24th European Conference on Artificial Intelligence - ECAI 2020
Santiago de Compostela, Spain

Abstract Argumentation with Markov Networks

Nico Potyka [1]

Abstract. We explain how abstract argumentation problems can be encoded as Markov networks. From a computational perspective, this allows reducing argumentation tasks like finding labellings or deciding credulous and sceptical acceptance to probabilistic inference tasks in Markov networks. From a semantical perspective, the resulting probabilistic argumentation models are interesting in their own right. In particular, they satisfy several of the properties proposed for epistemic probabilistic argumentation by Hunter and Thimm. We also consider an extension to frameworks with deductive support and show that it maintains many of the interesting guarantees of both approaches.

1 Introduction

Abstract argumentation [13] studies the acceptability of arguments based on their relationships and abstracted from their content. While abstract argumentation frameworks initially considered only attack relationships between arguments, bipolar argumentation allows an additional support relation [4, 29, 8, 11]. It is also often useful to go beyond the classical two-valued view that arguments can only be accepted or rejected. More fine-grained semantics are considered, for example, in ranking frameworks that can be based on fixed point equations [7, 27, 5, 12] or the graph structure [10, 2] and gradual argumentation frameworks [33, 3, 31]. Probabilistic argumentation captures more fine-grained degrees of belief by building up on probability theory [14, 21, 35, 36, 40, 30, 32].

Computational argumentation deals with the question how and how efficient argumentation problems can be solved algorithmically [15]. The development of algorithms has been driven forward by the biannual International Competition on Computational Models of Argumentation [41, 19]. Many of the most successful competitors apply SAT-, ASP- or CSP-solvers. These solvers are well suited for the structure of many argumentation problems and are hard to beat due to decades of research in these areas. However, when it comes to counting solutions with a particular structure, algorithms developed in the area of probabilistic reasoning may be an interesting alternative. In particular, the probabilistic graphical models community provides a variety of exact and approximate algorithms that offer a tradeoff between solution quality and computational performance [25]. These algorithms may be helpful in combinatorial tasks like counting labellings or performing sceptical inference.

In this paper, we study Markov networks as a tool for abstract argumentation. We explain how argumentation problems can be encoded as Markov networks and how inference tasks in argumentation frameworks can be reduced to inference tasks in Markov networks. As it turns out, the resulting Markov networks are interesting probabilistic argumentation models in their own right and the probabilities of arguments can sometimes be directly connected to the relative frequency of labellings that accept them. In particular, they satisfy several of the properties proposed by Hunter and Thimm for probabilistic argumentation [21]. We also study an extension to bipolar argumentation frameworks that respects supported and mediated attacks [11] and some properties by Hunter and Thimm that can be transferred to the bipolar setting naturally.

[1] University of Osnabrueck, Germany, email: npotyka@uni-osnabrueck.de

2 Markov Networks Basics

Intuitively, Markov networks give a graphical representation of the independency assumptions in a probabilistic model P. Theoretically, the graphical representation is negligible because all information is implicitly contained in P. The crucial assumption is that P can be decomposed into a product of factors such that each factor depends only on a small subset of the random variables. This special structure results in independencies between the random variables in P that can be exploited in order to compute probabilities more efficiently [25].

As usual in this area, we denote random variables by capital letters X, Y, Z and values of these random variables by small letters x, y, z. Bold capital letters $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ denote ordered sequences of random variables and bold small letters $\mathbf{x}, \mathbf{y}, \mathbf{z}$ denote assignments to these random variables. For example, if $\mathbf{X} = (X_1, X_2, X_3)$ and $\mathbf{x} = (x_1, x_2, x_3)$, then $\mathbf{X} = \mathbf{x}$ denotes the assignment $(X_1 = x_1, X_2 = x_2, X_3 = x_3)$. We write $\mathbf{Y} \subseteq \mathbf{X}$ if the random variables in $\mathbf{Y}$ form a subset of the random variables in $\mathbf{X}$. If $\mathbf{Y} \subseteq \mathbf{X}$, we denote by $\mathbf{X}|_{\mathbf{Y}}$ and $\mathbf{x}|_{\mathbf{Y}}$ the restriction of $\mathbf{X}$ and $\mathbf{x}$ to the random variables in $\mathbf{Y}$. For example, if $\mathbf{Y} = (X_1, X_3)$, we have $\mathbf{X}|_{\mathbf{Y}} = (X_1, X_3)$ and $\mathbf{x}|_{\mathbf{Y}} = (x_1, x_3)$ in our example above.

A *factor* with scope $\mathbf{X}$ is a function $\phi(\mathbf{X})$ that maps every assignment $\mathbf{x}$ to $\mathbf{X}$ to a non-negative real number. As a simple example, consider the factor $\phi(X_1, X_2) = X_1 \cdot X_2$. If X_1 and X_2 are boolean random variables, the factor yields $1 \cdot 1 = 1$ if $X_1 = X_2 = 1$ and 0 otherwise. It can therefore be interpreted as an encoding of logical conjunction. The intuition of factors is that they can increase or decrease the probabilities of states. Our example factor will move the probability of a state (variable assignment) where $X_1 \wedge X_2$ is not satisfied to 0 and leave the probability unchanged if it is satisfied.

Given a set of factors $\Phi = \{\phi_1(\mathbf{X}_1), \dots, \phi_k(\mathbf{X}_k)\}$, $\mathbf{X}_i \subseteq \mathbf{X}$, we define the *plausibility of a state* of $\mathbf{X}$ via

$$\mathrm{Pl}_\Phi(\mathbf{X}) = \prod_{i=1}^{k} \phi_i(\mathbf{X}|_{\mathbf{X}_i}).$$

By normalizing the plausibility, we obtain a probability distribution that is called the *Gibbs distribution* over $\mathbf{X}$ and is defined by

$$P_\Phi(\mathbf{X}) = \frac{1}{Z}\mathrm{Pl}_\Phi(\mathbf{X}),$$

where $Z = \sum_{\mathbf{x}} \mathrm{Pl}_\Phi(\mathbf{x})$ is a normalization constant that guarantees that the probabilities sum up to 1. In the literature, Z is usually called the *partition function*. The *Markov network structure* corresponding to P_Φ is the undirected graph obtained by introducing a node for every random variable and connecting each two nodes that occur together in a factor [25]. A *Markov network* is a Gibbs distribution $P_\Phi(\mathbf{X})$ along with the corresponding Markov network structure. However, in this work, the Gibbs distribution P_Φ is most important and we will not discuss the the graphical structure much.

One of the main reasoning problems in Markov networks is computing marginal probabilities. That is, given an assignment $\mathbf{Y} = \mathbf{y}$ to a subset $\mathbf{Y} \subset \mathbf{X}$ of the random variables, compute $P_\Phi(\mathbf{y})$. Note that this also allows computing conditional probabilities because conditional probabilities can be defined based on marginal probabilities via $P_\Phi(\mathbf{Y} \mid \mathbf{Z}) = \frac{P_\Phi(\mathbf{Y},\mathbf{Z})}{P_\Phi(\mathbf{Z})} = \frac{\mathrm{Pl}_\Phi(\mathbf{Y},\mathbf{Z})}{\mathrm{Pl}_\Phi(\mathbf{Z})}$. In general, computing marginal probabilities is intractable under the usual complexity-theoretical assumptions. However, the probabilistic graphical models community offers a considerable amount of work about tractable special cases and approximation algorithms [25]. The main family of exact algorithms is based on the idea of building up a clique tree from the Markov network and performing a message passing algorithm on the clique tree. Roughly speaking, probabilities can be computed in polynomial time if the size of cliques in the tree (not just in the original graph) can be bounded by a constant. In general, finding a small tree can be impossible or computationally intractable, but for some special cases, this can be done efficiently. Two other well investigated computational problems are computing the partition function Z and *MAP queries*, that is, finding an assignment to (a subset of) the random variables with maximum probability. Again, both problems are intractable in general, but can be solved exactly in special cases or can be approximated by a variety of algorithms [25].

3 Markov Networks for Abstract Argumentation

A Dung-style *(finite) abstract argumentation framework (AAF)* is a tuple $(\mathcal{A}, \mathrm{Att})$, where $\mathcal{A}$ is a finite set of arguments and $\mathrm{Att} \subseteq \mathcal{A} \times \mathcal{A}$ is the *attack relation* [13]. If $(A, B) \in \mathrm{Att}$, we say that A attacks B. With a slight abuse of notation, we let $\mathrm{Att}(A) = \{B \mid (B, A) \in \mathrm{Att}\}$ denote the set of attackers of A. Semantics of argumentation frameworks can be defined in terms of extensions or labellings in an equivalent way [9]. We will use labellings since they are more convenient for our purpose. A labelling is a function $\mathrm{L} : \mathcal{A} \to \{\mathrm{in}, \mathrm{out}, \mathrm{und}\}$ that assigns to each argument a label. With a slight abuse of notation, we let, for each label $l \in \{\mathrm{in}, \mathrm{out}, \mathrm{und}\}$, $l(\mathrm{L}) = \{A \mid \mathrm{L}(A) = l\}$ denote the set of arguments labelled with l by L. Intuitively, L interprets the arguments in $\mathrm{in}(\mathrm{L})$, $\mathrm{out}(\mathrm{L})$ and $\mathrm{und}(\mathrm{L})$ as accepted, rejected and undecided, respectively. Following [9], we call a labelling

Complete: if L satisfies

1. $\mathrm{L}(A) = \mathrm{in}$ if and only if $\mathrm{L}(B) = \mathrm{out}$ for all $B \in \mathrm{Att}(A)$.
2. $\mathrm{L}(A) = \mathrm{out}$ if and only if $\mathrm{L}(B) = \mathrm{in}$ for some $B \in \mathrm{Att}(A)$.

Furthermore, we say that a complete labelling is

Grounded: if $\mathrm{in}(\mathrm{L})$ is minimal with respect to set inclusion.
Preferred: if $\mathrm{in}(\mathrm{L})$ is maximal with respect to set inclusion.
Semi-stable: if $\mathrm{und}(\mathrm{L})$ is minimal with respect to set inclusion.
Stable: if $\mathrm{und}(\mathrm{L}) = \emptyset$.

For $\mathcal{S} \in \{$*Complete*, *Grounded*, *Prefered*, *Semi-stable*, *Stable*$\}$, we denote by $\mathcal{L}^{\mathcal{S}}$ the set of $\mathcal{S}$-Labellings. In the following, we abbreviate these semantics by c, g, p, ss, s . For $A \in \mathcal{A}$, we say that A is

credulously $\mathcal{S}$-accepted if $\mathrm{L}(A) = \mathrm{in}$ for some $\mathrm{L} \in \mathcal{L}^{\mathcal{S}}$,
sceptically $\mathcal{S}$-accepted if $\mathrm{L}(A) = \mathrm{in}$ for all $\mathrm{L} \in \mathcal{L}^{\mathcal{S}}$.

Classical computational problems are deciding whether an $\mathcal{S}$-Labelling exists, finding one or enumerating all $\mathcal{S}$-Labellings and deciding if arguments are credulously or sceptically $\mathcal{S}$-accepted. We will now show how some of these problems can be connected to probabilistic reasoning problems in Markov networks.

Given a set of arguments $\mathcal{A} = \{A_1, \ldots, A_n\}$, we introduce a corresponding set of ternary random variables $\mathbf{X} = \{X_1, \ldots, X_n\}$ each of which can take values from $\{\mathrm{in}, \mathrm{out}, \mathrm{und}\}$. For every labeling L, the corresponding variable assignment is denoted by $\mathbf{x}_\mathrm{L}$ and, conversely, the labelling corresponding to a complete variable assignment $\mathbf{x}$ is denoted by $\mathrm{L}_\mathbf{x}$. For each of the semantics that we reviewed, we will introduce a set of factors Φ such that labellings correspond to variable assignments with certain probabilities under P_Φ. We start with complete semantics.

Definition 1 (Complete Factors)**.** The set of complete factors $\Phi^c = \{\phi_1^c(S_1), \ldots, \phi_n^c(S_n)\}$ contains for every argument A_i exactly one factor $\phi_i^c(S_i)$ with scope $S_i = \{X_i\} \cup \mathrm{Att}(X_i)$ that is defined by the following case differentiation:

if $X_j = \mathrm{in}$ *for any* $X_j \in \mathrm{Att}(X_i)$ **then**
  if $X_i = \mathrm{out}$ **then return** 1 **else return** 0
else if $X_j = \mathrm{out}$ *for all* $X_j \in \mathrm{Att}(X_i)$ **then**
  if $X_i = \mathrm{in}$ **then return** 1 **else return** 0
else
  if $X_i = \mathrm{und}$ **then return** 1 **else return** 0

Intuitively, the factor goes through the three relevant cases of possible states of X_i's parents and returns 1 if X_i's state is valid under complete semantics and 0 otherwise. Let us also note that complete labellings always exist, so that the partition function Z is non-zero and the Gibbs distribution is well-defined. The following proposition summarizes some interesting relationships.

Proposition 1. *Let P_{Φ^c} be the Gibbs distribution over the set of complete factors. Then*

1. $\mathrm{Pl}_{\Phi^c}(\mathbf{x}) > 0$ *iff* $\mathrm{L}_\mathbf{x}$ *is a complete labelling.*
2. $\mathcal{L}^c = \{\mathrm{L}_\mathbf{x} \mid \mathrm{Pl}_{\Phi^c}(\mathbf{x}) > 0\}$ *is the set of complete labellings.*
3. *For all* $\mathrm{L} \in \mathcal{L}^c$, $\mathrm{Pl}_{\Phi^c}(\mathbf{x}_\mathrm{L}) = 1$ *and* $P_{\Phi^c}(\mathbf{x}_\mathrm{L}) = \frac{1}{|\mathcal{L}^c|}$.
4. *For all arguments* $A_i \in \mathcal{A}$,
   (a) $P_{\Phi^c}(X_i = \mathrm{in}) = \frac{c_i}{|\mathcal{L}^c|}$ *is the number c_i of complete labellings that accept A_i divided by the number of all complete labellings.*
   (b) *A_i is credulously accepted under complete semantics iff* $P_{\Phi^c}(X_i = \mathrm{in}) > 0$*,*
   (c) *A_i is sceptically accepted under complete semantics iff* $P_{\Phi^c}(X_i = \mathrm{in}) = 1$.

Proof. 1. First suppose that $\mathrm{L}_\mathbf{x}$ is a complete labelling. Consider an arbitrary argument A_i. If $\mathrm{L}_\mathbf{x}(A_i) = \mathrm{in}$, then $\mathrm{L}_\mathbf{x}(B) = \mathrm{out}$ for all attackers B of A_i. Hence, we are in the second branch of the if-condition and the factor returns 1. If $\mathrm{L}_\mathbf{x}(A_i) = \mathrm{out}$, then $\mathrm{L}_\mathbf{x}(B) = \mathrm{in}$ for some attacker B of A_i. Then we are in the first branch and the factor returns again 1. If $\mathrm{L}_\mathbf{x}(A_i) = \mathrm{und}$, the first two conditions are not met and we are in the third branch and the factor returns again 1. Hence, $\mathrm{Pl}_{\Phi^c}(\mathbf{x}) = 1 > 0$.

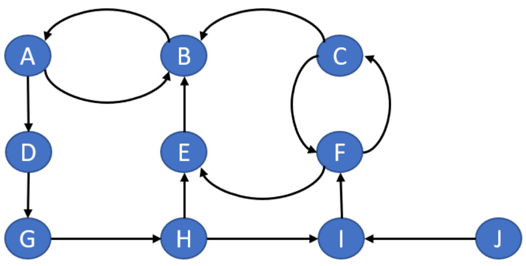

Figure 1. AAF: attack relations are denoted by directed edges.

Conversely, suppose that $\mathrm{Pl}_{\Phi^c}(\mathbf{x}) = 0$. Then there must be some $i \in \{1, \ldots, n\}$ such that $\phi_i^c(\mathbf{x}|_{S_i}) = 0$. If 0 was returned in the first branch, we have $\mathrm{L}_{\mathbf{x}}(A_i) \neq \mathrm{out}$ even though an attacker of A_i is in, hence $\mathrm{L}_{\mathbf{x}}$ is not complete. If 0 was returned in the second branch, we have $\mathrm{L}_{\mathbf{x}}(A_i) \neq \mathrm{in}$ even though all attackers of A_i are out and again $\mathrm{L}_{\mathbf{x}}$ is not complete. Finally, if 0 was returned in the third branch, no attacker of A_i is in (otherwise we were in the first branch), not all attackers of A_i are out (otherwise we were in the second branch) and A_i must be in or out. But none of these labellings is justified and so again $\mathrm{L}_{\mathbf{x}}$ is not complete.

2. Follows immediately from 1.

3. Follows from 1 because we showed that all labellings have plausibility 1 and all other labellings have plausibility 0. Hence, Z is just the number of complete labellings.

4. For a), note that we have $P_{\Phi^c}(X_i = \mathrm{in}) = \frac{1}{Z}\sum_{\mathrm{L}\in\mathcal{L}^c, \mathrm{L}(A_i)=\mathrm{in}} \mathrm{Pl}_{\Phi^c}(\mathbf{x}_\mathrm{L}) = \frac{c_i}{|\mathcal{L}^c|}$. b) and c) follow from a) by noting that credulous acceptance means that $c_i > 0$ and sceptical acceptance means $c_i = |\mathcal{L}^c|$. □

To illustrate the general idea of the reductions, let us first interpret the proposition from a computational perspective. Item 1 says that finding a complete labelling can be reduced to finding an assignment with positive probability. Therefore, it can, in particular, be reduced to a MAP query. Item 2 says that enumerating all complete labellings can be reduced to computing the support (the states with non-zero probability) of a Gibbs distribution. Item 3 says that all complete labellings are equally probable. This implies, in particular, that finding the number of complete labellings can be reduced to a MAP query because every preferred labelling L has maximum probability and the number of labellings is $|\mathcal{L}^c| = \frac{1}{P_{\Phi^c}(\mathbf{x}_\mathrm{L})}$. Item 4 shows that deciding credulous and sceptical acceptance can be reduced to computing marginal probabilities in Markov networks.

To illustrate the meaning of the Gibbs distribution, consider the AAF shown in Figure 1. The corresponding complete labellings are shown in the following table:

	A	B	C	D	E	F	G	H	I	J
L_1	out	in	out	in	out	in	out	in	out	in
L_2	und	und	out	und	out	in	und	und	out	in
L_3	und	und	und	und	und	und	und	und	out	in
L_4	in	out	out	out	out	in	in	out	out	in
L_5	in	out	und	out	und	und	in	out	out	in
L_6	in	out	in	out	in	out	in	out	out	in

The second column in Table 1 shows the marginal in and out probabilities for the arguments. Note that the probabilities for the labels have to sum up to 1 and therefore $P(X = \mathrm{und}) = 1 - P(X = \mathrm{in}) - P(X = \mathrm{out})$ for all random variables (arguments) X. We can

$\mathcal{A}$	c	g	p	ss	s
$P(A = \mathrm{in})$	$\frac{3}{6}$	0.22	0.6	0.67	$\frac{2}{3}$
$P(A = \mathrm{out})$	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
$P(B = \mathrm{in})$	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
$P(B = \mathrm{out})$	$\frac{3}{6}$	0.22	0.6	0.67	$\frac{2}{3}$
$P(C = \mathrm{in})$	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
$P(C = \mathrm{out})$	$\frac{3}{6}$	0.34	0.55	0.64	$\frac{2}{3}$
$P(D = \mathrm{in})$	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
$P(D = \mathrm{out})$	$\frac{3}{6}$	0.22	0.6	0.67	$\frac{2}{3}$
$P(E = \mathrm{in})$	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
$P(E = \mathrm{out})$	$\frac{3}{6}$	0.34	0.55	0.64	$\frac{2}{3}$
$P(F = \mathrm{in})$	$\frac{3}{6}$	0.34	0.55	0.64	$\frac{2}{3}$
$P(F = \mathrm{out})$	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
$P(G = \mathrm{in})$	$\frac{3}{6}$	0.22	0.6	0.67	$\frac{2}{3}$
$P(G = \mathrm{out})$	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
$P(H = \mathrm{in})$	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
$P(H = \mathrm{out})$	$\frac{3}{6}$	0.22	0.6	0.67	$\frac{2}{3}$
$P(I = \mathrm{in})$	0	0	0	0	0
$P(I = \mathrm{out})$	1	1	1	1	1
$P(J = \mathrm{in})$	1	1	1	1	1
$P(J = \mathrm{out})$	0	0	0	0	0

Table 1. Marginal probabilities (rounded) under different semantics for Figure 1. Note that $P(X = \mathrm{und}) = 1 - P(X = \mathrm{in}) - P(X = \mathrm{out})$.

see that, under complete semantics, all arguments but I are credulously accepted and that only J is sceptically accepted. Furthermore, the partition function is $Z = 6$, so that the number of states with non-zero probability is 6. These states correspond to the complete labellings and their probabilities are shown in Table 2. We can, in

	c	g	p	ss	s
L_1	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$
L_2	$\frac{1}{6}$	0.25	0.04	0.01	0
L_3	$\frac{1}{6}$	0.5	0.02	0.00	0
L_4	$\frac{1}{6}$	0.06	0.17	0.32	$\frac{1}{3}$
L_5	$\frac{1}{6}$	0.13	0.09	0.03	0
L_6	$\frac{1}{6}$	0.03	0.34	0.32	$\frac{1}{3}$

Table 2. Probabilities of complete labellings under different semantics.

particular, see that the probability of each argument is just the relative frequency of complete labellings that accept it as explained in item 4 of Proposition 1.

Roughly speaking, grounded, preferred and semi-stable semantics minimize or maximize the number of occurences of one particular label. In order to encode this intuition in Markov networks, we can add one unary factor for every variable that rewards or penalizes particular labels.

Definition 2 (l-w-Factor)**.** The set of l-w-factors $\Phi^{(l,w)} = \{\phi_1^{(l,w)}(X_1), \ldots, \phi_n^{(l,w)}(X_n)\}$ contains one unary factor

$$\phi_i^{(l,w)}(X_i) = \begin{cases} w, & \text{if } X_i = l \\ 1, & \text{else} \end{cases}$$

for every argument A_i.

If the labelling L assigns the label l to k arguments, then the corresponding plausibility distribution $\mathrm{Pl}_{\Phi^{(l,w)}}$ yields $\mathrm{Pl}_{\Phi^{(l,w)}}(\mathbf{x}_\mathrm{L}) = w^k$. In particular, when we combine the l-w-factors with the complete factors, we get $\mathrm{Pl}_{\Phi^c \cup \Phi^{(l,w)}}(\mathbf{x}_\mathrm{L}) = \mathrm{Pl}_{\Phi^c}(\mathbf{x}_\mathrm{L}) \cdot w^k$. If $w > 1$, this allows increasing the plausibility (and thus the probability) of a complete labelling with a particular label. Symmetrically, $w < 1$, allows decreasing the plausibility of complete labelling with a particular label. We consider the following l-w-factors.

Grounded Factors: $\Phi^g = \Phi^{(\mathrm{in},0.5)}$.
Preferred Factors: $\Phi^p = \Phi^{(\mathrm{in},2)}$.
Semi-stable Factors: $\Phi^{ss} = \Phi^{(\mathrm{und},0.5)}$.
Stable Factors: $\Phi^s = \Phi^{(\mathrm{und},0)}$.

For grounded, preferred and semi-stable semantics, we do not get as much information as before from the Gibbs distribution. However, labellings now correspond to states with maximum probability and can therefore be computed via MAP queries.

Proposition 2. *1.* $\mathrm{L}_\mathbf{x}$ *is the grounded labelling if and only if* $\mathrm{Pl}_{\Phi^c \cup \Phi^g}(\mathbf{x}) > \mathrm{Pl}_{\Phi^c \cup \Phi^g}(\mathbf{y})$ *for all other assignments* $\mathbf{y}$.
2. $\mathrm{L}_\mathbf{x}$ *is a preferred labelling if and only if* $\mathrm{Pl}_{\Phi^c \cup \Phi^p}(\mathbf{x}) \geq \mathrm{Pl}_{\Phi^c \cup \Phi^p}(\mathbf{y})$ *for all other assignments* $\mathbf{y}$.
3. $\mathrm{L}_\mathbf{x}$ *is a semi-stable labelling if and only if* $\mathrm{Pl}_{\Phi^c \cup \Phi^{ss}}(\mathbf{x}) \geq \mathrm{Pl}_{\Phi^c \cup \Phi^{ss}}(\mathbf{y})$ *for all other assignments* $\mathbf{y}$.

Proof. 1. The grounded labelling is unique and minimizes the number of arguments that are in. Therefore, its plausibility is minimally decreased and it has maximum probability under $\mathrm{Pl}_{\Phi^c \cup \Phi^g}(\mathbf{x})$.

2 and 3 follow analogously. The only difference is now that several labellings can exist. However, by construction, they all have maximum probability. □

The result is not particularly interesting for grounded semantics because a grounded labelling can be computed in polynomial time [15]. However, we include it for completeness and in order to compare the resulting probabilities. The third, fourth and fifth columns in Tables 1 and 2 show the marginal probabilities of arguments and the probabilities of labellings. For example, L_3 in Table 2 has maximum probability under grounded factors and is therefore the grounded labelling. The remaining complete labellings also have a non-zero probability, but the probability decreases with the number of arguments that are labelled in. Similarly, the preferred and semi-stable labellings have maximum probability under preferred and semi-stable factors, respectively. The resulting marginal probabilities do not have an easy interpretation, but seem to be plausible degrees of belief.

We get again stronger guarantees for stable labellings because we can now just set the plausibility of a labelling to 0 if it contains an und-label. However, stable labellings do not necessarily exist, so that $P_{\Phi^c \cup \Phi^s}$ may actually not be a well-defined probabilistic model (in this case, all plausibilities are 0 and therefore cannot be normalized to sum up to 1 as required for a probability model). Note that this is the case if and only if the partition function Z is 0. If $Z \neq 0$, there are again several interesting relationships that follow immediately from the definition of the factors. The proof is analogous to the proof of Proposition 1 and is therefore left out.

Proposition 3. *Let* $P_{\Phi^c \cup \Phi^s}$ *be the Gibbs distribution over the set of complete and stable factors.*

1. *The partition function* Z *is non-zero if and only if a stable labelling exists.*
2. *If* $Z \neq 0$*, then* $\mathrm{Pl}_{\Phi^c \cup \Phi^s}(\mathbf{x}) > 0$ *if and only if* $\mathrm{L}_\mathbf{x}$ *is a stable labelling.*
3. *If* $Z \neq 0$*, then* $\mathcal{L}^s = \{\mathrm{L}_\mathbf{x} \mid \mathrm{Pl}_{\Phi^c \cup \Phi^s}(\mathbf{x}) > 0\}$ *is the set of stable labellings.*
4. *If* $Z \neq 0$*, then for all stable labellings* L, $\mathrm{Pl}_{\Phi^c \cup \Phi^s}(\mathbf{x}_\mathrm{L}) = 1$ *and* $P_{\Phi^c \cup \Phi^s}(\mathbf{x}_\mathrm{L}) = \frac{1}{|\mathcal{L}^s|}$.
5. *If* $Z \neq 0$*, then for all arguments* $A_i \in \mathcal{A}$*,*
   (a) $P_{\Phi^c \cup \Phi^s}(X_i = \mathrm{in}) = \frac{s_i}{|\mathcal{L}^s|}$ *is the number* s_i *of stable labellings that accept* A_i *divided by the number of all stable labellings.*
   (b) A_i *is credulously accepted under stable semantics if and only if* $P_{\Phi^c \cup \Phi^s}(X_i = \mathrm{in}) > 0$*,*
   (c) A_i *is sceptically accepted under stable semantics if and only if* $P_{\Phi^c \cup \Phi^s}(X_i = \mathrm{in}) = 1$.

For our running example in Figure 1, the partition function Z is 3, which tells us that there are three stable labellings. Similar as for semi-stable factors, L_1, L_4 and L_6 in Table 2 have maximum probability under stable factors and are therefore the stable labellings. The remaining complete labellings have probability zero now. The resulting marginal probabilities can be seen in the Stbl-column of Table 1. As under complete semantics, all arguments but I are credulously accepted and only J is sceptically accepted. The probabilities of arguments correspond again to the relative frequency of labellings that accept them. Note that, since stable labellings never label arguments as undecided, the probabilities of in and out always sum up to 1 now.

4 Semantical Evaluation

Our constructed Markov networks allow us to draw conclusions about classical abstract argumentation problems from the generated probabilities. However, we could also ask, are the resulting probabilities meaningful on their own from a probabilistic reasoning perspective? Hunter and Thimm proposed several properties that can be interesting to compare different probabilistic argumentation approaches [21]. They considered binary arguments (accept/reject), whereas we consider ternary arguments here (in, out, und) since we build up on labellings. In order to transfer the ideas from [21] to our setting, we identify accepted arguments with in-labelled arguments as usual [9]. We focus on the atomic properties here and do not talk about combined properties that are satisfied when several atomic properties are satisfied. Consider the following properties from [21]:

COH: P is called *coherent* iff $P(B = \mathrm{in}) \leq 1 - P(A = \mathrm{in})$ whenever $(A, B) \in \mathrm{Att}$.
RAT: P is called *rational* iff $P(A = \mathrm{in}) > 0.5$ implies $P(B = \mathrm{in}) \leq 0.5$ whenever $(A, B) \in \mathrm{Att}$.
INV: P is called *involutary* iff $P(A = \mathrm{in}) = 1 - P(B = \mathrm{in})$ whenever $(A, B) \in \mathrm{Att}$.
OPT: P is called *optimistic* iff $P(A = \mathrm{in}) \geq 1 - \sum_{B \in \mathrm{Att}(A)} P(B = \mathrm{in})$.
FOU: P is called *founded* iff $\mathrm{Att}(A) = \emptyset$ implies $P(A = \mathrm{in}) = 1$.

Let us give a rough intuitive interpretation of these properties. COH says that the belief in an attacked argument should decrease as the belief in an attacker increases. RAT is more cautious: if the belief in an attacker is greater than 0.5 (we tend to accept the attacker), the belief in the attacked argument should not be larger than 0.5 (we do not tend to accept the attacked). INV is stronger than COH and demands that the bound is satisfied with equality. OPT gives a lower bound for the belief in arguments that decreases with the belief in the attackers. In particular, for unattacked arguments, it just demands that the belief in this argument must be 1. The latter is also what FOU demands.

Proposition 4. *$P_{\Phi^c}, P_{\Phi^c \cup \Phi^g}, P_{\Phi^c \cup \Phi^p}, P_{\Phi^c \cup \Phi^{ss}}, P_{\Phi^c \cup \Phi^s}$ satisfy COH, RAT and FOU. $P_{\Phi^c \cup \Phi^s}$ also satisfies OPT.*

Proof. Let $P \in \{P_{\Phi^c \cup \Phi^c}, P_{\Phi^c \cup \Phi^g}, P_{\Phi^c \cup \Phi^p}, P_{\Phi^c \cup \Phi^{ss}}, P_{\Phi^c \cup \Phi^s}\}$, $(A, B) \in \mathrm{Att}$ and let $\mathcal{L}^c_A = \{\mathrm{L} \mid L(A) = \mathrm{in}\}$ denote those complete labellings that label A in and let $\mathcal{L}^c_B$ denote those complete labellings that label B in. Since complete labellings cannot accept both A and B, we have $\mathcal{L}^c_A \cap \mathcal{L}^c_B = \emptyset$ and therefore $P(A = \mathrm{in}) + P(B = \mathrm{in}) \leq \sum_{\mathrm{L} \in \mathcal{L}^c_A} P(\mathrm{L}) + \sum_{\mathrm{L} \in \mathcal{L}^c_B} P(\mathrm{L}) \leq \sum_{\mathrm{L} \in \mathcal{L}^c} P(\mathrm{L}) = 1$. This implies COH.

RAT follows from COH because, if $P(A = \mathrm{in}) > 0.5$, then COH implies $P(B = \mathrm{in}) \leq 1 - P(A = \mathrm{in}) < 0.5$.

If $\mathrm{Att}(A) = \emptyset$, all complete labellings label A in. Since the complete factors assign zero probability to all other labellings and the probabilities of all labellings have to sum up to 1, we have $P(A = \mathrm{in}) = 1$. Thus, FOU is satisfied as well.

Finally, for OPT, let $\mathcal{L}^s_A$ denote those stable labellings that label A in. Since stable labellings do not use the und label, $\mathcal{L}^s \setminus \mathcal{L}^s_A$ is the set of stable labellings that label A out. But whenever A is labelled out, at least one attacker of A must be labelled in. Thus, $\sum_{B \in \mathrm{Att}(A)} P_{\Phi^c \cup \Phi^s}(B = \mathrm{in}) \geq \sum_{\mathrm{L} \in \mathcal{L}^s \setminus \mathcal{L}^s_A} P_{\Phi^c \cup \Phi^s}(\mathrm{L})$ and therefore $P_{\Phi^c \cup \Phi^s}(A = \mathrm{in}) = 1 - \sum_{\mathrm{L} \in \mathcal{L}^s \setminus \mathcal{L}^s_A} P_{\Phi^c \cup \Phi^s}(\mathrm{L}) \geq 1 - \sum_{B \in \mathrm{Att}(A)} P_{\Phi^c \cup \Phi^s}(B = \mathrm{in})$. □

To see that $P_{\Phi^c}, P_{\Phi^c \cup \Phi^g}, P_{\Phi^c \cup \Phi^p}, P_{\Phi^c \cup \Phi^{ss}}$ can violate OPT and INV, consider again the graph in Figure 1. We can see from Table 1 that they all satisfy $P(A = \mathrm{in}) < 1 - P(B = \mathrm{in})$. However, OPT demands that $P(A = \mathrm{in}) \geq 1 - P(B = \mathrm{in})$ and INV demands $P(A = \mathrm{in}) = 1 - P(B = \mathrm{in})$.

$P_{\Phi^c \cup \Phi^s}$ can violate INV as well. In Figure 1, INV demands that $P(I) = 1 - P(J)$ and $P(I) = 1 - P(H)$ and thus $P(H) = P(J)$, which is violated by $P_{\Phi^c \cup \Phi^s}$. Indeed, INV implies, in general, that all arguments that attack the same argument must have the same probability. This may be a too strong requirement. For example, one could argue that, in Figure 1, H should have a smaller probability than J because H is attacked by G, whereas J does not have any attackers. A case where INV is less controversial is when there is only one attacker. It is interesting to note that, in this case, $P_{\Phi^c \cup \Phi^s}$ does indeed satisfy the condition of INV.

Proposition 5. *If $\mathrm{Att}(A) = \{B\}$, then $P_{\Phi^c \cup \Phi^s}(A = \mathrm{in}) = 1 - P_{\Phi^c \cup \Phi^s}(B = \mathrm{in})$.*

Proof. Let again $\mathcal{L}^s_A$ denote those stable labellings that label A in. As before, $\mathcal{L}^s \setminus \mathcal{L}^s_A$ is the set of stable labellings that label A out and since B is the only attacker of A, B must be labelled in. Thus, $1 = \sum_{\mathrm{L} \in \mathcal{L}^s_A} P_{\Phi^c \cup \Phi^s}(\mathrm{L}) + \sum_{\mathrm{L} \in \mathcal{L}^s \setminus \mathcal{L}^s_A} P_{\Phi^c \cup \Phi^s}(\mathrm{L}) = P_{\Phi^c \cup \Phi^s}(A = \mathrm{in}) + P_{\Phi^c \cup \Phi^s}(B = \mathrm{in})$, which implies the claim. □

5 Extension to Bipolar Argumentation

In many applications it is useful to consider not only attack relations, but also support relations. A *bipolar argumentation framework (BAF)* is a tuple $(\mathcal{A}, \mathrm{Att}, \mathrm{Sup})$, where $\mathcal{A}$ is a finite set of arguments, $\mathrm{Att} \subseteq \mathcal{A} \times \mathcal{A}$ is the *attack relation* as before and Sup is the support relation. We also assume that $\mathrm{Att} \cap \mathrm{Sup} = \emptyset$. We let again $\mathrm{Sup}(A) = \{B \mid (B, A) \in \mathrm{Sup}\}$ denote the set of supporters of A.

The meaning of the support relation can be defined in different ways. The intuitve idea of *deductive support* [8, 11] is that if an argument is accepted, the argument that it supports must be accepted as well. In this way, arguments can also indirectly support arguments via chains of support relations. The dual idea (if an argument is accepted, all its supporters must be accepted as well) is referred to as *necessary support* [11]. However, deductive support relations can be tranlsated to necessary support relations by just reversing their direction, so we will focus on deductive support here. A deeper discussion of both and other notions of support can be found in [11]. One way to give formal meaning to deductive support relations is to translate the bipolar argumentation framework to an abstract argumentation framework with additional attacks [11]. These new attacks correspond to indirect attacks that are created by the interplay between attack and support relations. The following indirect attacks have been considered for this purpose in [11]:

Supported Attack from A to B: there is a sequence of arguments $A_1, \ldots, A_n$ such that $A_1 = A$, $A_n = B$, $(A_i, A_{i+1}) \in \mathrm{Sup}$ for $1 \leq i \leq n-2$ and $(A_{n-1}, A_n) \in \mathrm{Att}$.

Mediated Attack from A to B: there is a sequence of arguments $A_1, \ldots, A_n$ such that $A_1 = B$, $A_n = A$, $(A_i, A_{i+1}) \in \mathrm{Sup}$ for $1 \leq i \leq n-2$ and $(A_n, A_{n-1}) \in \mathrm{Att}$.

Intuitively, there is a supported attack from A to B iff A directly or indirectly supports an attacker of B. There is a mediated attack from A to B iff A attacks an argument that is directly or indirectly supported by B. The *Dung framework associated with* $(\mathcal{A}, \mathrm{Att}, \mathrm{Sup})$ is then defined as the Dung framework $(\mathcal{A}, \mathrm{Att} \cup \mathrm{Att}^s \cup \mathrm{Att}^m)$, where Att^s and Att^m contain edges that correspond to supported and mediated attacks in $(\mathcal{A}, \mathrm{Att}, \mathrm{Sup})$ [11]. This is an elegant way to extend established semantics to bipolar argumentation frameworks. However, this solution does not treat attack and support equally. For example, if we have three arguments A, B, C such that A attacks C and B supports C, it seems that C could as well be accepted as rejected. However, the translation takes only account of the mediated attack from A to B and ignores the fact that C has a supporter. Hence, the only complete labelling of the associated Dung framework labels A in and B and C out. The disparity between attack and support becomes more prevalent when we keep our attacker A, but add n supporters B_i of C. Then A will be accepted and C and all of its supporters B_i will be rejected even when C is supported by thousands of arguments.

In order to obtain equal treatment of attack and support, we adapt the definition of complete labellings. We will show that the resulting definition still respects supported attacks and mediated attacks. Completeness makes sufficient and necessary conditions for acceptance and rejection of arguments. In order to treat attack and support symmetrically, we weaken the conditions for attacks and add dual conditions for supports. We call a labelling

Deductive: if L satisfies

1. If $\mathrm{L}(A) = \mathrm{in}$, then $\mathrm{L}(B) = \mathrm{out}$ for all $B \in \mathrm{Att}(A)$.
2. If $\mathrm{L}(A) = \mathrm{out}$, then $\mathrm{L}(B) = \mathrm{out}$ for all $B \in \mathrm{Sup}(A)$.
3. If $\mathrm{L}(B) = \mathrm{in}$ for some $B \in \mathrm{Att}(A)$, then $\mathrm{L}(A) = \mathrm{out}$.
4. If $\mathrm{L}(B) = \mathrm{in}$ for some $B \in \mathrm{Sup}(A)$, then $\mathrm{L}(A) = \mathrm{in}$.

Conditions 1 and 2 correspond to necessary conditions for acceptance and rejection. We can accept (reject) only when all attackers (supporters) are out. Conditions 3 and 4 correspond to sufficient conditions. If one attacker (supporter) is in, the argument must be out (in). Condition 1 and 3 make sure that deductive labellings respect classical attacks. Conditions 2 and 4 make sure that they interpret support in a dual manner. The following proposition shows that deductive labellings also respect supported and mediated attacks.

Proposition 6. *Let* L *be a* deductive *labelling.*

1. *If* $\mathrm{L}(A) = \mathrm{in}$ *and there is a supported attack from* A *to* B*, then* $\mathrm{L}(B) = \mathrm{out}$.
2. *If* $\mathrm{L}(A) = \mathrm{in}$ *and there is a mediated attack from* A *to* B*, then* $\mathrm{L}(B) = \mathrm{out}$.

Proof. 1. Consider the sequence of arguments $A_1, \ldots, A_n$ given by the definition of supported attack. Condition 4 of a dual labelling implies that $\mathrm{L}(A_i) = \mathrm{in}$ for $1 \leq i \leq n-1$. Then condition 3 implies that $\mathrm{L}(A_n) = \mathrm{L}(B) = \mathrm{out}$.

2. Consider the sequence of arguments $A_1, \ldots, A_n$ given by the definition of mediated attack. Condition 3 of a dual labelling implies that $\mathrm{L}(A_{n-1}) = \mathrm{out}$. Condition 2 then implies that $\mathrm{L}(A_i) = \mathrm{out}$ for $1 \leq i \leq n-2$ and hence $\mathrm{L}(B) = \mathrm{out}$. □

We introduce again a set of factors for deductive labellings. It is interesting to note that our deductive semantics can be encoded by pairwise factors only, which can be computationally beneficial.

Definition 3 (Deductive Factors)**.** The set of deductive factors $\Phi^d = \{\phi_1^d(S_1), \ldots, \phi_m^d(S_m)\}$ contains for every edge $(A_j, A_k) \in \mathrm{Att} \cup \mathrm{Sup}$ exactly one factor $\phi_i^d(S_i)$ with scope $S_i = \{X_j, X_k\}$ that is defined by

$$\phi_i^d(X_j, X_k) = \begin{cases} 0 & \textit{if } X_j = \mathrm{in} \wedge X_k \neq \mathrm{out} \\ & \textit{or } X_j \neq \mathrm{out} \wedge X_k = \mathrm{in} \\ 1 & \textit{otherwise} \end{cases}$$

if $(A_j, A_k) \in \mathrm{Att}$ and by

$$\phi_i^d(X_j, X_k) = \begin{cases} 0 & \textit{if } X_j = \mathrm{in} \wedge X_k \neq \mathrm{in} \\ & \textit{or } X_j \neq \mathrm{out} \wedge X_k = \mathrm{out} \\ 1 & \textit{otherwise} \end{cases}$$

if $(A_j, A_k) \in \mathrm{Sup}$.

The resulting Gibbs distribution assigns again uniform probability to all deductive labellings and probability 0 to all others.

Proposition 7. *Let* $(\mathcal{A}, \mathrm{Att})$ *be an AAF and let* P_{Φ^d} *be the Gibbs distribution over the set of deductive factors. Then* $\mathrm{Pl}_{\Phi^d}(\mathbf{x}) > 0$ *iff* $\mathrm{L}_{\mathbf{x}}$ *is a deductive labelling.*

Proof. First, let $\mathrm{L}_{\mathbf{x}}$ be a deductive labelling and consider an arbitrary argument A_i. If $\mathrm{L}_{\mathbf{x}}(A_i) = \mathrm{in}$, then all attackers of A_i are out (condition 1 of deductive labelling) and all arguments attacked by A_i are out (condition 3). This guarantees that any attack factor involving X_i returns 1. Furthermore, all arguments supported by A_i must be in (condition 4). This guarantees that any support factor involving A_i must return 1. If $\mathrm{L}_{\mathbf{x}}(A_i) = \mathrm{out}$, every attack factor must return 1 and since every supporter of A_i must be out (condition 2), every support factor must return 1 as well. If $\mathrm{L}_{\mathbf{x}}(A_i) = \mathrm{und}$, no attacker and supporter of A_i can be in (contrapositive of conditions 3 and 4), and no argument attacked by A_i can be in (contrapositive of condition 1) and no argument supported by A_i can be out (contrapositive of condition 2). Hence, the attack and support factors involving X_i must return again 1. Hence, $\mathrm{Pl}_{\Phi^c}(\mathbf{x}) = 1 > 0$.

Conversely, suppose that $\mathrm{Pl}_{\Phi^d}(\mathbf{x}) = 0$. Then there must be some $i \in \{1, \ldots, n\}$ such that $\phi_i^d(X_j, X_k) = 0$. By going through the four possible cases that could return a 0, we can see that $\mathrm{L}_{\mathbf{x}}$ is not a deductive labelling. □

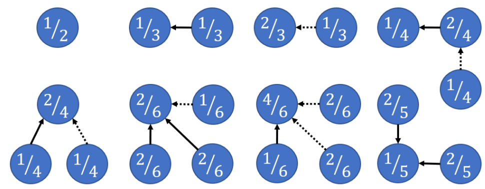

Figure 2. in-probabilities for arguments under deductive stable semantics for different bipolar graphs. Attack relations are denoted by dashed edges, support relations by dotted edges.

We could prove analogous properties to Proposition 1. However, we omit the explicit statements to save space. Instead, let us note that we can extend the deductive factors with grounded, preferred, semi-stable factors and stable factors as before. We obtained the strongest properties for the stable semantics and will therefore only look at this case. By a *deductive stable labelling*, we mean a labelling that is deductive and labels every argument as either in or out. There always exists a deductive stable labelling because the labelling that rejects all arguments is deductive. Hence, the partition function is guaranteed to be non-zero.

Proposition 8. *Let* $P_{\Phi^d \cup \Phi^s}$ *be the Gibbs distribution over the deductive and stable factors. Then*

1. $\mathrm{Pl}_{\Phi^d \cup \Phi^s}(\mathbf{x}) > 0$ *if and only if* $\mathrm{L}_{\mathbf{x}}$ *is a deductive stable labelling.*
2. $\mathcal{L}^{ds} = \{\mathrm{L}_{\mathbf{x}} \mid \mathrm{Pl}_{\Phi^d \cup \Phi^s}(\mathbf{x}) > 0\}$.
3. *For all deductive stable labellings* L*,* $\mathrm{Pl}_{\Phi^d \cup \Phi^s}(\mathbf{x}_\mathrm{L}) = 1$ *and* $P_{\Phi^d \cup \Phi^s}(\mathbf{x}_\mathrm{L}) = \frac{1}{|\mathcal{L}^{ds}|}$.
4. *For all arguments* $A_i \in \mathcal{A}$*,* $P_{\Phi^d \cup \Phi^s}(X_i = \mathrm{in}) = \frac{d_i}{|\mathcal{L}^s|}$ *is the number* d_i *of deductive stable labellings that accept* A_i *divided by the number of all deductive stable labellings.*

The proof is analogous to the previous ones and is therefore left out. Figure 2 shows some acceptance probabilities. The probabilities are again just the relative frequencies of deductive stable labellings that accept the argument. The deductive semantics is more credulous than the complete semantics because arguments cannot only be accepted, but can also be rejected without reason. An argument with an equal number of equally strong attackers and supporters has probability 0.5. If the number of attackers (supporters) is larger than the number of supporters (attackers), the belief in the argument decreases (increases). Note, in particular, that this happens in a symmetric fashion for attack and support. At first sight, it seems strange that, in the attack-only graph (top row, second graph from the left), attacker and attacked have equal probability. However, one may argue that since both are unsupported, one may belief one as well as the other. As we add 'support' in form of a supported attack or a second attack, the probabilities of the attacker increases and the probability of the attacked decreases as the graphs on the right in Figure 2 illustrate.

Let us note that $P_{\Phi^d \cup \Phi^s}$ still satisfies COH and RAT. In particular, it satisfues dual properties for support.

Proposition 9. $P_{\Phi^d \cup \Phi^s}$ *satisfies COH, RAT and*

1. *D-COH:* $P_{\Phi^d \cup \Phi^s}(B = \mathrm{in}) \geq P_{\Phi^{ds}}(A = \mathrm{in})$ *whenever* $(A, B) \in \mathrm{Sup}$.
2. *D-RAT:* $P_{\Phi^d \cup \Phi^s}(A = \mathrm{in}) > 0.5$ *implies* $P_{\Phi^d \cup \Phi^s}(B = \mathrm{in}) > 0.5$ *whenever* $(A, B) \in \mathrm{Sup}$.

Proof. The proofs for COH and RAT are analogous to the previous proofs and therefore left out. For D-COH, consider $(A, B) \in$

Sup and let $\mathcal{L}_A^{ds} = \{\mathrm{L} \mid L(A) = \mathrm{in}\}$ denote those deductive stable labellings that label A in and let $\mathcal{L}_B^{ds}$ denote those deductive stable labellings that label B in. Since deductive labellings must accept B whenever they accept A, we have $\mathcal{L}_A^{ds} \subseteq \mathcal{L}_B^{ds}$ and therefore $P_{\Phi^d \cup \Phi^s}(A = \mathrm{in}) = \sum_{\mathrm{L} \in \mathcal{L}_A^{ds}} P_{\Phi^d \cup \Phi^s}(\mathrm{L}) \leq \sum_{\mathrm{L} \in \mathcal{L}_B^{ds}} P_{\Phi^d \cup \Phi^s}(\mathrm{L}) = P_{\Phi^d \cup \Phi^s}(A = \mathrm{in})$. This proves D-COH. D-RAT follows from D-COH because, if $P_{\Phi^d \cup \Phi^s}(A = \mathrm{in}) > 0.5$, then D-COH implies $P_{\Phi^d \cup \Phi^s}(B = \mathrm{in}) \geq P_{\Phi^d \cup \Phi^s}(A = \mathrm{in}) > 0.5$. □

The remaining properties by Hunter and Thimm [21] need to be refined for the bipolar setting. For example, FOU would cause a conflict in the graph in the lower left of Figure 2 because we would have to accept both the attacker and the supporter. Since OPT implies FOU, similar problems occur for OPT. INV is not readily compatible with support edges either. A natural adaptation of FOU is to demand that arguments must be accepted when not only the attackers are out, but also no attacked arguments is in. This would lift the probability of all isolated arguments to 1, while still maintaining consistency with the meaning of support edges. We will discuss this and some other alternatives in more detail in future work.

6 Related Work

Markov networks are a very general and flexible tool and have been applied to various problems in artificial intelligence. Some recent applications in collaborative filtering, fake news detection and classification can be found in [39, 28, 38]. One of the most popular applications in the area of knowledge representation and reasoning are probably Markov Logic Networks that combine logic and probability theory [34]. The basic idea is not so different from the ideas here. The probability of interpretations can increase or decrease based on the probability of weighted formulas that they satisfy. One reason for the success of Markov logic networks is perhaps that the weights can naturally be learned from data, which makes them a natural tool for explainable AI. Some recent applications in question answering, entity resolution and activity recognition in smart homes can be found in [22, 18, 20]. It is, indeed, a natural idea to regard the weights in our factors as parameters that are not pre-defined, but can be learned from data by maximizing the probability of particular labels for output arguments when given the state of other arguments as input. Related ideas have been studied in [37] for the special case of Boltzmann machines. Some other related ideas for learning argumentation frameworks from data can be found in [23, 24].

The state-of-the-art for reasoning algorithms for abstract argumentation is perhaps best reflected by the International Competition on Computational Models of Argumentation [41, 19]. While also direct approaches compete, some of the most successful participants reduce the argumentation problems to SAT, ASP or CSP problems. For example, pyglaf [1], Cegartix [16] and LabSAT [6] use SAT-solvers, ASPARTIX [17] uses an ASP-solver and CoQuiAAS [26] uses a CSP solver.

We already discussed deductive and necessary support [8, 11] in the previous section. Another interesting support interpretation is evidential support [29]. Here, a sceptical stance is taken and nothing should be accepted unless supported by evidence. To this end, *prima-facie* arguments are introduced that must be accepted. Other arguments can only be accepted if they are directly or indirectly supported by these arguments. Let us note that we can incorporate prima-facie arguments in our deductive Gibbs model by just adding unary prima-facie factors that return 1 if the argument is in and 0 otherwise. Condition 4 of deductive labellings guarantees then that arguments that are directly or indirectly supported by prima-facie arguments are accepted as well. By this encoding, we do not capture the necessary condition, that an accepted argument must be supported by a prima-facie argument. However, the resulting semantics may be a nice combination of deductive and evidential semantics that may be worth studying in future work.

7 Discussion and Conclusions

We explained how argumentation problems can be encoded as Markov networks and discussed relationships between their reasoning problems. From a computational perspective, Markov networks seem most promising for combinatorial tasks like deciding sceptical inference or counting the number of labellings. We conducted some simple experiments with exact inference algorithms (Belief Propagation) to compute the marginal probabilities of arguments for randomly generated graphs under complete and stable semantics with pgmpy[2] (Python) and Mallet[3] (Java). Recall that this allows deciding sceptical inference. pgmpy is rather slow, but can deal with about 20 arguments in seconds on a regular desktop computer. Mallet is significantly faster and can deal with 60 arguments in seconds. Since the theoretical number of candidate labellings is 3^{60}, this is a decent runtime. We could probably scale up further by using a C++ library like OpenGM2[4]. However, this alone may not be sufficient to compete with state-of-the-art solvers. Some tracks in the International Competition of Computational Models of Argumentation[5] contain sceptical inference as a subtask. While there seem to be no explicit runtime results for sceptical inference available, it seems that the currently best solvers can deal with hundreds of arguments in reasonable time [19]. Out of the box, our reduction is probably not competitive because attack-only argumentation frameworks leave only little freedom in interpreting the arguments, so that algorithms profit greatly from preprocessing (e.g., accept all arguments without attackers, reject their direct successors and repeat) and constraint propagation (e.g., reject attackers and successors of in-labelled arguments). We are currently experimenting with different ideas to combine exact and approximate inference algorithms for Markov networks with preprocessing techniques and constraint solvers.

Our reductions may have a greater potential in bipolar settings. For example, consider again the graph with edges (A, C) and (B, C). If both edges are attacks, we would just accept A and B and reject C. However, provided that we treat attack and support relations equally, if (A, C) is an attack and (B, C) is a support, we could accept A and reject B, C or reject A and accept B, C. By adding n copies of this structure, we can see that there are graphs where replacing n attacks with supports leads from an attack-only graph with a single labelling to a bipolar graph with 2^n labellings. Of course, the number of labellings depends on the actual semantics, but in bipolar settings, the number of labellings can be significantly larger. Then exploiting symmetrical structure becomes more important and this is a well studied problem for Markov networks.

Arguably, our reduction also makes a nice connection between classical and probabilistic argumentation approaches. We will explore connections to other probabilistic reasoning approaches in more detail in future work and will also look at the properties of some alternative factorizations.

[2] https://github.com/pgmpy/pgmpy
[3] http://mallet.cs.umass.edu/
[4] https://hci.iwr.uni-heidelberg.de/opengm2/
[5] http://argumentationcompetition.org

REFERENCES

[1] Mario Alviano, 'The pyglaf argumentation reasoner', in *International Conference on Logic Programming (ICLP)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, (2018).

[2] Leila Amgoud and Jonathan Ben-Naim, 'Ranking-based semantics for argumentation frameworks', in *International Conference on Scalable Uncertainty Management (SUM)*, pp. 134–147. Springer, (2013).

[3] Leila Amgoud and Jonathan Ben-Naim, 'Evaluation of arguments in weighted bipolar graphs', in *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU)*, pp. 25–35. Springer, (2017).

[4] Leila Amgoud, Claudette Cayrol, and Marie-Christine Lagasquie-Schiex, 'On the bipolarity in argumentation frameworks.', in *International Workshop on Non-Monotonic Reasoning (NMR)*, volume 4, pp. 1–9, (2004).

[5] Howard Barringer, Dov M Gabbay, and John Woods, 'Temporal, numerical and meta-level dynamics in argumentation networks', *Argument & Computation*, **3**(2-3), 143–202, (2012).

[6] Christoph Beierle, Florian Brons, and Nico Potyka, 'A software system using a sat solver for reasoning under complete, stable, preferred, and grounded argumentation semantics', in *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*, pp. 241–248. Springer, (2015).

[7] Philippe Besnard and Anthony Hunter, 'A logic-based theory of deductive arguments', *Artificial Intelligence*, **128**(1-2), 203–235, (2001).

[8] Guido Boella, Dov M Gabbay, Leon van der Torre, and Serena Villata, 'Support in abstract argumentation', in *International Conference on Computational Models of Argument (COMMA)*, pp. 40–51. Frontiers in Artificial Intelligence and Applications, IOS Press, (2010).

[9] Martin WA Caminada and Dov M Gabbay, 'A logical account of formal argumentation', *Studia Logica*, **93**(2-3), 109, (2009).

[10] Claudette Cayrol and Marie-Christine Lagasquie-Schiex, 'Graduality in argumentation.', *Journal of Artificial Intelligence Research (JAIR)*, **23**, 245–297, (2005).

[11] Claudette Cayrol and Marie-Christine Lagasquie-Schiex, 'Bipolarity in argumentation graphs: Towards a better understanding', *International Journal of Approximate Reasoning*, **54**(7), 876–899, (2013).

[12] Marco Correia, Jorge Cruz, and Joao Leite, 'On the efficient implementation of social abstract argumentation.', in *European Conference on Artificial Intelligence (ECAI)*, pp. 225–230, (2014).

[13] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artificial intelligence*, **77**(2), 321–357, (1995).

[14] Phan Minh Dung and Phan Minh Thang, 'Towards (probabilistic) argumentation for jury-based dispute resolution.', *International Conference on Computational Models of Argument (COMMA)*, **216**, 171–182, (2010).

[15] Paul E Dunne and Michael Wooldridge, 'Complexity of abstract argumentation', in *Argumentation in artificial intelligence*, 85–104, Springer, (2009).

[16] Wolfgang Dvorák, Matti Järvisalo, Johannes Peter Wallner, and Stefan Woltran, 'Cegartix: A sat-based argumentation system', in *Pragmatics of SAT Workshop (POS)*, (2012).

[17] Uwe Egly, Sarah Alice Gaggl, and Stefan Woltran, 'Aspartix: Implementing argumentation frameworks using answer-set programming', in *International Conference on Logic Programming*, pp. 734–738. Springer, (2008).

[18] Jeffrey Fisher, Peter Christen, and Qing Wang, 'Active learning based entity resolution using markov logic', in *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, pp. 338–349. Springer, (2016).

[19] Sarah A Gaggl, Thomas Linsbichler, Marco Maratea, and Stefan Woltran, 'Design and results of the second international competition on computational models of argumentation', *Artificial Intelligence*, 103193, (2019).

[20] KS Gayathri, KS Easwarakumar, and Susan Elias, 'Probabilistic ontology based activity recognition in smart homes using markov logic network', *Knowledge-Based Systems*, **121**, 173–184, (2017).

[21] Anthony Hunter and Matthias Thimm, 'Probabilistic reasoning with abstract argumentation frameworks', *Journal of Artificial Intelligence Research*, **59**, 565–611, (2017).

[22] Tushar Khot, Niranjan Balasubramanian, Eric Gribkoff, Ashish Sabharwal, Peter Clark, and Oren Etzioni, 'Exploring markov logic networks for question answering', in *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 685–694, (2015).

[23] Hiroyuki Kido and Keishi Okamoto, 'A bayesian approach to argument-based reasoning for attack estimation.', in *International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 249–255, (2017).

[24] Hiroyuki Kido and Frank Zenker, 'Argument-based bayesian estimation of attack graphs: A preliminary empirical analysis', in *International Conference on Principles and Practice of Multi-Agent Systems (PRIMA)*, pp. 523–532. Springer, (2017).

[25] Daphne Koller and Nir Friedman, *Probabilistic graphical models: principles and techniques*, MIT press, 2009.

[26] Jean-Marie Lagniez, Emmanuel Lonca, and Jean-Guy Mailly, 'Coquiaas: A constraint-based quick abstract argumentation solver', in *International Conference on Tools with Artificial Intelligence (ICTAI)*, pp. 928–935. IEEE, (2015).

[27] Joao Leite and Joao Martins, 'Social abstract argumentation', in *International Joint Conference on Artificial Intelligence (IJCAI)*, volume 11, pp. 2287–2292, (2011).

[28] Duc Minh Nguyen, Tien Huu Do, Robert Calderbank, and Nikos Deligiannis, 'Fake news detection using deep markov random fields', in *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, pp. 1391–1400, (2019).

[29] Nir Oren and Timothy J. Norman, 'Semantics for evidence-based argumentation', in *International Conference on Computational Models of Argument (COMMA)*, pp. 276–284. IOS Press, (2008).

[30] Sylwia Polberg and Anthony Hunter, 'Empirical evaluation of abstract argumentation: Supporting the need for bipolar and probabilistic approaches', *International Journal of Approximate Reasoning*, **93**, 487–543, (2018).

[31] Nico Potyka, 'Continuous dynamical systems for weighted bipolar argumentation', in *International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pp. 148–157, (2018).

[32] Nico Potyka, Sylwia Polberg, and Anthony Hunter, 'Polynomial-time updates of epistemic states in a fragment of probabilistic epistemic argumentation', in *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU)*, pp. 74–86, (2019).

[33] Antonio Rago, Francesca Toni, Marco Aurisicchio, and Pietro Baroni, 'Discontinuity-free decision support with quantitative argumentation debates.', in *International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pp. 63–73, (2016).

[34] Matthew Richardson and Pedro Domingos, 'Markov logic networks', *Machine learning*, **62**(1-2), 107–136, (2006).

[35] Tjitze Rienstra, Matthias Thimm, Beishui Liao, and Leendert van der Torre, 'Probabilistic abstract argumentation based on scc decomposability', in *International Conference on Principles of Knowledge Representation and Reasoning (KR)*, (2018).

[36] Régis Riveret, Pietro Baroni, Yang Gao, Guido Governatori, Antonino Rotolo, and Giovanni Sartor, 'A labelling framework for probabilistic argumentation', *Annals of Mathematics and Artificial Intelligence*, 1–51, (2018).

[37] Régis Riveret, Dimitrios Korkinof, Moez Draief, and Jeremy Pitt, 'Probabilistic abstract argumentation: an investigation with boltzmann machines', *Argument & Computation*, **6**(2), 178–218, (2015).

[38] Sriram Srinivasan, Behrouz Babaki, Golnoosh Farnadi, and Lise Getoor, 'Lifted hinge-loss markov random fields'. AAAI Conference on Artificial Intelligence (AAAI), (2019).

[39] Harald Steck, 'Markov random fields for collaborative filtering', in *Advances in Neural Information Processing Systems (NIPS)*, pp. 5474–5485, (2019).

[40] Matthias Thimm, Federico Cerutti, and Tjitze Rienstra, 'Probabilistic graded semantics', in *International Conference on Computational Models of Argument (COMMA)*, pp. 369–380, (2018).

[41] Matthias Thimm and Serena Villata, 'The first international competition on computational models of argumentation: Results and analysis', *Artificial Intelligence*, **252**, 267–294, (2017).