $\mathcal{R}' = \emptyset$, we have $G' \Vdash_{\mathsf{ad}} \Gamma$ for each $\Gamma \subseteq \mathcal{A}$. Hence, $P_{\mathsf{ad}}(\Gamma) > 0$. □

When P is minimal, each attack has a probability of 0. Therefore, there is only one subgraph with non-zero probability, and this is the graph with no attacks. So when P is minimal, a probabilistic attack graph is reduced to a conflictfree set of arguments. Hence, any set of arguments is admissible with probability 1.

Proposition 7. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, if P is minimal (i.e. $P(\alpha) = 0$ for all $\alpha \in \mathcal{R}$), then for all $\Gamma \subseteq \mathcal{A}$, $P_{\mathsf{ad}}(\Gamma) = 1$.*

Proof. Assume that for all $(\alpha, \beta) \in \mathcal{R}$, $P(\alpha, \beta) = 0$. Then there is a subgraph $G' \sqsubseteq G$ where $G' = (\mathcal{A}, \mathcal{R}', P)$ and $\mathcal{R}' = \emptyset$. Since, $\mathcal{R}' = \emptyset$ and for all $(\alpha, \beta) \in \mathcal{R}$, $P(\alpha, \beta) = 0$, there is one subgraph in the sample space which has probability 1, and this is the subgraph with no attacks. □

In this section, we have seen how probabilistic attack graphs can be used to generate a probability function over admissible sets, and this will be extended to extensions in the next section.

4.2 Probability functions over extensions

The probabilistic information in a probabilistic attack graph can be harnessed to qualify the uncertainty that a given set of arguments is an extension. Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, and a set of arguments $\Gamma \subseteq \mathcal{A}$, the probability that Γ is an X extension, where $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, is defined as follows. Note, a similar definition to this definition has been given in [LON11].

Definition 10. *Let $G = (\mathcal{A}, \mathcal{R}, P)$ be a probabilistic attack graph, let $\Gamma \subseteq \mathcal{A}$ and let $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$. The probability that Γ is an X extension, denoted $P_X(\Gamma)$, is*

$$\sum_{G' \in Q_X(\Gamma)} P(G')$$

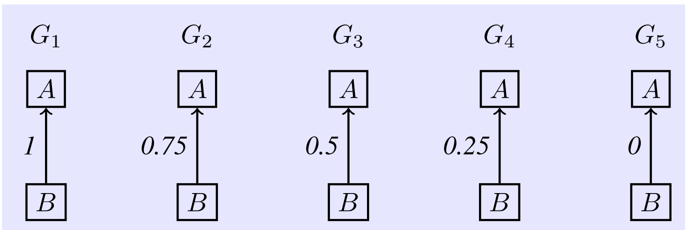
	Subgraph	Probability of subgraph	Grounded extension
G_1	$A \leftarrow B \leftarrow C$	0.5	$\{A, C\}$
G_2	$A \leftarrow B \quad C$	0.2	$\{B, C\}$
G_3	$A \quad B \leftarrow C$	0.2	$\{A, C\}$
G_4	$A \quad B \quad C$	0.1	$\{A, B, C\}$

Table 3: The graph G_1 and its subgraphs (with probability of subgraphs and extensions) for Example 12 where $P(B, A) = 0.8$ and $P(C, B) = 0.4$.

Example 12. *Consider the argument graph G_1 and its subgraphs, together with the probability distribution and extensions, given in Table 3. For $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $P_X(\{A, C\}) = 0.5 + 0.2 = 0.7$, $P_X(\{B, C\}) = 0.2$, and $P_X(\{A, B, C\}) = 0.1$.*

To illustrate the effect on the grounded extensions by decreasing the probability of attack, we consider the following simple example.

Example 13. *The following labelled graphs have the same structure but different probabilities of attack. For each, the probabilistic attack graph is $(\mathcal{A}, \mathcal{R}, P^i)$, and the probability distribution P^i can be obtained straightforwardly from the labelled graph G_i below since the sample space contains just two graphs $G = (\{A, B\}, \{(B, A)\})$ and $G' = (\{A, B\}, \{\})$. Essentially, for each graph G_i, $P^i(G) = 1 - P^i(G')$. For instance, for the first labelled graph (i.e. G_1), $P^1(G) = 1$ and $P^1(G') = 0$, and for the second labelled graph, $P_2(G) = 0.75$ and $P_2(G') = 0.25$*



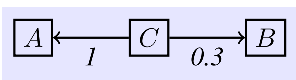
For each of these graphs, the probability for each of $\{B\}$ and $\{A, B\}$ being the grounded extension is tabulated below for each probabilistic attack graph.

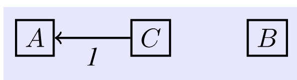
	G_1	G_2	G_3	G_4	G_5
$P^i_{\mathsf{gr}}(\{B\})$	1	0.75	0.5	0.25	0
$P^i_{\mathsf{gr}}(\{A, B\})$	0	0.25	0.5	0.75	1

So here we see that as the probability of attack drops, the probability that $\{B\}$ being the grounded extension drops and the the probability that $\{A, B\}$ being the grounded extension rises.

To illustrate how we can deal with situations involving explicit uncertainty about attacks, we return to an example we discussed in Section 1 concerning a bank robbery.

Example 14. *Returning to Example 1, let A be the argument "John says he was not in town when the robbery took place, and therefore denies being involved", let B be the argument "Peter says he was at home watching TV when the robbery took place, and therefore denies being involved", and let C be the argument "Harry says that he is certain that he saw John outside the bank just before the robbery took place, and he also thinks that possibly he saw Peter there too". We represent this by the labelled graph G below left. Let G' be the labelled graph below right. Suppose we only assign probability mass to G and to G' so that $P(G) = 0.3$ and $P(G') = 0.7$.*





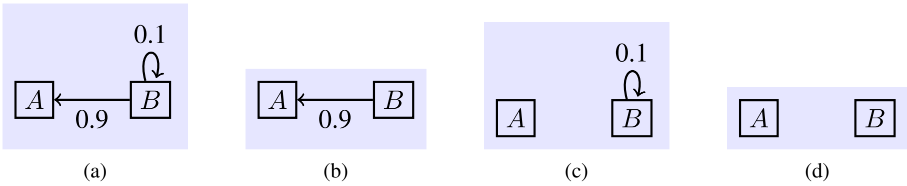


Figure 5: The subgraphs for Example 15. Each subgraph is labelled with the probability of attack.

Hence, for $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $P_X(\{C\}) = 0.3$, and $P_X(\{B, C\}) = 0.7$. Hence, with high probability, $\{B, C\}$ is the grounded extension.

An interesting, and not uncommon situation, where uncertainty about attacks occurs is with cycles, in particular with self-attacking arguments. Real world examples of self-attacking arguments can be usefully analyzed by identifying the uncertainty of the self attack. We see this next.

Example 15. *Consider the following statements taken by the police concerning a suspect called James. First, "James says he was at home all day", and second "Paul said he could be mistaken but he believes he saw James in the office this morning". Let us represent the first sentence by argument A and the second sentence by argument B where B attacks A and B attacks itself. Hence, this could be represented by the argument graph G_1 given in Figure 5a. For this example, we use the independence assumption.*

Given the uncertainty associated with these attacks, it would appear that at least the certainty of the attack by B on B should be low. The strict subgraphs are G_2 (Figure 5b), G_3 (Figure 5c), and G_4 (Figure 5d). The probabilities for each of these graphs is $P(G_1) = 0.09$, $P(G_2) = 0.81$, $P(G_3) = 0.01$, and $P(G_4) = 0.09$. Therefore, for $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}\}$, $P_X(\{B\}) = 0.81$, $P_X(\{A, B\}) = 0.09$, $P_X(\{\}) = 0.09$, and $P_X(\{A\}) = 0.01$. Hence, with high probability, $\{B\}$ is the grounded extension.

To illustrate how we can use these ideas for enthymemes, consider the following example concerning a forthcoming party.

Example 16. *Let Ann and Bob be two invitees to a party. Suppose there is a reason to believe that Ann will come, and there is a reason to believe that Bob will come, We can represent this situation by two arguments as follows.*

- A = *"Ann will come to the party"*
- B = *"Bob will come to the party"*

Now suppose that it is known that Ann and Bob dislike each other, and so if one comes, the other probably will not come. We can represent these incompatibilities by the attack relation so that A and B attack each other. For this, each of these arguments is an enthymeme since the claim does not explicitly reflect the inconsistencies for identifying these attacks. But, via common knowledge, the attack can be identified. This means we have an argument graph G_1 containing A and B where A and B attack each other. Let G_2 be the subgraph where A attacks B, let G_3 be the subgraph where B attacks A and let G_4 be the subgraph where neither argument attacks each other. So G_1 reflects the situation where the reason to believe one of them coming will cause the other to not come, G_2 reflects the situation where the reason to believe A coming will cause B to not come (though not vice versa), G_3 reflects the situation where the reason to believe B coming will cause A to not come (though not vice versa), and G_4 reflects the situation to believe the reason to believe either of them coming will not affect the other coming. Suppose, we assume the probability distribution $P(G_1) = 0.7$, $P(G_2) = 0.1$, $P(G_3) = 0.1$, and $P(G_4) = 0.1$. Hence, we get the following probabilities for the extensions. So for the grounded case, it is possible that $\{A\}$ or $\{B\}$ is the extension, and for both the grounded and preferred case, it is possible that $\{A, B\}$ is the extension.

	$\{\}$	$\{A\}$	$\{B\}$	$\{A, B\}$
Preferred	*0*	*0.8*	*0.8*	*0.1*
Grounded	*0.7*	*0.1*	*0.1*	*0.1*

As our final example for this section, we consider the sensor fusion problem again.

Example 17. *We return to Example 11. For this probabilistic attack graph, we get the following probabilities for the preferred and grounded extensions.*

	$\{\}$	$\{A\}$	$\{B\}$	$\{C\}$	$\{A,B\}$	$\{A,C\}$	$\{B,C\}$	$\{A,B,C\}$
Preferred	*0*	*0.2*	*0.4*	*0.5*	*0.6*	*0.5*	*0*	*0*
Grounded	*0.7*	*0.3*	*0*	*0*	*0*	*0*	*0*	*0*

So the most likely interpretation, with probability of 0.7, according to the grounded semantics is the empty set (i.e. none of the arguments for any of the colours can be accepted). Though with low probability, A could be the grounded extension. For the preferred semantics, $\{A, B\}$ is the most likely extension.

In the case of grounded semantics, the following result shows that the probability distribution over the grounded extensions sums to 1.

Proposition 8. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, $\sum_{\Gamma\subseteq\mathcal{A}} P_{\mathsf{gr}}(\Gamma) = 1$.*

Proof. For each $G' \sqsubseteq G$, if $G' \Vdash_{\mathsf{gr}} \Gamma_1$ and $G' \Vdash_{\mathsf{gr}} \Gamma_2$, then $\Gamma_1 = \Gamma_2$. Also for each $G' \sqsubseteq G$, there is a $\Gamma \subseteq \mathcal{A}$ such that $G' \Vdash_{\mathsf{gr}} \Gamma$. So each subgraph entails exactly one grounded extension. Recall that, by definition, $Q_{\mathsf{gr}}(\Gamma) = \{G' \sqsubseteq G \mid G' \Vdash_{\mathsf{gr}} \Gamma\}$. Therefore,

$$\text{for all } \Gamma_i, \Gamma_j \subseteq \mathcal{A}, \text{ if } \Gamma_i \neq \Gamma_j, \text{ then } Q_{\mathsf{gr}}(\Gamma_i) \cap Q_{\mathsf{gr}}(\Gamma_j) = \emptyset$$

Furthermore, for each $G' \sqsubseteq G$, there is a $\Gamma \subseteq \mathcal{A}$ such that $G' \Vdash_{\mathsf{gr}} \Gamma$. Therefore, we have the following.

$$\bigcup_{\Gamma\subseteq\mathcal{A}} Q_{\mathsf{gr}}(\Gamma) = \{G' \mid G' \sqsubseteq G\}$$

Hence, we have

$$\sum_{\Gamma\subseteq\mathcal{A}} \sum_{G'\in Q_{\mathsf{gr}}(\Gamma)} P(G') = 1 \tag{1}$$

By definition , $P_{\mathsf{gr}}(\Gamma) = \sum_{G'\in Q_{\mathsf{gr}}(\Gamma)} P(G')$, and so by substitution in (1), $\sum_{\Gamma\subseteq\mathcal{A}} P_{\mathsf{gr}}(\Gamma) = 1$. □

For the special case that the probabilistic attack graph has an assignment of zero for all its attacks, then the set of all arguments $\mathcal{A}$ is a complete, preferred, grounded, and stable extension with probability 1, as shown by the next result.

Proposition 9. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, if for all $(\alpha, \beta) \in \mathcal{R}$, $P(\alpha, \beta) = 0$, then for all $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $P_X(\mathcal{A}) = 1$*

Proof. Assume that for all $(\alpha, \beta) \in \mathcal{R}$, $p(\alpha, \beta) = 0$. Then there is a subgraph $G' \sqsubseteq G$ where G' = $(\mathcal{A}, \mathcal{R}', p)$ and $\mathcal{R}' = \emptyset$ and $P(G')$ = 1. Furthermore, since $\mathcal{R}' = \emptyset$, we have $G' \Vdash_X \mathcal{A}$ for each $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$. Therefore, $P_X(\mathcal{A}) = 1$ for each $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$. □

The next result shows that for any set of arguments Γ, it is more likely that Γ is admissible than Γ is a complete, preferred, grounded, or stable, extension.

Proposition 10. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, if $\Gamma \subseteq \mathcal{A}$, $P_X(\Gamma) \leq P_{\mathsf{ad}}(\Gamma)$, where $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$.*

Proof. For all $G' \sqsubseteq G$, and for all $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $G' \Vdash_X \Gamma$ implies $G' \Vdash_{\mathsf{ad}} \Gamma$. So, $\sum_{G'\in Q_X(\Gamma)} P(G')$ $\leq \sum_{G'\in Q_{\mathsf{ad}}(\Gamma)} P(G')$, and therefore $P_X(\Gamma) \leq P_{\mathsf{ad}}(\Gamma)$. □

In this section, we have seen how probabilistic attack graphs can be used to generate probabilistic qualification of extensions. This provides an intuitive way of analysing a variety of examples.

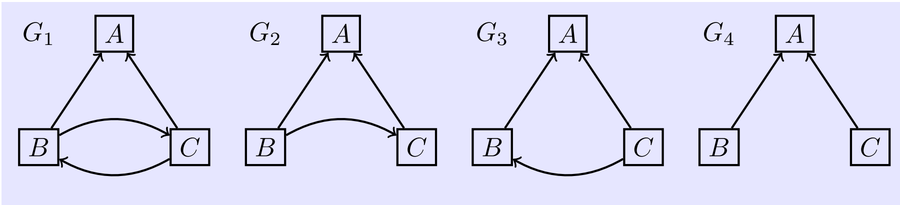


Figure 6: Argument graph G_1 and subgraphs with non-zero assignment according to the probability distribution given in Table 4 and used in Example 18.

Subgraph	Probability of subgraph	Grounded extension	Preferred extensions
G_1	0.04	$\{\}$	$\{B\}, \{C\}$
G_2	0.16	$\{B\}$	$\{B\}$
G_3	0.16	$\{C\}$	$\{C\}$
G_4	0.64	$\{B, C\}$	$\{B, C\}$

Table 4: Probability distribution for the subgraphs in Figure 6 and used in Example 18.

4.3 Probability function over inferences

Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, and an argument $\alpha \in \mathcal{A}$, we also want to calculate the probability that α is an X inference (i.e. the probability that α is in an X extension), which we denote by $P_X(\alpha)$, where $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$. For this, we use the following definition which calculates the probability of a formula being in an X extension as the sum of the probabilities for the subgraphs that entail an X extension containing the formula, and recalling $I_X(\alpha) = \{G' \mid G' \sqsubseteq G$ and $G' \Vdash_X \Gamma$ and $\alpha \in \Gamma\}$.

Definition 11. *Let $G = (\mathcal{A}, \mathcal{R}, P)$ be a probabilistic attack graph. For an argument α, the probability that it is in an X extension (where $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$) is*

$$P_X(\alpha) = \sum_{G' \in I_X(\alpha)} P(G')$$

As we have seen so far in examples, the same argument can appear in multiple extensions with non-zero probability. So the above definition for probability of inferences allows this information to be drawn out.

Example 18. *Consider the following arguments concerning a murder case which we represent in the argument graph G_1 in Figure 6*

- *A = "Andrew says he didn't shoot the victim"*
- *B = "Billy says he saw Andrew shoot the victim three times"*
- *C = "Charlie says he saw Andrew shoot the victim four times"*

Clearly B attacks A and C attacks A, and so the probability of attack is 1 for each of them. There is also a possibility that B and C attack each other since they do appear to contradict each other. But because there appears to be some imprecision in the arguments, they only attack each other with low probability. We represent this in the probability distribution in Table 4. here, we see that the independence assumption holds. We also present the three proper subgraphs with non-zero probability in Figure 6. For this probabilistic attack graph, we get the following probabilities for grounded and preferred inferences.

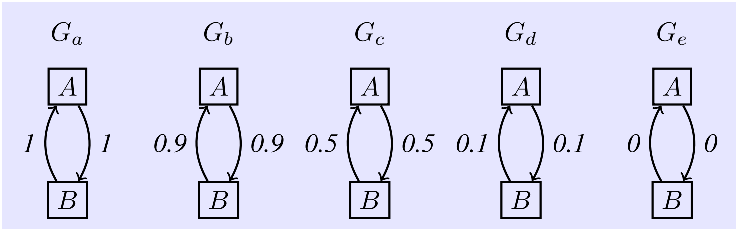
$$P_{\mathsf{pr}}(A) = 0 \quad P_{\mathsf{pr}}(B) = 0.84 \quad P_{\mathsf{pr}}(C) = 0.84$$
$$P_{\mathsf{gr}}(A) = 0 \quad P_{\mathsf{gr}}(B) = 0.8 \quad P_{\mathsf{gr}}(C) = 0.8$$

For an argument graph with the structure of G_1 we would get B and C as preferred inferences, but we would not get B and C grounded inferences. So by introducing the uncertainty into the argument graph, we are able to get a more graduated influence on the grounded extensions.

The above example shows how we get a more graded approach with probabilistic inference. Without probabilities, G_1 gives no grounded inferences, and it gives either B or C as preferred inferences, whereas with probabilities, we get B or C as grounded inferences and as preferred inferences with significant probability.

To further illustrate the effect of the probability of attack on inferences, we use a simple argument graph where A attacks B and B attacks A, and consider different combinations of probabilities.

Example 19. *The following labelled graphs have the same structure but different probabilities of attack. For each probabilistic attack graph used to produce these labelled graph, the sample space is given in Figure 2. In each case we have used a probability distribution over the subgraphs that has the attacks being independent. Therefore, for each labelled graph below, the probability distribution can be obtained from the probability of attack as explained earlier. For instance, for the case of G_a, the probability distribution would assign 1 to the graph in Figure 2a (i.e. $P(G_1) = 1$), and for the case of G_b, the assignment would be $P(G_1) = 0.81$, $P(G_2) = 0.09$, $P(G_3) = 0.09$, and $P(G_4) = 0.01$.*

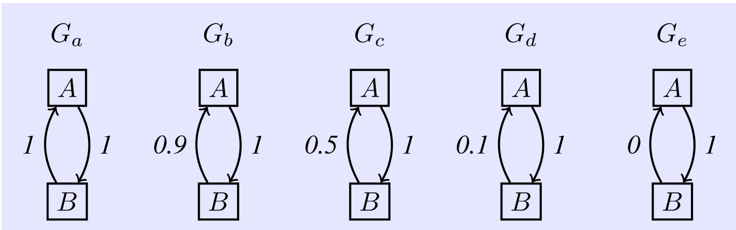


For each of these labelled graphs, the probability for A being a grounded and preferred inference is tabulated below.

	G_a	G_b	G_c	G_d	G_e
$P_{\mathsf{gr}}(A)$	0	0.10	0.50	0.90	1
$P_{\mathsf{pr}}(A)$	1	0.91	0.75	0.91	1

So here we see that the probability that A is a grounded inference rises from 0 to 1, as the probability of the symmetrical attack falls from 1 to 0, whereas the probability that A is a preferred inference falls from 1 to 0.75 as the probability of the symmetrical attack falls from 1 to 0.5, and then the probability that A is a preferred inferences rises from 0.75 to 1 as the probability of the symmetrical attack falls from 0.5 to 0.

Example 20. *The following labelled graphs have the same structure but an increasing difference between the two probabilities of attack. For each probabilistic attack graph used to produce these labelled graph, the sample space is given in Figure 2. In each case we have used a probability distribution over the subgraphs that has the attacks being independent, and so we can obtained the probability distribution over subgraphs as Example 19.*



For each of these graphs, using the probability distribution over subgraphs, the probability for the grounded and preferred inferences is tabulated below.

	G_a	G_b	G_c	G_d	G_e
$P_{\mathsf{pr}}(A)$	1	1	1	1	1
$P_{\mathsf{gr}}(A)$	0	0.1	0.5	0.9	1
$P_{\mathsf{pr}}(B)$	1	0.9	0.5	0.1	0
$P_{\mathsf{gr}}(B)$	0	0	0	0	0

Here we see that the probability that A is a grounded inference rises from 0 to 1, as the probability of the attack on it falls from 1 to 0, whereas the probability that A is a preferred inferences is always 1. In contrast, the probability that B is a grounded inferences is always 0, and the probability that B is a preferred inference falls from 1 to 0, as the probability of the attack by it on A falls from 1 to 0.

We turn to considering some properties for Definition 11. If $P(\alpha, \alpha) = 1$ holds for some argument α (i.e. α is self-attacking with probability 1), then there is no extension containing α, and so it has zero probability.

Proposition 11. *If $P(\alpha, \alpha) = 1$, then $P_X(\alpha) = 0$.*

Proof. Let G be a graph containing node α. Assume $P(\alpha, \alpha) = 1$. Therefore, for all $\Gamma \in \mathcal{A}$, if $\alpha \in \Gamma$, then Γ is not conflictfree. Therefore, for all $G' \sqsubseteq G$, for all $\Gamma \subseteq \mathcal{A}$, if $\alpha \in \Gamma$, then $G' \not\Vdash_{\mathsf{ad}} \Gamma$. Therefore, for all $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $I_X(\alpha) = \emptyset$, and hence $P_X(\alpha) = 0$. □

If an argument α is unattacked, then the probability that it is in a grounded extension is 1.

Proposition 12. *If α is unattacked, then $P_{\mathsf{gr}}(\alpha) = 1$.*

Proof. Let G be a graph containing node α. Assume α is unattacked. Therefore, for all $G' \sqsubseteq G$, α is unattacked in G'. Therefore, for all $G' \sqsubseteq G$, there is a $\Gamma \subseteq \mathcal{A}$, such that $G' \Vdash_X \Gamma$ and $\alpha \in \Gamma$. Therefore, from Proposition 8, we have that $P_{\mathsf{gr}}(\alpha) = 1$. □

The probability that an argument α is in a complete extension is greater than being in a preferred, which in turn is greater than being in a grounded extension.

Proposition 13. *For all $\alpha \in \mathcal{A}$, $P_{\mathsf{gr}}(\alpha) \leq P_{\mathsf{pr}}(\alpha)$, and $P_{\mathsf{pr}}(\alpha) \leq P_{\mathsf{co}}(\alpha)$.*

Proof. Let G be a graph containing node α. For all $G' \sqsubseteq G$, for all Γ_i, Γ_j, if $G' \Vdash_{\mathsf{gr}} \Gamma_i$ and $G' \Vdash_{\mathsf{pr}} \Gamma_j$, then $\Gamma_i \subseteq \Gamma_j$. For all $G' \sqsubseteq G$, for all Γ, if $G' \Vdash_{\mathsf{pr}} \Gamma$ then $G' \Vdash_{\mathsf{co}} \Gamma$. Therefore, $I_{\mathsf{gr}}(\alpha) \subseteq I_{\mathsf{pr}}(\alpha)$ and $I_{\mathsf{pr}}(\alpha) \subseteq I_{\mathsf{co}}(\alpha)$. Therefore, $\sum_{G' \in I_{\mathsf{gr}}(\alpha)} P(G') \leq \sum_{G' \in I_{\mathsf{pr}}(\alpha)} P(G')$, and $\sum_{G' \in I_{\mathsf{pr}}(\alpha)} P(G') \leq \sum_{G' \in I_{\mathsf{co}}(\alpha)} P(G')$. Therefore, $P_{\mathsf{gr}}(\alpha) \leq P_{\mathsf{pr}}(\alpha)$, and $P_{\mathsf{pr}}(\alpha) \leq P_{\mathsf{co}}(\alpha)$. □

The probability that $\{\alpha\}$ is an admissible set is greater than the probability that α is in an extension.

Proposition 14. *For all $\alpha \in \mathcal{A}$, $P_X(\alpha) \leq P_{\mathsf{ad}}(\{\alpha\})$.*

Proof. Let G be a graph containing node α. For all $G' \sqsubseteq G$, and for all $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $G' \Vdash_X \Gamma$ implies $G' \Vdash \Gamma$. Therefore, for all $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $I_X(\alpha) \subseteq I_{\mathsf{ad}}(\alpha)$. Hence, $\sum_{G' \in I_X(\alpha)} P(G') \leq \sum_{G' \in I_{\mathsf{ad}}(\alpha)} P(G')$, and so, $P_X(\alpha) \leq P_{\mathsf{ad}}(\{\alpha\})$. □

The next result shows that we can equivalently define the probability of a formula being in a grounded extension in terms of the probability function over grounded extensions.

Proposition 15. *Given a probabilistic attack graph $G = (\mathcal{A}, \mathcal{R}, P)$, an argument $\alpha \in \mathcal{A}$*

$$P_{\mathsf{gr}}(\alpha) = \sum_{\Gamma \subseteq \mathcal{A} \text{ } s.t. \text{ } \alpha \in \Gamma} P_{\mathsf{gr}}(\Gamma)$$

Proof. $P_{\mathsf{gr}}(\alpha)$

$= \sum_{G' \in I_{\mathsf{gr}}(\alpha)} P(G')$	(according to Definition 11)
$= \sum_{G' \in \{G' \sqsubseteq G \mid G' \Vdash_{\mathsf{gr}} \Gamma \text{ and } \alpha \in \Gamma\}} P(G')$	(by subsitution with definition for I_{gr})
$= \sum_{\Gamma \subseteq \mathcal{A} \text{ s.t. } \alpha \in \Gamma} \sum_{G' \in \{G' \sqsubseteq G \mid G' \Vdash_{\mathsf{gr}} \Gamma\}} P(G')$	(since there is one grounded extension per subgraph)
$= \sum_{\Gamma \subseteq \mathcal{A} \text{ s.t. } \alpha \in \Gamma} \sum_{G' \in Q_{\mathsf{gr}}(\Gamma)} P(G')$	(by definition for Q_{gr})
$= \sum_{\Gamma \subseteq \mathcal{A} \text{ s.t. } \alpha \in \Gamma} P_{\mathsf{gr}}(\Gamma)$	(according to Definition 10)

□

In this subsection, we have considered the probability of a formula being in an extension as the sum of the probabilities of the subgraphs that entail an extension containing the formula. As seen from the examples, this provides an additional insight into the uncertainty of attack.

5 Modelling enthymemes

In the real-world, arguments are often enthymemes (i.e. arguments with some premises being implicit). The missing premises have to be abduced by the recipient. These may be obtained from contextual and background knowledge, and more challengingly, common-sense knowledge.

To help understand this point, we consider logic-based formalisms for argumentation. There are a number of proposals for logic-based formalisations of argumentation (for reviews see [CML00, PV02, BH08]). In a number of key examples of them, an argument is a pair where the first item in the pair is a minimal consistent set of formulae that proves the second item, which is a formula. Furthermore, in these approaches, a key form of counterargument is an undercut: One argument undercuts another argument when the claim of the first argument negates the premises of the second argument. For example, the logical argument $\langle\{c, c \rightarrow \neg b\}, \neg b\rangle$ is a counterargument to the logical argument $\langle\{b, b \rightarrow a\}, a\rangle$.

We can identify the attacks relation in terms of arguments and counterarguments. So if we have logical arguments α and β, where β is a counterargument to α, then β attacks α. Because attack is defined in terms of the logic, this attack is certain. In other words, there is no doubt that β attacks α. Hence, the probability of attack is 1. Moreover, if we have a logical argumentation system, for any pair of arguments, the probability of attack is either 0 or 1.

Unfortunately, real arguments do not normally fit this mould of logical arguments. Real arguments (i.e. those presented by people in general) are normally enthymemes [Wal89]). 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 Γ is the set of premises explicitly given for an implicit support enthymeme, and α is the claim, then Γ does not entail α, but there are some implicitly assumable premises Γ' such that $\Gamma \cup \Gamma'$ is a minimal consistent set of formulae that entails α. 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 example, 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.

We can see the use of enthymemes both in monological argumentation, for example in an essay in a current affairs magazine or by a politician giving a lecture, and in dialogical argumentation, for example in a group of politicians in a debate, or a group of healthcare professionals discussing a patient.

5.1 Enthymemes arising when listening to debates

To investigate the use of probabilistic attack graphs for modelling enthymemes, we will focus in this paper on enthymemes arising when listening to debates. We use this focus in order to be able to present some formal definitions and results, but it would be reasonable to consider the solutions presented as being relevant to a wide range of situations where enthymemes arise.

When an audience is listening to participants in a discussion or debate, the participants present arguments including enthymemes. The way these are presented gives some idea to the audience of which are meant to attack which. However, if an argument being attacked is an enthymeme, and the attacking and attacked arguments are from different participants, then there is uncertainty about whether it is indeed a valid attack. Each enthymeme is a representative for an intended argument, but for the audience it may be uncertain which decoding is the intended argument. The audience may have zero or more choices. This means that the audience takes an argument graph as input, and tries to determine the intended argument graph (i.e. the graph obtained by instantiating each node with its intended argument and deleting the arcs that are not valid attacks). This intended argument graph has a structure that would be isomorphic to a spanning subgraph of the original argument graph. To illustrate our concerns, we consider the following informal example (note in the rest of the section, our solution makes this precise by dealing only with logical arguments).

Example 21. *Consider the following arguments where E_1 and E_3 are by a proponent for expanding Heathrow airport with a third runway, and E_2 is by an opponent, and so from the discussion it appears* *that E_3 attacks E_2 and E_2 attacks E_1.*

- E_1 = *We should build a third runway at Heathrow because everyone will benefit from the increased capacity.*
- E_2 = *It is not true that everyone will benefit in the community.*
- E_3 = *Local residents won't have problems with traffic because we will increase public transport to the airport.*

Let us suppose that with the common knowledge we have available, we judge that E_2 decodes to either E_2' or E_2''.

- E_2' = *It is not true that everyone will benefit in the community. There are local residents who will suffer from increased noise from the increased number of aircraft.*
- E_2'' = *It is not true that everyone will benefit in the community. There are local residents who will have problems from increased traffic on the roads to the airport.*

Given these decodings, we see that E_1 is attacked by both interpretations of E_2 whereas only one interpretation of E_2 is attacked by E_3.

In this section, we assume the arguments are given by some participants, and that somehow, these arguments are translated into an argument graph (which we call the starting graph) instantiated with logical arguments some of which may be enthymemes. This argument graph might, for instance, be constructed from watching a TV discussion where as TV viewers, the audience is unable to interact with the participants to clarify their arguments.

Our solution is to formalize how the audience of the argumentation evaluates this starting graph. We assume a knowledgebase that we believe is common knowledge for the participants. We then construct all the possible argument graphs based on the possible decodings of the enthymemes. Then we derive a probability distribution over the spanning subgraphs of the starting graph. So the probability of a subgraph is the probability that it is the correct structure (i.e. the probability that it is the correct set of nodes and arcs) for the intended argument graph (where the intended graph is the argument obtained by each participant giving all the premises for each of its arguments).

5.2 Definitions for logical argumentation

In this subsection, we review definitions for logical argumentation [BH01] and for approximations of them [Hun07].

A **logical argument** is a pair $\langle\Phi, \alpha\rangle$ where $\mathcal{L}$ is the language of propositional logic and $\Phi \subseteq \mathcal{L}$ is the support and $\alpha \in \mathcal{L}$ is the claim. This is a very general definition. It does not assume that Φ is consistent, or that it even entails α, though we will consider these as optional further conditions below. To support our use of logical arguments, we require the following definition: For a logical argument $\langle\Phi, \alpha\rangle$, let $\mathsf{Support}(\langle\Phi, \alpha\rangle)$ be Φ, and let $\mathsf{Claim}(\langle\Phi, \alpha\rangle)$ be α.

Next, we consider particular kinds of logical argument, where $\vdash$ denotes the classical consequence relation. If $\Phi \vdash \alpha$, then $\langle\Phi, \alpha\rangle$ is **supported**; If $\Phi \not\vdash \bot$, then $\langle\Phi, \alpha\rangle$ is **consistent**; If $\Phi \vdash \alpha$, and there is no $\Phi' \subset \Phi$ such that $\Phi' \vdash \alpha$, then $\langle\Phi, \alpha\rangle$ is **minimal**; And if $\Phi \vdash \alpha$, and $\Phi \not\vdash \bot$, then $\langle\Phi, \alpha\rangle$ is **valid** (i.e. it is a consistent and supported argument, but it may have unnecessary premises). A logical argument is **strict** iff it is valid and minimal. In addition, we require a further kind of logical argument that has the potential to be transformed into a valid argument: If $\Phi \not\vdash \alpha$, and $\Phi \not\vdash \neg\alpha$, then $\langle\Phi, \alpha\rangle$ is a **precursor** (i.e. it is a precursor for a valid argument). Therefore, if $\langle\Phi, \alpha\rangle$ is a precursor, then there exists some $\Psi \subset \mathcal{L}$ such that $\Phi \cup \Psi \vdash \alpha$ and $\Phi \cup \Psi \not\vdash \bot$, and hence $\langle\Phi \cup \Psi, \alpha\rangle$ is a valid argument.

Example 22. *Let $\Delta = \{a, \neg a \vee b, c, \neg b, b, \neg c, \neg b \vee c\}$. Some logical arguments from Δ are $A_1 = \langle\{a, \neg a \vee b, c, b\}, b\rangle$, $A_2 = \langle\{c, \neg c\}, b\rangle$, $A_3 = \langle\{a, \neg a \vee b, c\}, b\rangle$, $A_4 = \langle\{a, \neg a \vee b, c, \neg c\}, b\rangle$, $A_5 = \langle\{a, \neg a \vee b\}, b\rangle$, $A_6 = \langle\{\neg a \vee b\}, b\rangle$, and $A_7 = \langle\{\neg a \vee b, \neg b \vee c, \neg c\}, b\rangle$. Of these, $\{A_1, A_2, A_3, A_4, A_5\}$ are supported, $\{A_1, A_3, A_5\}$ are valid, $\{A_2, A_5\}$ are minimal, and A_5 is a strict argument. Also, $\{A_6, A_7\}$ are not supported and A_6 is a precursor.*

Now we consider enthymemes. If a proponent has a strict argument that it wishes a recipient to be aware of, which we refer to as the **intended argument**, then the proponent may send an enthymeme instead of the intended argument to the recipient. Given an intended argument, we use the following definition of an enthymeme (which is a special case of [BH12]): For this paper, a precursor $\langle \Phi, \alpha \rangle$ is an **enthymeme** for a strict argument $\langle \Psi, \alpha \rangle$ iff $\Phi \subset \Psi$. This is what we referred to earlier as an implicit support enthymeme.

Example 23. *Let u be "you need an umbrella today", and r be "the weather report predicts rain". So for an intended argument $\langle \{r, r \rightarrow u\}, u \rangle$, the enthymeme sent by the proponent to the recipient may be $\langle \{r\}, u \rangle$.*

Next we present a standard definition for logical attack which subsumes a range of definitions for logical attack such as undercut, direct undercut, and rebut.

Definition 12. *For logical arguments A and B, A is a* **defeater** *of B iff $\mathsf{Claim}(A) \vdash \neg \bigwedge \mathsf{Support}(B)$.*

Example 24. *Let $\Delta = \{a \vee b, a \leftrightarrow b, \neg a, c \rightarrow a, \neg a \wedge \neg b, a, b, c, a \rightarrow b, \neg a, \neg b, \neg c\}$.*

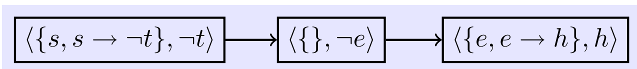
$\langle \{a \vee b, a \leftrightarrow b\}, a \wedge b \rangle$ *is a defeater of* $\langle \{\neg a, c \rightarrow a\}, \neg c \rangle$
$\langle \{a, a \rightarrow b\}, b \vee c \rangle$ *is a defeater of* $\langle \{\neg b, \neg c\}, \neg(b \vee c) \rangle$

Representing enthymemes as logical arguments makes precise the incompleteness in the argumentation. Furthermore, given a valid argument, and some knowledge that a proponent of the valid argument assumes is common knowledge, it is straightforward to construct an enthymeme.

5.3 Framework for posit graphs

For our proposal, we assume the input, which we call the **starting graph**, is a type of argument graph called a posit graph defined as follows: A **posit graph** is a tuple $(\mathcal{E}, \mathcal{S})$ where $\mathcal{E}$ is a set of logical arguments, which we call **posits**, that are either valid or precursors, and $\mathcal{S} \subseteq \mathcal{E} \times \mathcal{E}$ is a set of attacks. We assume that for each attack (E_i, E_j) in the posit graph, the proponent of E_i is different from the proponent of E_j. This means that no agent can attack his/her own arguments. Hence, we assume there are no self-cycles. Note, we do not assume that (E_i, E_j) implies that E_i is a defeater of E_j; As we suggested above, we assume that there is some external process that identifies the logical arguments and the attacks, and that this is the starting point for our formal analysis. This external process could for instance be the listener of a radio and television debate, and the listener writes down the enthymemes that he or she hears, and what appear to be attacks between those enthymemes, for instance based on the way the participants speak to each other in the debate.

Example 25. *Consider Example 21. This can be represented by the following posit graph where h = "we should build a third runway at Heathrow", e = "everyone will benefit from the increased capacity", t = "some local residents will have problems from increased traffic", and s = "there is a solution to the road traffic problem".*



We also assume that there is a set of formulae $\mathcal{K} \subseteq \mathcal{L}$, called the **cobase**. Each formula in $\mathcal{K}$ is a formula that the audience regards as being common knowledge between the participants. However, we do not impose any constraints on $\mathcal{K}$. It may be the empty set, it may be consistent, or it may be inconsistent. Furthermore, we assume it may contain common-sense knowledge, contextual knowledge that the participants are aware of, previously exchanged knowledge, etc.

Availability of a cobase is potentially an issue. But this does not appear to be a weakness of this proposal, *per se*, but rather it is a reflection of a difficult problem in modelling communications between agents in the real-world. Enthymemes occur in the real-world, and they appear to be handled by finding common knowledge. We therefore need to acknowledge the need for finding common knowledge, and then to identify ways of obtaining it.

Given a starting graph as input, we formalize the following steps in the following sections: (1) Decode the enthymemes given the available knowledge in the cobase, and use this to generate a set of candidates for the intended argument graph; (2) Interpret each candidate to take account of enthymemes that cannot be decoded given the cobase, and use this to generate a set of interpretations of the starting graph where for each enthymeme that is attacked, there is uncertainty whether or not the attack holds in the intended graph; (3) Use the set of interpretations to generate a probability distribution over the spanning subgraphs of the starting graph, and probabilistic argumentation to generate a probabilistic qualification of each of the possible extensions of the starting graph.

5.4 Decoding starting graphs

For a starting graph, we seek decodings of the posits. For this, we take a minimal subset of the cobase that when added to the support in the posit results in a valid argument.

Definition 13. *A valid argument A is a* **decoding** *of posit E with respect to cobase $\mathcal{K}$ iff* $\mathsf{Claim}(A) = \mathsf{Claim}(E)$ *and there is a $\Gamma \subseteq \mathcal{K}$ s.t.* $\mathsf{Support}(A) = \mathsf{Support}(E) \cup \Gamma$ *and it is not the case that there is a valid argument A' such that* $\mathsf{Claim}(A') = \mathsf{Claim}(E)$ *and* $\mathsf{Support}(E) \subset \mathsf{Support}(A')$ *and* $\mathsf{Support}(A') \subset \mathsf{Support}(A)$*. The set of valid arguments that are decodings of posit E with respect to $\mathcal{K}$ is* $\mathsf{Decode}(E, \mathcal{K})$.

Example 26. *Continuing Example 25, we consider the enthymeme $E_1 = \langle \{\}, \neg e \rangle$, with the cobase $\mathcal{K} = \{r, r \rightarrow n, n \rightarrow \neg e, r \rightarrow t, t \rightarrow \neg e\}$, where the additional atom is n = "some local residents will suffer from noise from the increased number of aircraft". Each of the following is a valid decoding.*

$$A_1 = \langle \{n, n \rightarrow \neg e\}, \neg e \rangle$$
$$A_2 = \langle \{t, t \rightarrow \neg e\}, \neg e \rangle$$

Clearly, for a posit E, and a cobase $\mathcal{K}$, if $\mathcal{K} = \emptyset$, and E is not a valid argument, then $\mathsf{Decode}(E, \mathcal{K}) = \emptyset$, if E is a valid argument, then $\mathsf{Decode}(E, \mathcal{K}) = \{E\}$, and if $\mathcal{K} \subseteq \mathcal{K}'$, then $\mathsf{Decode}(E, \mathcal{K}) \subseteq \mathsf{Decode}(E, \mathcal{K}')$.

Next, we use decoding of posits to decode a posit graph, and provide an illustration in Fig 7.

Definition 14. *A posit graph $G' = (\mathcal{E}', \mathcal{S}')$ is a* **graph decoding** *of a posit graph $G = (\mathcal{E}, \mathcal{S})$ iff there is a bijection $F : \mathcal{E} \rightarrow \mathcal{E}'$ s.t.*

- *for each $E \in \mathcal{E}$, if* $\mathsf{Decode}(E, \mathcal{K}) \neq \emptyset$*, then* $F(E) \in \mathsf{Decode}(E, \mathcal{K})$*, otherwise $F(E) = E$*
- *for each $(F(E_i), F(E_j)) \in \mathcal{S}'$, $(E_i, E_j) \in \mathcal{S}$*

If G' is a graph decoding of G, and $F : \mathcal{E} \rightarrow \mathcal{E}'$ is a bijection such that the above two conditions are satisfied, then we call F the **decoding function** *from G to G'.*

Once an enthymeme is decoded to a valid argument, it might be the case that an attack on it is not a logical attack. To address this, we define a candidate as a kind of graph decoding in which the non-logical attacks are dropped.

Definition 15. *A posit graph $G' = (\mathcal{E}', \mathcal{S}')$ is a* **candidate** *for a posit graph $G = (\mathcal{E}, \mathcal{S})$ iff G' is a decoding of G and $F : \mathcal{E} \rightarrow \mathcal{E}'$ is the decoding function from G to G' such that for each $(E_i, E_j) \in \mathcal{S}$,*

- *if $F(E_j)$ is a valid argument and $F(E_i)$ is a defeater of $F(E_j)$, then $(F(E_i), F(E_j)) \in \mathcal{S}'$.*
- *if $F(E_j)$ is a valid argument and $F(E_i)$ is not a defeater of $F(E_j)$, then $(F(E_i), F(E_j)) \notin \mathcal{S}'$.*
- *if $F(E_j)$ is a precursor, then $(F(E_i), F(E_j)) \in \mathcal{S}'$.*

Let $\mathsf{Candidates}(G, \mathcal{K})$ *be the set of candidates for G.*

So a candidate has the same structure as a spanning subgraph of the original posit graph, and for each posit, if there is a decoding of it, then one of these decodings is in the candidate. Also, if every posit in a posit graph is valid, then we call it a **valid graph** as illustrated in Figure 7.

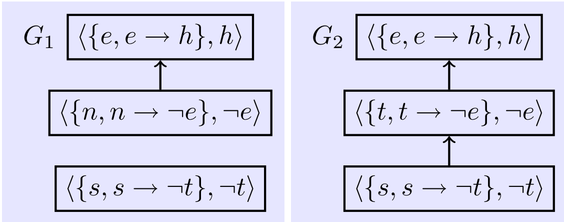


Figure 7: Consider the posit graph G in Example 25. Using the decodings in Example 26, we obtain the graph decodings G_1 and G_2 above. Each is a candidate and each is valid.

Proposition 16. *Let G be a posit graph. If $\mathcal{K} = \mathcal{L}$, then each $G' \in \mathsf{Candidates}(G, \mathcal{K})$ is a valid graph. If $\mathcal{K} = \emptyset$, then $\mathsf{Candidates}(G, \mathcal{K}) = \{G\}$.*

Proof. Let G be a posit graph. (Case 1) Assume $\mathcal{K} = \mathcal{L}$. Therefore, for each $E \in \mathsf{Nodes}(G)$, $\mathsf{Decode}(E, \mathcal{K}) \neq \emptyset$. So for each graph decoding G' of G, there is an $E' \in \mathsf{Decode}(E, \mathcal{K})$ s.t. $F(E) = E'$. Therefore, for each candidate $G'' \in \mathsf{Candidates}(G, \mathcal{K})$ is a valid graph. (Case 2) Assume $\mathcal{K} = \emptyset$. Therefore, for each $E \in \mathsf{Nodes}(G)$, $\mathsf{Decode}(E, \mathcal{K}) = \emptyset$ or $\mathsf{Decode}(E, \mathcal{K}) = \{E\}$. So there is exactly one graph decoding of G, which is G. Therefore, $\mathsf{Candidates}(G, \mathcal{K}) = \{G\}$. □

In this subsection, we have shown how we can generate a set of graph decodings, called candidates, from a starting graph as input. Clearly, decoding needs common knowledge, and if there is insufficient common knowledge, then enthymemes remain with missing premises, but if there is too much knowledge, then there is ambiguity (i.e. multiple choices for the decoding).

5.5 Interpreting candidates

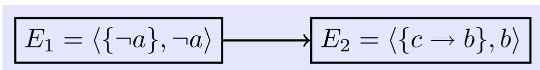
Given a candidate $G' = (\mathcal{E}', \mathcal{S}')$ in $\mathsf{Candidates}(G, \mathcal{K})$, we obtain the partition of the attack relation in G' as follows.

$$\mathsf{Closed}(G') = \{(E_i, E_j) \in \mathcal{S}' \mid \{\mathsf{Claim}(E_i)\} \cup \mathsf{Support}(E_j) \vdash \bot\}$$
$$\mathsf{Open}(G') = \{(E_i, E_j) \in \mathcal{S}' \mid \{\mathsf{Claim}(E_i)\} \cup \mathsf{Support}(E_j) \not\vdash \bot\}$$

So $\mathsf{Closed}(G')$ contains the arcs that are definitely logical attacks, whereas $\mathsf{Open}(G')$ contains the arcs where E_j is a precursor and it is not known what the missing premises are for it, and hence it is unknown whether the attack is indeed a logical attack. We use this partition to give the following set of options for arcs in the candidate. An option is a possible set of arcs that could hold for a candidate (given an appropriate completion of the missing premises).

$$\mathsf{Options}(G') = \{\mathsf{Closed}(G') \cup \Pi \mid \Pi \in \wp(\mathsf{Open}(G'))\}$$

Example 27. *Consider the starting graph G below. Let $\mathcal{K} = \emptyset$. So $\mathsf{Candidates}(G, \mathcal{K}) = \{G\}$, and $\mathsf{Options}(G) = \{\{(E_1, E_2)\}, \{\}\}$.*



From each candidate, we can form the set of posit graphs (which we call combos) where for each of them, the set of arcs is an option.

Definition 16. *For a posit graph $G' = (\mathcal{E}', \mathcal{S}')$, the set of* **combos** *is $\mathsf{Combos}(G') = \{(\mathcal{E}', \mathcal{S}'') \mid \mathcal{S}'' \in \mathsf{Options}(G')\}$.*

The set of interpretations for a starting graph is based on the set of combos that can be formed from each candidate for the starting graph. So for each candidate, we consider all possible ways that the arcs in $\mathsf{Open}(G')$ can be dropped for each candidate (since we are uncertain whether they hold or not).

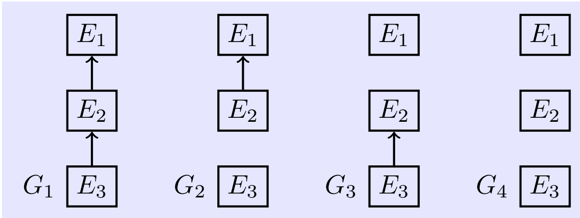


Figure 8: A graph $G' = (\mathcal{A}', \mathcal{R}')$ is a **spanning subgraph** of a graph $G = (\mathcal{A}, \mathcal{R})$, denoted $G' \sqsubseteq G$, when $\mathcal{R}' \subseteq \mathcal{R}$. For example, the spanning subgraphs of G_1 are G_1 to G_4.

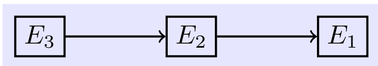
Definition 17. *For a starting graph $G = (\mathcal{E}, \mathcal{S})$, and a cobase $\mathcal{K}$, the set of* **interpretations** *is*

$$\mathsf{Interpretations}(G, \mathcal{K}) = \bigcup_{G' \in \mathsf{Candidates}(G,\mathcal{K})} \mathsf{Combos}(G')$$

Example 28. *Continuing Ex 27,* $\mathsf{Interpretations}(G, \mathcal{K}) = \{G, G'\}$ *where* $G' \sqsubset G$ *and* $\mathsf{Arcs}(G') = \emptyset$.

Example 29. *Consider the starting graph G_1, and its subgraphs in Figure 8. Let $\mathcal{K} = \emptyset$, and so for each E_i,* $\mathsf{Decode}(E_i, \mathcal{K}) = \{\}$*. So,* $\mathsf{Candidates}(G_1, \mathcal{K}) = \{G_1\}$*,* $\mathsf{Closed}(G_1) = \emptyset$*,* $\mathsf{Open}(G_1) = \{(E_3, E_2), (E_2, E_1)\}$*,* $\mathsf{Options}(G_1) = \wp(\{(E_3, E_2), (E_2, E_1)\})$*, and therefore* $\mathsf{Interpretations}(G_1, \mathcal{K}) = \{G_1, G_2, G_3, G_4\}$.

Example 30. *Continuing the example in Figure 7 with the starting graph below. So,* $\mathsf{Candidates}(G, \mathcal{K}) = \{G_1, G_2\}$, $\mathsf{Closed}(G_1) = \{(E_2, E_1)\}$, $\mathsf{Closed}(G_2) = \{(E_3, E_2), (E_2, E_1)\}$, $\mathsf{Open}(G_1) = \emptyset$*, and* $\mathsf{Open}(G_2) = \emptyset$*. So* $\mathsf{Options}(G_1) = \{\mathsf{Closed}(G_1)\}$ *and* $\mathsf{Options}(G_2) = \{\mathsf{Closed}(G_2)\}$*. Therefore, the set of interpretations is* $\mathsf{Interpretations}(G, \mathcal{K}) = \{G_1, G_2\}$.



In the following result we show that if all the arguments in a candidate are valid, then each candidate is an interpretation. At the other extreme, if all the arguments in a candidate are precursors, then each spanning subgraph of that candidate is an interpretation. So for each candidate there are between 1 and 2^n interpretations where n is the number of attacks (i.e. arcs) in the candidate.

Proposition 17. *For a candidate $G' = (\mathcal{E}', \mathcal{S}')$,*

- *If* $\mathsf{Open}(G') = \emptyset$*, then* $\mathsf{Combos}(G') = \{G'\}$
- *If* $\mathsf{Closed}(G') = \emptyset$*, then* $\mathsf{Combos}(G') = \{G'' \mid G'' \sqsubseteq G'\}$

Proof. Let $G' = (\mathcal{E}', \mathcal{S}')$ be a candidate. (Case 1) Assume $\mathsf{Open}(G') = \emptyset$. Therefore, $\mathsf{Closed}(G') = \mathcal{S}'$. Hence, $\mathsf{Options}(G') = \{\mathsf{Closed}(G') \cup \emptyset\} = \{\mathcal{S}' \cup \emptyset\} = \{\mathcal{S}'\}$. Therefore, $\mathsf{Combos}(G') = \{(\mathcal{E}', \mathcal{S}'') \mid \mathcal{S}'' \in \mathsf{Options}(G')\} = \{(\mathcal{E}', \mathcal{S}')\} = \{G'\}$. (Case 2) Assume $\mathsf{Closed}(G') = \emptyset$. So, $\mathsf{Open}(G') = \mathcal{S}'$. So, $\mathsf{Options}(G') = \{\mathsf{Closed}(G') \cup \Pi \mid \Pi \in \wp(\mathcal{S}')\} = \wp(\mathcal{S}')$. So, $\mathsf{Combos}(G') = \{(\mathcal{E}', \mathcal{S}'') \mid \mathcal{S}'' \in \mathsf{Options}(G')\} = \{(\mathcal{E}', \mathcal{S}'') \mid \mathcal{S}'' \in \wp(\mathcal{S}')\} = \{G'' \mid G'' \sqsubseteq G'\}$. □

Proposition 18. *If there is a valid graph in* $\mathsf{Candidates}(G, \mathcal{K})$*, then all graphs in* $\mathsf{Candidates}(G, \mathcal{K})$ *are valid graphs.*

Proof. Let G be a posit graph and let $\mathcal{K}$ be a cobase. Assume there is a $G' \in \mathsf{Candidates}(G, \mathcal{K})$ such that G' is valid. So, for each $E \in \mathsf{Nodes}(G)$, $\mathsf{Decode}(E, \mathcal{L}) \neq \emptyset$. So, for each graph decoding G'' of G, and for each $E \in \mathsf{Nodes}(G)$, $F(E)$ is valid. So, for each $G'' \in \mathsf{Candidates}(G, \mathcal{K})$, and for each $E \in \mathsf{Nodes}(G)$, $F(E)$ is valid. So, each $G'' \in \mathsf{Candidates}(G, \mathcal{K})$ is valid. □

Proposition 19. *For posit graph G, there is a candidate in* $\mathsf{Candidates}(G,\mathcal{K})$ *that is valid iff*

$$\mathsf{Interpretations}(G,\mathcal{K}) = \mathsf{Candidates}(G,\mathcal{K})$$

Proof. Let G be a posit graph and let $\mathcal{K}$ be a cobase. ($\Rightarrow$) Assume there is a candidate in $\mathsf{Candidates}(G,\mathcal{K})$ that is valid. Therefore, by Proposition 18, all graphs in $\mathsf{Candidates}(G,\mathcal{K})$ are valid graphs. Therefore, for each graph $G' \in \mathsf{Candidates}(G,\mathcal{K})$, $\mathsf{Open}(G') = \emptyset$, and so by Proposition 17, $\mathsf{Combos}(G') = \{G'\}$. By Definition 17, $\mathsf{Interpretations}(G,\mathcal{K}) = \bigcup_{G' \in \mathsf{Candidates}(G,\mathcal{K})} \mathsf{Combos}(G')$. Therefore, by substituting $\{G'\}$ for $\mathsf{Combos}(G')$, we get $\mathsf{Interpretations}(G,\mathcal{K}) = \bigcup_{G' \in \mathsf{Candidates}(G,\mathcal{K})} \{G'\}$ So, $\mathsf{Interpretations}(G,\mathcal{K}) = \mathsf{Candidates}(G,\mathcal{K})$. ($\Leftarrow$) Assume $\mathsf{Interpretations}(G,\mathcal{K}) = \mathsf{Candidates}(G,\mathcal{K})$. Therefore, for each graph $G' \in \mathsf{Candidates}(G,\mathcal{K})$, $\mathsf{Combos}(G') = \{G'\}$. Therefore, for each graph $G' \in \mathsf{Candidates}(G,\mathcal{K})$, and for each node E' in the graph G', E' is valid. Therefore, for each graph $G' \in \mathsf{Candidates}(G,\mathcal{K})$, G' is valid. Hence, there is a candidate in $\mathsf{Candidates}(G,\mathcal{K})$ that is valid. □

In this subsection, we have shown how we can take a set of candidates, and determine the set of interpretations, where each interpretation gives a possible structure for the intended argument graph.

5.6 Probabilistic analysis of interpretations

We now use the probabilistic attack graph to capture the uncertainty associated with decoding enthymeme graphs. Essentially, we identify an appropriate subgraph probability distribution for the starting graph.

Definition 18. *Let* $G = (\mathcal{E},\mathcal{S})$ *and* $G' = (\mathcal{E}',\mathcal{S}')$ *be posit graphs, with* $G' \sqsubseteq G$*, and let* $G'' = (\mathcal{E}'',\mathcal{S}'')$ *be an interpretation of G.* G'' **reflects** G' *iff there is a bijection* $F : \mathcal{E}' \rightarrow \mathcal{E}''$ *such that for all* $E_i, E_j \in \mathcal{E}'$

$$(E_i, E_j) \in \mathcal{S}' \textit{ iff } (F(E_i), F(E_j)) \in \mathcal{S}''$$

When G'' reflects G', G'' is a graph that is isomorphic to G' and each argument in G' is mapped to a decoding of it in G'' (or itself if no decoding exists).

Definition 19. *, Let G be a posit graph. A* **simple distribution** *is a subgraph probability distribution* $P : \{G' \mid G' \sqsubseteq G\} \rightarrow [0,1]$ *such that for each* $G' \sqsubseteq G$*,* $P(G')$ *is*

$$\frac{|\{G'' \in \mathsf{Interpretations}(G,\mathcal{K}) \mid G'' \textit{ reflects } G'\}|}{|\mathsf{Interpretations}(G,\mathcal{K})|}$$

Proposition 20. *If* $P : \{G' \mid G' \sqsubseteq G\} \rightarrow [0,1]$ *is a simple distribution, then* $[\sum_{G' \sqsubseteq G} P(G')] = 1$

Proof. Let $\mathsf{Part}(G_i) = \{G'' \in \mathsf{Interpretations}(G,\mathcal{K}) \mid G'' \text{ reflects } G_i\}$. For all G_i, $G_j \sqsubseteq G$, if $G_i \neq G_j$, then $\mathsf{Part}(G_i) \cap \mathsf{Part}(G_j) = \emptyset$. Furthermore, $\bigcup_{G_i \sqsubseteq G} \mathsf{Part}(G_i) = \mathsf{Interpretations}(G,\mathcal{K})$. Therefore, if $\{G' \mid G' \sqsubseteq G\}$ is $\{G_1, ..., G_n\}$, then $\mathsf{Part}(G_1), ..., \mathsf{Part}(G_n)$ is a partition of $\mathsf{Interpretations}(G,\mathcal{K})$. Hence, $[\sum_{G' \sqsubseteq G} P(G')] = 1$. □

Next, we consider two scenarios for a starting graph containing an enthymeme. The first scenario arises when there is no decoding of the enthymeme, and the second scenario arises when there are multiple decodings of the enthymeme.

Example 31. *Consider the starting graph G in Example 25. Let* $\mathsf{Nodes}(G) = \{E_1, E_2, E_3\}$ *and* $\mathsf{Arcs}(G) = \{(E_3, E_2), (E_2, E_1)\}$*. Let* $G' \sqsubset G$ *be s.t.* $\mathsf{Arcs}(G') = \{(E_2, E_1)\}$*.*

- *(Scenario 1) If* $\mathcal{K} = \emptyset$*, then* $\mathsf{Interpretations}(G,\mathcal{K}) = \{G, G'\}$*.*
- *(Scenario 2) For the candidates in Figure 7 (i.e.* $\mathsf{Candidates}(G,\mathcal{K}) = \{G_1, G_2\}$*), we have* G_2 *reflects* G *and* G_1 *reflects* G'*, and so* $\mathsf{Interpretations}(G,\mathcal{K}) = \{G_1, G_2\}$*.*

For each scenario, the simple distribution is therefore $P(G) = 1/2$ *and* $P(G') = 1/2$*. So,* $P_X(\{E_1, E_3\}) = P_X(\{E_2, E_3\}) = 1/2$ *for* $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$*.*

In general, $0 \leq P(G') \leq 1$ for any $G' \sqsubseteq G$, and for any cobase. For instance, $P(G') = 1$ iff for all $G'' \in \mathsf{Interpretations}(G, \mathcal{K})$, G'' reflects G', and $P(G') = 0$ iff for all $G'' \in \mathsf{Interpretations}(G, \mathcal{K})$, it is not the case that G'' reflects G'. However, when the cobase is empty, we get the following result.

Proposition 21. *For a posit graph $G = (\mathcal{E}, \mathcal{S})$, let $n = |\mathcal{S}|$, let cobase $\mathcal{K} = \emptyset$, and let P be the simple distribution. For $G' \sqsubseteq G$, $1/2^n \leq P(G') \leq 1$*

Proof. Let G be a posit graph and let $\mathcal{K}$ be a cobase, with $\mathcal{K} = \emptyset$. So by Proposition 16, $\mathsf{Candidates}(G, \mathcal{K}) = \{G\}$. So for each node E in G, $\mathsf{Decode}(E) = \{E\}$, when E is valid, and $\mathsf{Decode}(E) = \emptyset$, when E is a precursor. Let $m = |\{(E_i, E_j) \in \mathsf{Arcs}(G) \mid E_j \text{ is a precursor }\}|$. So $|\mathsf{Open}(G)| \leq m$. Therefore, $|\mathsf{Interpretations}(G, \mathcal{K})| \leq 2^m$. Hence, for $G' \sqsubseteq G$, $1/2^m \leq P(G') \leq 1$. Since, $m \leq n$, we have $1/2^n \leq P(G') \leq 1$. □

Proposition 22. *For a posit graph G, and $G' \sqsubseteq G$, if $P(G') = 1$ where P is the simple distribution, then for all $G'' \in \mathsf{Interpretations}(G, \mathcal{K})$, G'' is valid.*

Proof. Let G be a posit graph, with $G' \sqsubseteq G$, and let $\mathcal{K}$ be a cobase. Assume $P(G') = 1$. Therefore, from the definition of the simple distribution, for all $G'' \in \mathsf{Interpretations}(G, \mathcal{K})$, G'' reflects G'. Therefore, for all $G'' \in \mathsf{Interpretations}(G, \mathcal{K})$, G'' is valid. □

By harnessing subgraph probability distributions, we are able to provide a simple and effective qualification of the uncertainty associated with decoding enthymemes.

5.7 Uncertainty of attack in posit graphs

So far we have considered the probability distribution over the subgraphs of a starting graph. However, we may wish to consider the uncertainty associated with an attack relationship in a starting graph, and in particular assign a probability value to the attack. For this, we first consider some subsidiary definitions below which we will use for some postulates for our probability assignments to attacks.

Definition 20. *Let $G = (\mathcal{E}, \mathcal{S})$ be a posit graph. For $(E_i, E_j) \in \mathcal{S}$, the* **set of potential attacks***, denoted* $\mathsf{Potential}(E_i, E_j)$*, is defined as follows.*

$$\{(E_i', E_j') \mid E_i' \in \mathsf{Poss}(E_i, \mathcal{K}) \text{ and } E_j' \in \mathsf{Poss}(E_j, \mathcal{K})\}$$

where $\mathsf{Poss}(E, \mathcal{K}) = \{E\}$ *when* $\mathsf{Decode}(E, \mathcal{K}) = \emptyset$*, and* $\mathsf{Poss}(E, \mathcal{K}) = \mathsf{Decode}(E, \mathcal{K})$ *when* $\mathsf{Decode}(E, \mathcal{K}) \neq \emptyset$.

Next, we provide a partitioning of $\mathsf{Potential}(E_i, E_j)$ into the set of valid attacks when the decoding results in a logical attack holding, the set of invalid attacks when the decoding results in the claim of the attacker being consistent with the support of the attacked valid argument, and the set of unknown attacks when the attacked argument is a precursor.

Definition 21. *Let $G = (\mathcal{E}, \mathcal{S})$ be a posit graph. For $(E_i, E_j) \in \mathcal{S}$, and for each $(E_i', E_j') \in \mathsf{Potential}(E_i, E_j)$,*

- $(E_i', E_j') \in \mathsf{Valid}(E_i, E_j)$ *iff E_i' is a defeater of E_j'*
- $(E_i', E_j') \in \mathsf{Invalid}(E_i, E_j)$ *iff (E_i' is not a defeater of E_j') and E_j' is valid*
- $(E_i', E_j') \in \mathsf{Unknown}(E_i, E_j)$ *iff (E_i' is not a defeater of E_j') and E_j' is a precursor.*

Proposition 23. *For a posit graph $G = (\mathcal{E}, \mathcal{S})$, and for each $(E_i, E_j) \in \mathcal{S}$, exactly one of the following hold.*

- $\mathsf{Potential}(E_i, E_j) = \mathsf{Valid}(E_i, E_j) \cup \mathsf{Invalid}(E_i, E_j)$
- $\mathsf{Potential}(E_i, E_j) = \mathsf{Unknown}(E_i, E_j)$

Postulate	Definition
P1	If $\mathsf{Potential}(E_i, E_j) = \mathsf{Valid}(E_i, E_j)$ then $P(E_i, E_j) = 1$
P2	If $\mathsf{Potential}(E_i, E_j) = \mathsf{Invalid}(E_i, E_j)$ then $P(E_i, E_j) = 0$
P3	If $\mathsf{Potential}(E_i, E_j) = \mathsf{Unknown}(E_i, E_j)$ then $P(E_i, E_j) = 0.5$
P4	If $\mathsf{Valid}(E_i, E_j) \neq \emptyset$ then $P(E_i, E_j) = \lvert\mathsf{Valid}(E_i, E_j)\rvert / \lvert\mathsf{Potential}(E_i, E_j)\rvert$

Table 5: Postulates for the probability of attack where $(\mathcal{E}, \mathcal{S})$ is a posit graph and $(E_i, E_j) \in \mathcal{S}$. We explain these as follows: (P1) If all the possible interpretations for attacks are valid attacks, then the probability of attack is 1; (P2) If all the possible interpretations for attacks are invalid attacks, then the probability of attack is 0; (P3) If an attack is unknown (i.e. the attacked argument is a precursor where the support is not negated by the claim of the attacker), then the ignorance is captured by a probability of attack of 0.5; (P4) If some of the potential attacks are valid, then the probability of attack is the proportion of potential attacks that are valid attacks.

Proof. Let $G = (\mathcal{E}, \mathcal{S})$ be a posit graph and let $(E_i, E_j) \in \mathcal{S}$ be an attack.

(Case 1) Suppose $(E'_i, E'_j) \in \mathsf{Valid}(E_i, E_j) \cup \mathsf{Invalid}(E_i, E_j)$. So either ($E'_i$ is a defeater of E'_j) or (E'_i is a not defeater of E'_j and E'_j is valid). (Case 1a) Suppose $\mathsf{Decode}(E_j, \mathcal{K}) = \emptyset$. Therefore, for all $(E''_i, E''_j) \in \mathsf{Potential}(E_i, E_j)$, E''_j is E'_j and E'_j is E_j. So E''_j is a precursor and therefore E''_i is a defeater of E''_j. So $(E''_i, E''_j) \in \mathsf{Valid}(E_i, E_j) \cup \mathsf{Invalid}(E_i, E_j)$. (Case 1b) Suppose $\mathsf{Decode}(E_j, \mathcal{K}) \neq \emptyset$. Therefore, for all $(E''_i, E''_j) \in \mathsf{Potential}(E_i, E_j)$, $E''_j \in \mathsf{Decode}(E_j, \mathcal{K})$. Hence, E''_j is valid. So either (E''_i is a defeater of E''_j) or (E'_i is a not defeater of E'_j and E'_j is valid). So $(E''_i, E''_j) \in \mathsf{Valid}(E_i, E_j) \cup \mathsf{Invalid}(E_i, E_j)$.

(Case 2) Suppose $(E'_i, E'_j) \in \mathsf{Unknown}(E_i, E_j)$. So E'_j is a precursor and E'_j is E_j. Therefore, $\mathsf{Decode}(E_j, \mathcal{K}) = \emptyset$. Therefore, for all $(E''_i, E''_j) \in \mathsf{Potential}(E_i, E_j)$, E''_j is E'_j and E'_j is E_j. and E''_i is not a defeater of E''_j. Therefore, $(E''_i, E''_j) \in \mathsf{Unknown}(E_i, E_j)$.

From Case 1 and Case 2, we obtain that for each $(E_i, E_j) \in \mathcal{S}$, $\mathsf{Potential}(E_i, E_j)$ is exactly one of $\mathsf{Valid}(E_i, E_j) \cup \mathsf{Invalid}(E_i, E_j)$ and $\mathsf{Unknown}(E_i, E_j)$. □

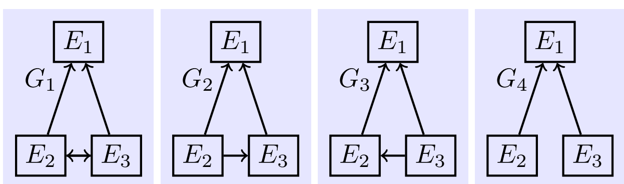
Now we give our postulates for the probability of attack in Table 5, and then show how they correspond to the marginal distribution obtained from the simple distribution given in the previous section. Given a subgraph probability distribution, such as the simple distribution given in the previous section, the marginal distribution for a particular attack relationship in the posit graph is the sum of the probabilities assigned to each subgraph that contains that attack relationship.

Definition 22. *For a posit graph $G = (\mathcal{E}, \mathcal{S})$, and a probability distribution P over $\{G' \mid G' \sqsubseteq G\}$, the marginal distribution for an attack $(E_i, E_j) \in \mathcal{S}$ is the following, where $\mathsf{Arcs}(G')$ is the set of arcs in G'.*

$$P(E_i, E_j) = \sum_{G' \text{ s.t. } G' \sqsubseteq G \text{ and } (E_i, E_j) \in \mathsf{Arcs}(G')} P(G')$$

Example 32. *Continuing Example 29, $P(E_2, E_1) = P(G_1) + P(G_2) = 1/2$.*

Example 33. *Consider the posit graph G_1, and three of its spanning subgraphs, below. Suppose, $E_1 = \langle\{b, c, \neg b \vee \neg c \vee a\}, a\rangle$, $E_2 = \langle\{\}, \neg b\rangle$, and $E_3 = \langle\{g\}, \neg c\rangle$.*



So E_1 decodes to itself. Suppose that there are three decodings of E_2: $E_2^1 = \langle\{c, c \rightarrow \neg b\}, \neg b\rangle$, $E_2^2 = \langle\{d, d \rightarrow \neg b\}, \neg b\rangle$, and $E_2^3 = \langle\{e, \neg b \rightarrow \neg e\}, \neg b\rangle$. Also suppose that there are no decodings of E_3. So there are three candidates. Each with E_1, E_3, and one of E_2^1, E_2^2 or E_2^3. Because, E_3 is an enthymeme, there are two interpretations per candidate:

Interpretation	*Nodes*	*Reflects subgraph*
G^1	E_1, E_2^1, E_3	G_1
G^2	E_1, E_2^1, E_3	G_3
G^3	E_1, E_2^2, E_3	G_2
G^4	E_1, E_2^2, E_3	G_4
G^5	E_1, E_2^3, E_3	G_2
G^6	E_1, E_2^3, E_3	G_4

Therefore, the simple distribution is $P(G_1) = 1/6$, $P(G_2) = 1/3$, $P(G_3) = 1/6$, and $P(G_4) = 1/3$. Therefore, $P_{\mathsf{gr}}(\{\}) = 1/6$, $P_{\mathsf{gr}}(\{E_2\}) = 1/3$, $P_{\mathsf{gr}}(\{E_3\}) = 1/6$, and $P_{\mathsf{gr}}(\{E_2, E_3\}) = 1/3$. Furthermore, $P_{\mathsf{pr}}(\{\}) = 0$, $P_{\mathsf{pr}}(\{E_2\}) = 1/2$, $P_{\mathsf{pr}}(\{E_3\}) = 1/3$, and $P_{\mathsf{pr}}(\{E_2, E_3\}) = 1/3$. Finally, $P(E_2, E_1) = 1$, $P(E_3, E_1) = 1$, $P(E_2, E_3) = 1/2$, and $P(E_3, E_2) = 1/3$.

We now consider our characterization theorem that relates the marginal distribution obtained from the simple distribution to the postulates given above.

Proposition 24. *Let $(\mathcal{E}, \mathcal{S})$ be a posit graph and let P be a probability function $P : \mathcal{S} \rightarrow [0, 1]$. P is the marginal distribution over $\mathcal{S}$ obtained from the simple distribution over $\{G' \mid G' \sqsubseteq G\}$ iff P satisfies Postulates P1 to P4 (Table 5).*

Proof. ($\Rightarrow$) From Proposition 23, for a posit graph $G = (\mathcal{E}, \mathcal{S})$, and for each $(E_i, E_j) \in \mathcal{S}$, exactly one of the following hold: $\mathsf{Potential}(E_i, E_j) = \mathsf{Valid}(E_i, E_j) \cup \mathsf{Invalid}(E_i, E_j)$ or $\mathsf{Potential}(E_i, E_j) = \mathsf{Unknown}(E_i, E_j)$. We proceed by assuming each of these cases, and show that the postulates hold.

(Case 1) Assume $\mathsf{Potential}(E_i, E_j) = \mathsf{Valid}(E_i, E_j) \cup \mathsf{Invalid}(E_i, E_j)$. Therefore, the proportion of candidates containing a decoding of (E_i, E_j), i.e. the proportion of candidates containing an arc from $\mathsf{Potential}(E_i, E_j)$, is given by Definition 22 (assuming the simple distribution)

$$P(E_i, E_j) = \sum_{G' \text{ s.t. } G' \sqsubseteq G \text{ and } (E_i, E_j) \in \mathsf{Arcs}(G)} P(G')$$

Since each candidate with an arc corresponding to (E_i, E_j) has a decoding from $\mathsf{Valid}(E_i, E_j)$ and each candidate without an arc corresponding to (E_i, E_j) has a decoding from $\mathsf{Invalid}(E_i, E_j)$, the proportion of candidates with an arc corresponding to (E_i, E_j) is the following, and hence we obtain P4.

$$P(E_i, E_j) = \frac{|\mathsf{Valid}(E_i, E_j)|}{|\mathsf{Valid}(E_i, E_j) + \mathsf{Invalid}(E_i, E_j)|} = \frac{|\mathsf{Valid}(E_i, E_j)|}{|\mathsf{Potential}(E_i, E_j)|}$$

Furthermore, if $\mathsf{Invalid}(E_i, E_j) = \emptyset$, the above implies P1, and if $\mathsf{Valid}(E_i, E_j) = \emptyset$, the above implies P2.

(Case 2) Assume $\mathsf{Potential}(E_i, E_j) = \mathsf{Unknown}(E_i, E_j)$. Therefore, there are no decodings of E_j in (E_i, E_j), i.e. $\mathsf{Decode}(E_j, \mathcal{K}) = \emptyset$. Therefore, for all candidates $G' \in \mathsf{Candidates}(G, \mathcal{K})$, $F(E_j)$ is a precursor. Therefore, $(E'_i, E'_j) \in \mathsf{Open}(G')$. Therefore, from Definition 17, half the interpretations G'' of G' will have the arc (E'_i, E'_j), and half the interpretations G'' of G' will not have the arc (E'_i, E'_j). Therefore, half of all the interpretations in $\mathsf{Interpretations}(G, \mathcal{K})$ will have the arc (E'_i, E'_j) and half of all the interpretations in $\mathsf{Interpretations}(G, \mathcal{K})$ will not have the arc (E'_i, E'_j). Therefore, by Definition 19, half the mass assigned by the simple distribution will be assigned to subgraphs of G that contain (E_i, E_j) and half the mass assigned by the simple distribution will be assigned to subgraphs of G that do not contain (E_i, E_j). Therefore, by Definition 22, $P(E_i, E_j) = 0.5$. Hence, implying P3.

($\Leftarrow$) (Case 1) Assume postulates P1, P2 and P4. For any $(E_i, E_j) \in \mathcal{S}$ where $\mathsf{Unknown}(E_i, E_j) = \emptyset$, these postulates imply the following.

$$P(E_i, E_j) = \frac{|\mathsf{Valid}(E_i, E_j)|}{|\mathsf{Valid}(E_i, E_j) + \mathsf{Invalid}(E_i, E_j)|}$$

Therefore, $P(E_i, E_j)$ is the proportion of candidates with an arc corresponding to (E_i, E_j) having a decoding from $\mathsf{Valid}(E_i, E_j)$ normalized by the total number of candidates. Since $\mathsf{Unknown}(E_i, E_j) = \emptyset$, $P(E_i, E_j)$ is also the proportion of interpretations with an arc corresponding to (E_i, E_j) having a decoding from $\mathsf{Valid}(E_i, E_j)$ normalized by the total number of interpretations. Therefore, $P(E_i, E_j)$ is the proportion of mass assigned by the simple distribution to subgraphs of G that contain (E_i, E_j) normalized by the total number of interpretations. Therefore, when $\mathsf{Unknown}(E_i, E_j) = \emptyset$, P is the marginal distribution over $\mathcal{S}$ obtained from the simple distribution over $\{G' \mid G' \sqsubseteq G\}$.

(Case 2) Assume the postulate P3. We consider any $(E_i, E_j) \in \mathcal{S}$ where $\mathsf{Unknown}(E_i, E_j) \neq \emptyset$. This postulate implies that $P(E_i, E_j) = 0.5$. Also, by Proposition 23, $\mathsf{Unknown}(E_i, E_j) \neq \emptyset$ implies that $\mathsf{Potential}(E_i, E_j) = \mathsf{Unknown}(E_i, E_j)$. Therefore, there are no decodings of E_j in (E_i, E_j), i.e. $\mathsf{Decode}(E_j, \mathcal{K}) = \emptyset$. Therefore, for all candidates $G' \in \mathsf{Candidates}(G, \mathcal{K})$, $F(E_j)$ is a precursor. Therefore, $(E'_i, E'_j) \in \mathsf{Open}(G')$. Therefore, from Definition 6, half the interpretations G'' of G' will have the arc (E'_i, E'_j), and half the interpretations G'' of G' will not have the arc (E'_i, E'_j). Therefore, half of all the interpretations in $\mathsf{Interpretations}(G, \mathcal{K})$ will have the arc (E'_i, E'_j) and half of all the interpretations in $\mathsf{Interpretations}(G, \mathcal{K})$ will not have the arc (E'_i, E'_j). Since, $P(E_i, E_j) = 0.5$, this implies that there is a uniform distribution assigned to the interpretations, and that half the mass is assigned to subgraphs of G that contain (E_i, E_j) and half the mass assigned by the uniform distribution will be assigned to subgraphs of G that do not contain (E_i, E_j). Therefore, when $\mathsf{Unknown}(E_i, E_j) \neq \emptyset$, P is the marginal distribution over $\mathcal{S}$ obtained from the simple distribution over $\{G' \mid G' \sqsubseteq G\}$ □

Following from this, for each attack in the posit graph, if the attacked argument is a precursor that cannot be decoded, then the probability of attack will be 0.5, otherwise the decoding of the attacked argument is valid, so the probability of attack is given by the proportion of the decoded attacks that are logical attacks.

5.8 Discussion of modelling enthymemes

Enthymemes are a very common phenomenon in argumentation, and yet relatively little progress has been made in developing formalisms for modelling and analysing them.

In this section, we have focused on the problem of how discussion involving enthymemes can be difficult to analyse since it is uncertain whether attacks by one participant on the arguments of another participant are indeed valid attacks. This is a common problem when attempting to judge discussions, particularly when it is difficult to interact with the participants. Therefore, our approach could be incorporated into technology to analyse discussions. In the long term such technology should listen to the participants, understanding the meaning of individual arguments, and be able to draw conclusions from the discussion. In the shorter term, we could envisage that the starting graph can be constructed by the audience when analysing a discussion perhaps using a systematic methodology for identifying the premises, claims, and the attacks. Once the starting graph is obtained, it can be analysed to get the likely extensions, or to determine if specific attacks are likely to be valid. This may help the audience draw conclusions from the discussion. It would also appear that the ideas are sufficiently general to be adapted for use in key logic-based proposal for argumentation including ASPIC+ [Pra10] and ABA [DKT06].

Whilst, we have focussed on an audience judging participants in a discussion or debate, we plan to develop the approach so that a participant in a discussion or debate can attempt to judge whether its opponents have made incorrect attacks on its enthymemes, or whether it has made incorrect attacks on its opponents enthymemes. This could then take into account of how what one participant thinks is common knowledge could differ from what another participant thinks is common knowledge.

We also want to generalize our approach to consider enthymemes where the claim is not explicit. To explain this, we consider the following conversation between two members of a family prior to Christmas.

- John says “Tabby is a member of the family. Let’s buy her an xmas present”.
- Mary replies “She is a cat”.

So John has made an argument for buying an xmas present for Tabby, and Mary has replied with what may or may not be a counterargument. From the point of view of abstract argumentation, we have the following arguments which can be assembled into an argument graph where B attacks A.

- A = "Tabby is a member of the family. Let's buy her an xmas present."
- B = "She is a cat."

The enthymemetic nature of B means that we are not sure whether indeed B does attack A, and so at this point we can only denote the uncertainty by the variable x. In order to determine the value, we need to determine the possible interpretations of the argument. For this, we need to identify possible missing premises and claim. And for this, we have identified (in text form), four possible interpretations of the enthymeme B. Whilst each is textual (and still open to some further interpretation), it is clearer in each case whether or not the interpretation attacks argument A.

- B_1 = "She is a cat, I hate cats, and I don't buy presents for those that I hate, therefore we shouldn't buy her a present".
- B_2 = "She is a cat, and so she is not a member of the family".
- B_3 = "She is a cat, she doesn't have a concept of xmas, and so it's impossible to say that she'll like a xmas present".
- B_4 = "She is a cat, they love to be spoiled, and so she'll love a present."

Here we see that B_1 is a rebuttal to A (i.e. their claims contradict each other), B_2 and B_3 are undercuts to A, and B_4 is not a counterargument to A. So we see that if we have B_1, B_2 or B_3 as the representative for B, then the argument attacks A, whereas if we have B_4 as the representative for B, then the argument does not attack A. We envisage that we can generalise the logic-based framework presented in this section to handle this kind of example. It would then be interesting to see how this enhanced framework could be used in dialogs involving enthymemes as investigated in [Dup11b, Dup11a].

In [Hun07, BH12], proposals have been made for how common knowledge can be used to construct enthymemes from logical arguments (so that a proponent can send an enthymeme to an intended recipient by removing common knowledge) and deconstruct (i.e. decode) enthymemes (so that a recipient can rebuild the intended logical argument from the enthymeme by reintroducing the common knowledge). These proposal are based on logical argumentation, with preferences over the cobase, and with the deconstruction process being based on abduction. In future work, the role of abduction in the decoding process (e.g. [EGL97]) could be further investigated as well as how the cobase could be better structured to also take account of uncertainty that a formula is a premise in a particular context.

It is worth adding that the formal nature of enthymemes are not well understood. There are few proposals for formal handling of enthymemes, it is far from clear what are the best approaches. Clearly there is the need for further research to investigate and compare a range of possibilities drawing on a range of approaches in knowledge representation and reasoning field including for example possibilistic and fuzzy methods. Already there are interesting proposals for combining possibility theory and defeasible argumentation (e.g. [ACGS08]) that may offer insights into the handling enthymemes.

6 Comparisons with developments of abstract argumentation

There have been a number of proposals for extending Dung's framework in order to allow for more sophisticated modelling and analysis of conflicting information. A common theme among some of these proposals is the observation that not all attacks are equal.

One of the first proposals for capturing this idea was preference-based argumentation frameworks (PAF) [AC02]. This generalizes Dung's definition for an argument graph by introducing a preference relation over arguments that in effect causes an attack to be ignored when the attacked argument is preferred over the attacker. So in PAF, we assume a preference relation over arguments, denoted $\mathcal{P}$, as well as a set of arguments $\mathcal{A}$ and an attack relation $\mathcal{R}$. From this, we need to define a defeat relation $\mathcal{D}$ as follows, and then, $(\mathcal{A}, \mathcal{D})$ is used as the argument graph, instead of $(\mathcal{A}, \mathcal{R})$, with Dung's usual definitions for extensions.

$$\mathcal{D} = \{(\alpha, \beta) \in \mathcal{R} \mid (\beta, \alpha) \notin \mathcal{P}\}$$

So with this definition for defeat, extensions for a preference-based argument graph $(\mathcal{A}, \mathcal{R}, \mathcal{P})$ can be obtained as follows: For $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $\Gamma \subseteq \mathcal{A}$, Γ is an extension of $(\mathcal{A}, \mathcal{R}, \mathcal{P})$ w.r.t. X iff Γ is an extension of $(\mathcal{A}, \mathcal{D})$ w.r.t. X.

The following result shows that we get the same extensions with PAF and probabilistic attack graphs when we set the probability assignment to be 1 for each attack in $\mathcal{D}$ and we set the probability assignment to be 0 for each attack not in $\mathcal{D}$.

Proposition 25. *Let $(\mathcal{A}, \mathcal{R}, \mathcal{P})$ be a PAF, and let $(\mathcal{A}, \mathcal{R}, P)$ be a probabilistic attack graph where the probability of attack P is defined as follows.*

$$P(\alpha, \beta) = \begin{cases} 1 \text{ if } (\alpha, \beta) \in \mathcal{D} \\ 0 \text{ if } (\alpha, \beta) \notin \mathcal{D} \end{cases}$$

For $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $\Gamma \subseteq \mathcal{A}$, Γ is an extension of $(\mathcal{A}, \mathcal{R}, \mathcal{P})$ w.r.t. X iff $P_X(\Gamma) = 1$.

Proof. Let $G = (\mathcal{A}, \mathcal{R}, P)$ and let $G' = (\mathcal{A}, \mathcal{R}', P)$ where $\mathcal{R}'$ is $\mathcal{D}$. Assume P is defined as above. Here we get the following probability distribution over subgraphs.

$$P(G') = (\prod_{(\alpha,\beta)\in\mathcal{D}} P(\alpha, \beta)) \times (\prod_{(\alpha,\beta)\in\mathcal{R}\setminus\mathcal{D}} (1 - P(\alpha, \beta)))$$

Since, for each $(\alpha, \beta) \in \mathcal{D}$, $P(\alpha, \beta) = 1$, and for each $(\alpha, \beta) \in \mathcal{R} \setminus \mathcal{D}$, $(1 - P(\alpha, \beta) = 1)$, we have $P(G') = 1 \times ... \times 1 = 1$. ($\Rightarrow$) Assume for some $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $\Gamma \subseteq \mathcal{A}$, Γ is an extension of $(\mathcal{A}, \mathcal{R}, \mathcal{P})$ w.r.t. X. Therefore, by definition, $G' \in Q_X(\Gamma)$. Since, we have shown that $P(G') = 1$, we have $\sum_{G'\in Q_X(\Gamma)} P_{\mathsf{ad}}(\Gamma') = 1$. Therefore, by definition, $P_{\mathsf{ad}}(\Gamma) = 1$. ($\Leftarrow$) Assume, for some $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $P_X(\Gamma) = 1$. Therefore, by definition, $\sum_{G'\in Q_X(\Gamma)} P_{\mathsf{ad}}(\Gamma') = 1$. Hence, $G' \in Q_X(\Gamma)$, since $P(G') = 1$. Therefore, by definition, $\Gamma \subseteq \mathcal{A}$, Γ is an extension of $(\mathcal{A}, \mathcal{R}, \mathcal{P})$ w.r.t. X. $\square$

The idea of explicitly adding weights to attacks was first proposed in [BGW05], and we can see a probability assignment on attacks as being a special case of that idea. However, the emphasis of that work was on how weights can be changed dynamically and over time, rather than on extending Dung's framework for determining extensions, or for harnessing probability theory in argumentation.

A more general approach, proposed in [DHM$^+$09, DHM$^+$11], extends argument graphs by associating a weight with each attack, where the weight indicates the relative strength of the attack. A key concept in the proposal is the notion of an inconsistency budget, which characterizes how much inconsistency we are prepared to tolerate: given an inconsistency budget, we would be prepared to disregard attacks up to a total weight of the budget. The key advantage of this approach is that it permits a much finer-grained level of analysis of argument systems than unweighted systems, and gives useful solutions when conventional (unweighted) argument systems have none. Furthermore, it is shown in [DHM$^+$09, DHM$^+$11] that the approach subsumes other approaches such as preference-based argumentation frameworks [AC02], value-based argumentation frameworks [Ben03], resolution-based argumentation frameworks [BG08], and extended argumentation frameworks [Mod09].

A weighted argument graph is a tuple $(\mathcal{A}, \mathcal{R}, w)$ where $w : \mathcal{R} \to \mathbb{R}^+$ assigns a weight to each attack and $\mathbb{R}^+$ denotes the set of reals greater than or equal to zero. In order to harness this definition, we require the following subsidiary definitions. Given a set of arcs and a weight function, the WeightSum function gives the sum of weights associated with the arcs in the set.

$$\mathsf{WeightSum}(\mathcal{R}, w) = \sum_{(\alpha,\beta)\in\mathcal{R}} w((\alpha, \beta))$$

Then for a weighted argument graph, the TolerantGraphs function gives the set of sets of arcs each of which is obtained by removing from the set of arcs $\mathcal{R}$ a subset that have a weighted sum less than or equal to a given budget $b \in \mathbb{R}^+$. So the TolerantGraphs function gives the arcs for spanning subgraphs of the original graph that have some arcs removed and the sum of those removed arcs is less than or equal to the budget b.

$$\mathsf{TolerantGraphs}(\mathcal{R}, w, b) = \{\mathcal{R}' \subseteq \mathcal{R} \mid \mathsf{WeightSum}(\mathcal{R} \setminus \mathcal{R}', w) \leq b\}$$

Now we can define an extension for a semantics $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$ in a weighted graph $(\mathcal{A}, \mathcal{R}, w)$ with budget b as follows.

Γ is an extension of $(\mathcal{A}, \mathcal{R}, w)$ w.r.t. semantics X and budget b
iff there is a $\mathcal{R}' \in \mathsf{TolerantGraphs}(\mathcal{R}, w, b)$ and Γ is an extension of $(\mathcal{A}, \mathcal{R}')$.

So a set of arguments is an extension if and only if there exists a tolerant graph (i.e. graph with some edges removed and the sum of those removed edges is less than or equal to the budget b) for which Γ is an extension. This generalization means that semantics that have a unique extension in Dung's original definition, such as grounded semantics, may have multiple extensions in this version.

Proposition 26. *Let $(\mathcal{A}, \mathcal{R}, w)$ be a weighted argument graph, and let $(\mathcal{A}, \mathcal{R}, P)$ be a probabilistic attack graph where P be a probability distribution over subgraphs $G' \sqsubseteq G$ defined as follows. For this, let* $\mathsf{Arcs}(G')$ *be the set of arcs in the argument graph, and π be the number of tolerant graphs (i.e. the cardinality of* $\mathsf{TolerantGraphs}(\mathcal{R}, w, b)$*).*

$$P(G') = \begin{cases} 1/\pi \text{ if } \mathsf{Arcs}(G') \in \mathsf{TolerantGraphs}(\mathcal{R}, w, b) \\ 0 \text{ if } \mathsf{Arcs}(G') \notin \mathsf{TolerantGraphs}(\mathcal{R}, w, b) \end{cases}$$

For $X \in \{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$, $b \in \mathbb{R}^+$, $\Gamma \subseteq \mathcal{A}$, Γ is an extension of $(\mathcal{A}, \mathcal{R}, w)$ w.r.t. semantics X and budget b iff $P_X(\Gamma) > 0$.

Proof. Assume $(\mathcal{A}, \mathcal{R}, w)$ is a weighted argument graph, $(\mathcal{A}, \mathcal{R}, P)$ is a probabilistic attack graph where P be a probability distribution over subgraphs defined as above, X is a semantics in $\{\mathsf{co}, \mathsf{pr}, \mathsf{gr}, \mathsf{st}\}$ and b is a budget in $\mathbb{R}^+$.

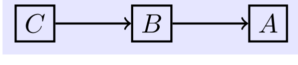
$\Gamma \subseteq \mathcal{A}$, Γ is an extension of $(\mathcal{A}, \mathcal{R}, w)$ w.r.t. X and b
($\Leftrightarrow$) there is a $\mathcal{R}' \in \mathsf{TolerantGraphs}(\mathcal{R}, w, b)$ and Γ is an extension of $(\mathcal{A}, \mathcal{R}')$
($\Leftrightarrow$) there is a $G' \sqsubseteq G$ such that $\mathsf{Arcs}(G') \in \mathsf{TolerantGraphs}(\mathcal{R}, w, b)$ and Γ is an extension of $(\mathcal{A}, \mathcal{R}')$
($\Leftrightarrow$) there is a $G' \sqsubseteq G$ such that $\mathsf{Arcs}(G') \in \mathsf{TolerantGraphs}(\mathcal{R}, w, b)$ and $G' \in Q_X(\Gamma)$.
($\Leftrightarrow$) there is a $G' \sqsubseteq G$ such that $P(G') > 0$ and $G' \in Q_X(\Gamma)$.
($\Leftrightarrow$) $\sum_{G' \in Q_X(\Gamma)} P(G') > 0$.
($\Leftrightarrow$) $P_X(\Gamma) > 0$.

□

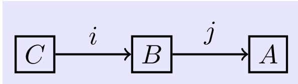
Therefore, our approach of probabilistic attack graphs subsumes weighted argument graphs (WAG), and therefore by using the results in [DHM+11], we also get that the approach of probabilistic attack graphs subsumes the approaches of preference-based argumentation frameworks (PAF) [AC02], value-based argumentation frameworks (VAF) [Ben03], resolution-based argumentation frameworks (RAF) [BG08], and extended argumentation frameworks (EAF) [Mod09]. Each of these proposals involves knocking out attacks, and this can be captured by defining the probability distribution over $\{G' \sqsubseteq G\}$ appropriately.

It is straightforward to see that the proposal in this paper is not subsumed by the WAG, PAF, VAF, RAF, or EAF approaches, since the approach of probabilistic attack graphs gives a quantitative evaluation of each set of arguments being an admissible set or an extension, as well as a quantitative evaluation of each argument being an inference, whereas these other approaches do not provide a quantitative evaluation, rather they just provide a Boolean evaluation on whether a set of arguments is an admissible set or an extension. Therefore these other approaches cannot give the richer qualifications to extensions that can be obtained by our probabilistic approach.

In another proposal for developing Dung's proposal, which we will refer to as varied-strength attacks (or VSA) approach, extra information representing the relative strength of attack is incorporated [MGS08]. In the VSA approach, each arc is assigned a type, and there is a partial ordering over the types. As a simple example, consider the following argument graph.



Here, C defends the attack on A, and as a result $\{C, A\}$ is the preferred, grounded and complete extension. Now consider the following version of the graph, according to the VAS approach, where the attack by C is of type i and the attack by B is of type j.



This gives a finer grained range of defence depending on whether type j is higher, or lower, or equally, ranked than i, or incomparable with i. Furthermore, this allows for a finer-grained definition of acceptable extension that specifies the required level of the defence of any argument in the extension. For instance, it can be insisted in the VSA approach that every defence of an argument should be by an attack that is stronger, so in the above graph that would mean that the type of $\rightarrow_i$ needs to be stronger than the type of $\rightarrow_j$ in order for $\{C, A\}$ to be an admissible set.

So in the VSA approach, a variable strength attack graph is a tuple $(\mathcal{A}, \mathcal{R}, \mathcal{S})$ where $\mathcal{S} \subseteq \mathcal{R} \times \mathcal{R}$ is a partial order relation over attacks according to relative strength, so if $((\alpha, \beta), (\gamma, \delta)) \in \mathcal{S}$, then the attack (α, β) is of higher or equal strength to (γ, δ). In order to simplify the presentation, we will restrict consideration to strict preference, though this can be generalised to a partial order. So we assume that for all $((\alpha, \beta), (\gamma, \delta)) \in \mathcal{S}$, it is not the case that $((\gamma, \delta), (\beta, \alpha)) \in \mathcal{S}$. Using this, the notion of defence is defined as follows, where $\Gamma \subseteq \mathcal{A}$ is a conflictfree set of arguments.

Γ defends α with respect to $\mathcal{S}$
iff if there is a $(\beta, \alpha) \in \mathcal{R}$,
then there is a $\gamma \in \Gamma$ such that $(\gamma, \beta) \in \mathcal{R}$, and $((\beta, \alpha), (\gamma, \beta)) \notin \mathcal{S}$.

Using the above definition for defence, the notion of admissibility for the VSA approach is defined as follows, where $\Gamma \subseteq \mathcal{A}$ is a conflictfree set of arguments.

Γ is an admissible set with respect to $\mathcal{S}$ iff for each $\alpha \in \Gamma, \Gamma$ defends α with respect to $\mathcal{S}$

As with the other approaches, the VSA approach does not subsume the probabilistic attack graphs approach, since the approach of probabilistic attack graphs gives a quantitative evaluation of each set of arguments being an admissible set, whereas the VSA approach only provides a qualitative evaluation, (i.e. a set of arguments is an admissible set or not), given a preference relation $\mathcal{S}$.

However, the VSA approach is not subsumed by the probabilistic attack graphs approach. In the above subsumption results, we were able to take the extra information that comes with the variant (e.g. preferences over arguments, weights on attacks with a budget, etc), to select one or more subgraphs with an appropriate probability assignment. This allowed the selection of extensions by the original approach to be captured in a simple way. In contrast, in the VSA approach, the way the extra information about the relative strength of attacks is used depends upon the current state of arguments being considered. Furthermore, the VSA is concerned with reinstatement rather than the initial attack. So an argument is only reinstated when the defender has an attack that is preferred to that of the original attacker. This means that the weights are used not to determine which attacks to use and which not to use, but rather to determine which arguments to reinstate given a potential defender. So the VSA approach offers a substantially different view on using weights.

7 Discussion and future work

The idea to assign a probability value to attacks and use this to obtain a probability distribution over subgraphs was first presented by Li *et al* [LON11]. Unfortunately, that proposal used an independence assumption between attacks. Whilst the independence assumption is useful and appropriate in some situations, it is not always appropriate (as we discussed in Section 3.2).

To address this weakness, we have considered a probability distribution over spanning subgraphs of the argument graph in this paper. This has resulted in the following contributions: (1) A discussion of a range of situations where the approach offers interesting opportunities for capturing real-world argumentation situations where there is a need to handle phenomena such as explicit qualification of attacks, implicit imprecision in attacks, and enthymemes with incomplete premises and/or incomplete claims. (2) A number of theoretical results concerning the use of a probability distribution over subgraphs to clarify the nature of the proposal and to identify benefits of the approach, including the idea of probability functions over inferences; (3) An investigation of modelling the uncertainty arising from decoding enthymemes; and (4) A comparison with other approaches including preference-based argumentation frameworks [AC02] and weighted argument graphs [DHM^{+}09, DHM^{+}11] showing that probabilistic attack graphs subsumes these, and thereby also subsuming value-based argumentation frameworks [Ben03], resolution-based argumentation frameworks [BG08], and extended argumentation frameworks [Mod09], and in addition, an explanation of how the approach is not subsumed by, nor subsumes, varied-strength attacks [MGS08];

Dung and Thang [DT10] have also extended abstract argumentation with probability theory. However, they do so by introducing a probability distribution over sets of arguments. This is used to give a probability distribution over subgraphs of an argument graph, and then the probability of an argument A is the sum of the probabilities for the subgraphs that contain A in their grounded extensions. This work is complementary to the proposal in this paper since their probability assignment is to arguments. It appears that for some modelling problems, assignment to arguments is more appropriate, and for other modelling problems, assignment to attacks is more appropriate.

Another proposal for making a probability assignment to arguments has been made by Thimm [Thi12] where given an argument graph, a probability distribution meeting some intuitive constraints is sought. A number of interesting results are shown such as the relationship between minimum and maximum entropy distributions and types of extension. Again this is complementary to our approach since it does not consider qualification of attacks.

There are relatively few other attempts to bring probability theory into computational models of argument. In the ABEL framework, reasoning with propositional information is augmented with probabilistic information so that individual arguments are qualified by a probability value [Hae98, HKL00, Hae01, Hae05]. This is a logic-based formalism rather than an abstract formalism, and there is no consideration of how this probabilistic information relates to Dung's proposals, or how it could be used to decide which arguments are acceptable. Another logic-based framework for argumentation, the LA system, also introduces probabilities into the rules, and these probabilities are propagated by the inference rules so that arguments are qualified by probabilities (such as via labels such as "likely", "very likely", etc), but again there is no consideration of how this probabilistic information relates to Dung's proposals [EGKF93, FD00]. In another rule-based system for argumentation, Riveret *et al* [RRS^{+}07], the belief in the premises of an argument is used to calculate the belief in the argument. However, as with the LA system, there is no consideration of how this probabilistic information relates to Dung's proposals for abstract argumentation.

Probabilistic reasoning with logical statements has also been considered by Pollock [Pol95]. However, the approach taken is to assign probabilities to formulae without considering the meaning of this in terms of models. Various issues arising from an assignment based on frequency that a consequent holds when the antecedent holds are considered, as well as how such an assignment could be used for statistical syllogism. The emphasis of the work is therefore different as it does not consider how a probabilistic perspective relates to abstract argumentation.

Probabilities have also been proposed in the qualitative probabilistic reasoning (QPR) framework, which is a qualitative form of Bayesian network, and whilst this can be regarded as a form of argumentation, it does not incorporate the dialectical reasoning seen with the generation of arguments and counterarguments, and so again there is no consideration of how this probabilistic information relates to Dung's proposals [Par96, Par98, Par01]. In another approach based on Bayesian networks, defeasible reasoning is simulated with such a network, so that conditional probabilities represent defeasible rules, and the probabilities are used to decide which inference is propagated. By harnessing existing software for Bayesian networks, it has been shown to be a computationally viable approach, though again there is no consideration of how this probabilistic information relates to Dung's proposals [Vre05]. Argumentation has also been used for merging multiple Bayesian networks, but then the Bayesian networks are the subject of argumentation, rather than probabilistic information being used to quantify the uncertainty of arguments in general [NP06].

The proposal for weighted argument graphs has been extended in various ways, such as merging weighted argument graphs [CLS11], and defining notions of acceptability based on weights [CDLS10, CMKM12]. It would be interesting in future work to investigate how these developments could be harnessed for probabilistic attack graphs, thereby leading to further insights into uncertainty in argumentation.

In an argumentation system based on classical logic, the formalism has been extended with a probability distribution over the models of the language [Hun13]. This allows for the uncertainty in the premises to be quantified, thereby giving a representation of the uncertainty in the arguments. This then raises the question of how the uncertainty in the logical arguments affects the uncertainty in the attacks. It would also be interesting to compare this with an interesting proposal for combining logical and probabilisitic perspectives in argumentation by Verheij [Ver12] which draws on axiomatizations of non-monotonic reasoning in what is called "ampliative reasoning".

Another topic for research is the development of algorithms and implementation for using the framework presented in this paper. A number of proposals have been made for algorithms for abstract argumentation (e.g. [BG02, CDM03, DMT07]). A good starting point would be the ASP-based ASPARTIX system [EGW08] as this would appear to support development of algorithms as ASP programs, which in turn would appear appropriate for also reasoning with the subgraphs. For instance, it would appear straightforward to implement an ASP program that considers each subgraph in turn, and then calls ASPARTIX for each subgraph to determine the extensions and inferences for the subgraph. Li *et al* give a Monte Carlo algorithm for querying probabilistic argument graphs, together with an empirical evaluation [LON11], and so any algorithm for determining the probability of admissible sets, extensions, or inferences that could be developed on ASPARTIX should be compared empirically with the Monte Carlo algorithm.

Finally, it would be valuable to compare and contrast the use of probability theory with the use of alternative uncertainty formalisms, in particular possibility theory, for qualifying arguments and attacks. There are proposals for possibilistic forms of logical argumentation (e.g. [AP04, ACGS08]). It would be interesting to draw on these proposals to develop an alternative justification and formalisation for representing and reasoning with uncertainty in abstract argumentation. Some of the concepts in probabilistic abstract argumentation proposals (such as considered in [DT10, LON11, Thi12, Hun12] as well as in this paper) would translate directly to a possibilistic formalization, but there may be some valuable differences for instance in modelling ambiguity and vagueness.

Acknowledgements

The author is very grateful to V.S. Subrahmanian for a valuable discussion that led to this paper and to the anonymous reviewers who provided numerous suggestions for improving this paper.

References

[AC02] L. Amgoud and C. Cayrol. A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, 34:197–215, 2002.

[ACGS08] T. Alsinet, C. Chesñevar, L. Godo, and G. Simari. A logic programming framework for possibilistic argumentation: Formalization and logical properties. *Fuzzy Sets and Systems*, 159(10):1208–1228, 2008.

[AP04] L. Amgoud and H. Prade. Reaching agreement through argumentation: A possibilistic approach. In *In the Proceedings of the Ninth Conference on Principles of Knowledge Representation and Reasoning (KR'04)*, pages 175–182. AAAI Press, 2004.

[BCD07] T. Bench-Capon and P. Dunne. Argumentation in artificial intelligence. *Artificial Intelligence*, 171(10-15):619–641, 2007.

[Ben03] T. Bench-Capon. Persuasion in practical argument using value based argumentationframeworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[BG02] P. Baroni and M. Giacomin. Argumentation through a distributed self-stabilizing approach. *Journal of Experimental and Theoretical Artificial Intelligence*, 14(4):273–301, 2002.

[BG08] P. Baroni and M. Giacomin. Resolution-based argumentation semantics. In *Computational Models of argument (COMMA'08)*, pages 25–36. IOS Press, 2008.

[BGW05] H. Barringer, D. Gabbay, and J. Woods. Temporal dynamics of support and attack networks: From argumentation to zoology. In *Mechanizing Mathematical Reasoning*, volume 2605 of *Lecture Notes in Computer Science*, pages 59–98. Springer, 2005.

[BH01] Ph. Besnard and A. Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128:203–235, 2001.

[BH08] Ph. Besnard and A. Hunter. *Elements of Argumentation*. MIT Press, 2008.

[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.

[Cam06] M. Caminada. Semi-stable semantics. In *Computational Models of Argument (COMMA'06)*, pages 121–130, 2006.

[CDLS10] C. Cayrol, C. Devred, and M-C Lagasquie-Schiex. Acceptability semantics accounting for strength of attacks. In *Proceedings of the European Conference on Artificial Intelligence*, 2010.

[CDM03] C. Cayrol, S. Doutre, and J. Mengin. On the decision problems related to the preferred semantics for argumentation frameworks. *Journal of Logic and Computation*, 13(3):377–403, 2003.

[CLS11] C. Cayrol and M-C Lagasquie-Schiex. A tool for merging argumentation systems. In *Proceedings of the International Conference on Tools with Artificial Intelligence*, pages 629–632. IEEE, 2011.

[CMKM12] S. Coste-Marquis, Sebastien Konieczny, and P. Marquis. Selecting extensions in weighted argumentation frameworks. In *Computational Models of Argument (COMMA'12)*, pages 342–349. IOS Press, 2012.

[CML00] C. Chesñevar, A. Maguitman, and R. Loui. Logical models of argument. *ACM Computing Surveys*, 32:337–383, 2000.

[DHM$^+$09] P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Inconsistency tolerance in weighted argument systems. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, pages 851–858. IFAAMAS, 2009.

[DHM$^+$11] P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175(2):457–486, 2011.

[DKT06] P. Dung, R. Kowalski, and F. Toni. Dialectical proof procedures for assumption-based admissible argumentation. *Artificial Intelligence*, 170:114–159, 2006.

[DMT07] P. Dung, P. Mancarella, and F. Toni. Computing ideal skeptical argumentation. *Artificial Intelligence*, 171:642–674, 2007.

[DT10] P. Dung and P. Thang. Towards (probabilistic) argumentation for jury-based dispute resolution. In *Computational Models of Argument (COMMA'10)*, pages 171–182. IOS Press, 2010.

[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.

[Dup11a] F. Dupin de Saint Cyr - Bannay. A first attempt to allow enthymemes in persuasion dialogs. In *Proceedings of the DEXA International Workshop: Data, Logic and Inconsistency (DALI'11)*, pages 332–336. IEEE Computer Society, 2011.

[Dup11b] F. Dupin de Saint Cyr - Bannay. Handling enthymemes in time-limited persuasion dialogs. In *Scalable Uncertainty Management (SUM'11)*, number 6929 in LNAI, pages 149–162. Springer, 2011.

[EGKF93] M. Elvang-Gøransson, P. Krause, and J. Fox. Acceptability of arguments as 'logical uncertainty'. In *Proceedings of the European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'93)*, pages 85–90. Springer-Verlag, 1993.

[EGL97] T. Eiter, G. Gottlob, and N. Leone. Semantics and complexity of abduction from default theories. *Artificial Intelligence*, 90:177–223, 1997.

[EGW08] U. Egly, S. Gaggl, and S. Woltran. Aspartix: Implementing argumentation frameworks using answer-set programming. In *Proceedings of the Twenty-Fourth International Conference on Logic Programming (ICLP'08),*, volume 5366 of *LNCS*, pages 734–738. Springer, 2008.

[FD00] J. Fox and S. Das. *Safe and Sound: Artificial Intelligence in Hazardous Applications*. MIT Press, 2000.

[Hae98] R. Haenni. Modelling uncertainty with propositional assumptions-based systems. In *Applications of Uncertainty Formalisms*, pages 446–470. Springer, 1998.

[Hae01] R. Haenni. Cost-bounded argumentation. *International Journal of Approximate Reasoning*, 26(2):101–127, 2001.

[Hae05] R. Haenni. Using probabilistic argumentation for key validation in public-key cryptography. *Internation Journal of Approximate Reasoning*, 38(3):355–373, 2005.

[HKL00] R. Haenni, J. Kohlas, and N. Lehmann. Probabilistic argumentation systems. In D. Gabbay and Ph Smets, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Volume 5*, pages 221–288. Kluwer, 2000.

[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 abstract argumentation. In *Computational Mdoels of Argument (COMMA'12)*, pages 117–128, 2012.

[Hun13] A. Hunter. A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning*, 54(1):47–81, 2013.

[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 (TAFA'11)*, 2011.

[MGS08] D. Martinez, A. Garcia, and G. Simari. An abstract argumentation framework with varied-strength attacks. In *Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR'08)*, 2008.

[Mod09] S. Modgil. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, 173(9-10):901–934, 2009.

[NP06] S. Neilson and S. Parsons. An application of formal argumentation: Fusing Bayes nets in MAS. In *Computational Models of Argument (COMMA'06)*, pages 33–46. IOS Press, 2006.

[Par96] S. Parsons. Defining normative systems for qualitative argumentation. In *Proceedings of the International Conference on Formal and Applied Practical Reasoning*, volume 1085 of *LNCS*, pages 449–463, 1996.

[Par98] S. Parsons. A proof theoretic approach to qualitative probabilistic reasoning. *International Journal of Approximate Reasoning*, 19:265–297, 1998.

[Par01] S. Parsons. *Qualitative Approaches to Reasoning Under Uncertainty*. MIT Press, 2001.

[Pol95] J. Pollock. *Cognitive Carpentry*. MIT Press, 1995.

[Pra10] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1:93–124, 2010.

[PV02] H. Prakken and G. Vreeswijk. Logical systems for defeasible argumentation. In D. Gabbay, editor, *Handbook of Philosophical Logic*, pages 219–318. Kluwer, 2002.

[RRS+07] R. Riveret, A. Rotolo, G. Sartor, H. Prakken, and B. Roth. Success chances in argument games: a probabilistic approach to legal disputes. In *Legal Knowledge and Information Systems (JURIX'07)*, pages 99–108. IOS Press, 2007.

[RS09] I. Rahwan and G. Simari, editors. *Argumentation in Artifical Intelligence*. Springer, 2009.

[Thi12] M. Thimm. A probabilistic semantics for abstract argumentation. In *Proceedings of the European Conference on Artificial Intelligence (ECAI'12)*, pages 750–755, 2012.

[Ver12] B. Verheij. Jumping to conclusions: A logic-probabilistic foundation for defeasible rule-based arguments. In *Logics in Artificial Intelligence (JELIA'12)*, volume 7519 of *LNCS*, pages 411–423. Springer, 2012.

[Vre05] G. Vreeswijk. Argumentation in Bayesian belief networks. In *Argumentation in Multi-agent Systems (ArgMAS'04)*, volume 3366 of *LNAI*. Springer, 2005.

[Wal89] D. Walton. *Informal Logic: A Handbook for Critical Argumentation*. Cambridge University Press, 1989.