Log. Univers. 9 (2015), 345–382

1661-8297/15/030345-38, *published online* April 25, 2015
DOI 10.1007/s11787-015-0120-1

Logica Universalis

Probabilistic Argumentation: An Equational Approach

D. M. Gabbay and O. Rodrigues

Abstract. There is a generic way to add any new feature to a system. It involves (1) identifying the basic units which build up the system and (2) introducing the new feature to each of these basic units. In the case where the system is *argumentation* and the feature is probabilistic we have the following. The basic units are: (a) the nature of the arguments involved; (b) the membership relation in the set S of arguments; (c) the attack relation; and (d) the choice of extensions. Generically to add a new aspect (probabilistic, or fuzzy, or temporal, etc) to an argumentation network $\langle S, R\rangle$ can be done by adding this feature to each component (a–d). This is a brute-force method and may yield a non-intuitive or meaningful result. A better way is to meaningfully translate the object system into another target system which does have the aspect required and then let the target system endow the aspect on the initial system. In our case we translate argumentation into classical propositional logic and get probabilistic argumentation from the translation. Of course what we get depends on how we translate. In fact, in this paper we introduce probabilistic semantics to abstract argumentation theory based on the equational approach to argumentation networks. We then compare our semantics with existing proposals in the literature including the approaches by M. Thimm and by A. Hunter. Our methodology in general is discussed in the conclusion.

Mathematics Subject Classification. Primary 68T27; Secondary 60B99, 68T30.

Keywords. Argumentation, probability theory, numerical methods.

1. Introduction

The objective of this paper is to provide some orientation to underpin probabilistic semantics for abstract argumentation. We feel that a properly developed probabilistic argumentation framework cannot be obtained by simply imposing an arbitrary probability distribution on the components of an argumentation system that does not agree with the dynamic aspects of these networks. We need to find a probability distribution that is compatible with their underlying motivation.

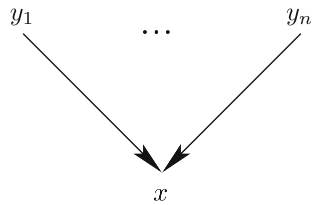

Figure 1. Basic attack formation in an argumentation network

We shall use the methodology of "Logic by Translation", which works as follows: Given a new area for which we want to study certain aspect properties AP, we translate this area to classical logic, study AP in classical logic and then translate back and evaluate what we have obtained.

Let us start by looking at interpretations of an abstract argumentation network $\langle S, R\rangle$, $S \neq \varnothing$, $R \subseteq S \times S$, into logics which already have probabilistic versions. This way we can import the probability aspect from there and it will have a meaning. We begin with translating abstract argumentation frames into classical propositional logic. In the abstract form, the elements of S are just atoms waiting to be instantiated as arguments coming from another application system. R may be defined using the source application system or may represent additional constraints. At any rate, in this abstract form, S is just a set of atoms and all we have about it is R. In translating $\langle S, R\rangle$ into classical propositional logic, we view S as a set of atomic propositions and we use R to generate a classical theory $\Delta_{\langle S,\, R\rangle}$. Consider Fig. 1, which describes the basic attack formation of all the attackers $Att(x) = \{y \in S \mid (y, x) \in R\} = \{y_1, \ldots, y_n\}$ of the node x in a network $\langle S, R\rangle$.

The essential logic translation of the attack on each node x is given by (E1) below, where x, y_i are propositional symbols representing the elements $x, y_i \in S$:

$$x \leftrightarrow \bigwedge_i \neg y_i \tag{E1}$$

So $\langle S, R\rangle$ corresponds to a classical *propositional* theory $\Delta_{\langle S,\, R\rangle} = \{x \leftrightarrow \bigwedge_i \neg y_i \mid x \in S\}$.[1] Note that in classical logic, this theory may be inconsistent and have no models. For example, if S contains a single node x and R is $\{(x, x)\}$, i.e., the network has a single self-attacking node, then the associated theory is $\{x \leftrightarrow \neg x\}$, which has no model. For this reason it is convenient to regard these theories as theories of Kleene three-valued logic, with values in

[1] If there is a logical relationship between the arguments of S that can be captured by formulae, then we can alternatively instantiate $x \longmapsto \varphi_x$, giving $\Delta_{\langle S,\, R\rangle} = \{\varphi_x \leftrightarrow \bigwedge_i \neg\varphi_y \mid x, y \in S\}$.

$\{0, \frac{1}{2}, 1\}$. In this 3-valued semantics, a valuation would satisfy $x \leftrightarrow \neg x$ if and only if it gives the value $\frac{1}{2}$ to x.[2]

If we consider the *equational approach* [5], then we can write

$$x = \bigwedge_i \neg y_i \tag{E2}$$

where (E2) is a numerical equation over the real interval $[0, 1]$, with conjunction and negation interpreted as numerical functions expressing the correspondence of the values of the two sides.

A complete extension of $\langle S, R\rangle$ is a solution to the equations of the form of (E2) when they are viewed as a set of Boolean equations in Kleene's 3-valued logic with values $\{0, \frac{1}{2}, 1\}$, where

$$x = 0 \;\text{ means that }\; x = \mathbf{out} \text{ (at least one attacker } y_i = \mathbf{in}) \tag{1}$$

$$x = 1 \;\text{ means that }\; x = \mathbf{in} \text{ (all attackers } y_i = \mathbf{out}) \tag{2}$$

$$x = \frac{1}{2} \;\text{ means that }\; x = \mathbf{und} \quad \text{(no attacker } y_i = \mathbf{in} \text{ and at least one attacker } y_j = \mathbf{und}) \tag{3}$$

The acceptability semantics above can be re-written in terms of the semantics of Kleene's logic as

$$v(x) = \min\{1 - v(y_i)\}$$

which in equational form can be simplified to

$$x = 1 - \max\{y_i\} \tag{E2*}$$

The reader should note that we actually solve the equations over the unit interval $[0, 1]$ and project onto Kleene's 3-valued logic by letting

$$\begin{array}{lll} x = 0 & \text{mean } x = \mathbf{out} & \text{(at least one attacker } y_i = \mathbf{in}) \\ 0 < x < 1 & \text{mean } x = \mathbf{und} & \text{(no attacker } y_i = \mathbf{in} \text{ and at least} \\ & & \quad \text{one attacker } y_j = \mathbf{und}) \\ x = 1 & \text{mean } x = \mathbf{in} & \text{(all attackers } y_i = \mathbf{out}) \end{array}$$

Now there are probabilistic approaches to two-valued classical logic. The simplest two methods are described in Gabbay's book *Logic for Artificial Intelligence and Information Technology* [4]. Our idea is to bring the probabilistic approach through the above translation into argumentation theory.

Let us start with a description of the probabilistic approaches to classical propositional logic.

Method 1: Syntactic Impose probability $P(q)$ on the atoms q of the language and propagate this probability to arbitrary well-formed formulas (wffs). So if $\varphi(q_1, \ldots, q_m)$ is built up from the atoms $q_1, \ldots, q_m$, we can calculate $P(\varphi)$ if we know $P(q_i)$, $i = 1, \ldots, m$.

[2] In Kleene's logic, one can interpret $\neg$ as complement to 1; $\wedge$ as min; and $\vee$ as max. Thus, if the values of A, B are $v(A), v(B)$, then $v(\neg A) = 1 - v(A)$, $v(A \wedge B) = \min(v(A), v(B))$ and $v(A \vee B) = \max(v(A), v(B))$.

Method 2: Semantic Impose probability on the models of the language of $\{q_1, \ldots, q_m\}$. The totality of models is the space W of all $\{0, 1\}$-vectors in 2^m. We give values $P(\varepsilon)$, for any $\varepsilon \in 2^m$, with the restriction that $\Sigma_{\varepsilon \in 2^m} P(\varepsilon) = 1$. The probability of any wff φ is then

$$P(\varphi) = \Sigma_{\varepsilon \Vdash \varphi} P(\varepsilon) \tag{P1}$$

The motivation for the syntactical Method 1 is that the atoms $\{q_1, \ldots, q_m\}$ are all independent. So for example, the date of birth of a person (p) is independent of whether it is going to rain heavily on that person's 21st birthday (q). However, if we want to hold a birthday party r in the garden on the 21st birthday, then we have that q attacks r.

If, on the other hand, we have:

$a =$ John comes to the party

$b =$ Mary comes to the party

then a and b may be dependent, especially if some relationship exists between John and Mary. We may decide that the probability of $a \wedge b$ is 0, but the probabilities of $\neg a \wedge b$ and of $a \wedge \neg b$ are $\frac{1}{4}$ each and the probability of $\neg a \wedge \neg b$ is $\frac{1}{2}$. Assigning probability in this way depends on the likelihood we attach to a particular situation (model). This is the semantic approach.

Example 1.1 shows that these two methods are orthogonal.

Example 1.1. What can $\Delta_{\langle S, R \rangle}$ mean in classical logic? It is a generalisation of the "Liar's paradox". x attacking itself is like x saying "I am lying": $x = \top$ if and only if $x = \bot$. Figure 1 represents y_i saying x is a lie. $\Delta_{\langle S, R \rangle}$ represents a system of lying accusations: a *community liar paradox*.

Similarly, S can represent people possibly invited to a birthday party. $y \rightarrow x$ means y saying "if I come, x cannot come". So Fig. 1 is saying "invite x if and only if you do not invite any of the y_i".

Suppose we instantiate $x \longmapsto \varphi_x$. Then we must have

$$P(\varphi_x) = P\left(\bigwedge_i \neg \varphi_{y_i}\right).$$

However, there may be also a connection between φ_x and some φ_{y_k}, e.g., $\varphi_x \vdash \varphi_{y_k}$. This will impose further restrictions on $P(\varphi_x)$ and $P(\varphi_{y_k})$, and it may be the case that no such probability function exists.

Remark 1.2. The two approaches are of course, connected. If we are given a probability on each q_i, then we get probability on each $\varepsilon \in 2^m$ by letting

$$P(\varepsilon) = \Pi_{\varepsilon \Vdash q} P(q) \times \Pi_{\varepsilon \Vdash \neg q}(1 - P(q)) \tag{P2}$$

The q_i's are considered independent, so the probability of $\bigwedge_i \pm q_i$ is the product of the probabilities

$$P(\bigwedge_i \pm q_i) = \Pi_i P(\pm q_i)$$

where $P(\neg q_i) = 1 - P(q_i)$ and the probability of $A \vee B$ is

$$P(A \vee B) = P(A) + P(B)$$

when $\Vdash \neg(A \wedge B)$, as is the case with disjuncts in a disjunctive normal form.

So, for example

$$\begin{aligned} P((a \wedge b) \vee (a \wedge \neg b)) &= P(a \wedge b) + P(a \wedge \neg b) \\ &= P(a)P(b) + P(a)(1 - P(b)) \\ &= P(a)(P(b) + 1 - P(b)) \\ &= P(a). \end{aligned}$$

2. The Syntactical Approach (Method 1)

Let us investigate the use of the syntactical approach.

Let $\langle S, R \rangle$ be an argumentation network. In the equational approach, according to the syntactical Method 1, we assign probabilities to all the atoms and are required to solve the Eq. (E3) below for each x, where $Att(x) = \{y_i\}$ and x and all y_i are numbers in $[0, 1]$:

$$P(x) = P\left(\bigwedge_i \neg y_i\right), \tag{E3}$$

Since in Method 1, all atoms are independent, (E3) is equivalent to (E3*):

$$P(x) = \Pi_i(1 - P(y_i)). \tag{E3*}$$

Such equations always have a solution.

Let us check whether this makes sense. Let us try to identify the argument x equationally with its probability, namely we let $P(x) = x$.

If $x = \mathbf{in}$, let $P(x) = 1$

If $x = \mathbf{out}$, let $P(x) = 0$.

If $x = \mathbf{und}$, let $0 < P(x) < 1$

to be determined by the solution to the equations.

Equation (E3*) becomes, under $P(x) = x$, the following:

$$x = \Pi(1 - y_i) \text{ for } x \in S. \tag{E4}$$

This is the $\mathrm{Eq_{inv}}$ equation in the equational approach (see [5]).

The following definition will be useful in the interpretation of values from $[0, 1]$ and their counterparts in Caminada's labelling functions.

Definition 2.1. A valuation function $\mathbf{f}$ can be mapped into a labelling function $\lambda(\mathbf{f})$ as follows.

$\mathbf{f}(x) = 1$	$\rightarrow$	$\lambda(\mathbf{f})(x) = \mathbf{in}$
$\mathbf{f}(x) = 0$	$\rightarrow$	$\lambda(\mathbf{f})(x) = \mathbf{out}$
$\mathbf{f}(x) \in (0, 1)$	$\rightarrow$	$\lambda(\mathbf{f})(x) = \mathbf{und}$

What do we know about $\mathrm{Eq_{inv}}$? We quote the following from [5].

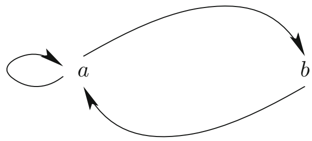

Figure 2. A sample argumentation network having a complete extension that cannot be found via Equations (E4)

Theorem 2.2. *Let* $\mathbf{f}$ *be a solution to Eq.* (E4)*. Then* $\lambda(\mathbf{f})$ *defined according to Definition* 2.1 *is a legal Caminada labelling* (*see* [1]) *and leads to a complete extension.*

Theorem 2.3. *Let* λ_0 *be a legal Caminada labelling leading to a preferred extension. Then there exists a solution* f_0*, such that* $\lambda_0 = \lambda(f_0)$*.*

Remark 2.4. There are (complete) extensions λ' such that there does not exist an f' with $\lambda' = \lambda(f')$.

For example, in Fig. 2, the extension $a = b = \mathbf{und}$ cannot be obtained by any f. Only $b = \mathbf{in}$, $a = \mathbf{out}$ can be obtained as a solution to Eq. (E4).[3]

Example 2.5. Let $\langle S, R\rangle$ be given and let λ be a complete extension which is not preferred! The reason that λ is not preferred, is that we have by definition, a λ_1 extending λ, which gives more $\{\mathbf{in}, \mathbf{out}\}$ values to points z, for which λ gives the value **und**. Therefore, we can prevent the existence of such an extension λ_1, if we force such points z to be undecided. This we do by attacking such points z by a new self-attacking point u. The construction is therefore as follows. We are given $\langle S, R\rangle$ and a complete extension λ, which is not preferred. We now construct a new $\langle S', R'\rangle$ which is dependent on λ. Consider $\langle S', R'\rangle$ where $S' = S \cup \{u\}$, where $u \notin S$, is a new point. Let R' be

$$R' = R \cup \{(u, u)\} \cup \{(u, v) \mid v \in S \text{ and } \lambda(v) = \mathbf{und}\}.$$

Then $\lambda' = \lambda \cup \{(u, \mathbf{und})\}$ is a preferred extension of $\langle S', R'\rangle$ and can therefore be obtained from a function f' using the Eq. (E4).

Let us see what the construction above does to our example in Fig. 2, and let us look at the extension $\lambda(a) = \lambda(b) = \mathbf{und}$.

Consider the network in Fig. 3. Its Eq. (E4) are:

1. $u = 1 - u$
2. $a = (1 - u)(1 - a)(1 - b)$
3. $b = (1 - u)(1 - a)$

From (1) we get $u = \frac{1}{2}$. So we have:

---

[3] The equations are

1. $a = (1 - a) \times (1 - b)$
2. $b = 1 - a$.

From the above two equations we get

3. $a = (1 - a) \times a$

The only possibility is $a = 0$.

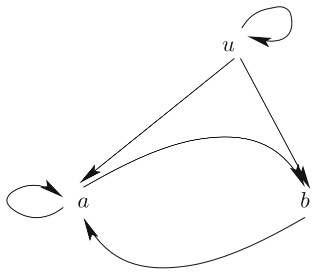

FIGURE 3. The network of Fig. 2 with an extra undecided node u attacking all nodes

2. $a = \frac{1}{2}(1-a)(1-b)$
3. $b = \frac{1}{2}(1-a)$

$$\begin{aligned} 1-b &= 1-\tfrac{1}{2}(1-a) \\ &= \tfrac{2-1+a}{2} \\ &= \tfrac{1+a}{2} \end{aligned}$$

therefore substituting in (1) we get

$$\begin{aligned} a &= \tfrac{1}{2}(1-a)(\tfrac{1+a}{2}) \\ &= \tfrac{1}{4}(1-a^2) \\ &\quad 4a + a^2 - 1 = 0 \\ &\quad (a+2)^2 - 4 - 1 = 0 \\ &\quad (a+2)^2 = 5 \\ a &= \sqrt{5} - 2 \approx 0.236 \\ b &= \tfrac{1}{2}(1-a) \\ &= \tfrac{1}{2}(1-\sqrt{5}+2) \\ &= \tfrac{3-\sqrt{5}}{2} \approx 0.382. \end{aligned}$$

The extension of the network is $a = b =$ **und**.

Summary of the results so far for the syntactical probabilistic method. Given an argumentation network $\langle S, R\rangle$, we can find all Method 1 complete probabilistic extensions for it by solving all $\mathrm{Eq}_{\mathrm{inv}}$ equations. Such complete probabilistic extensions will also be complete extensions in the traditional sense (i.e., Dung's), which will also include all preferred extensions (Theorems 2.2 and 2.3).[4]

[4] Note that in traditional Dung semantics a preferred extension E is maximal in the sense that there is no extension E' such that

1. If x is considered **in** (resp. **out**) by E then x is also considered **in** (resp. **out**) by E'.
2. There exists at least one node considered**in** (resp. **out**) by E' and considered **und** by E.

The above definition holds for numerical or probabilistic semantics, where the value 1 (resp. 0) is understood as **in** (resp. **out**) and values in $(0, 1)$ are understood as **und**.

However, not all complete extensions can be obtained in this manner (i.e., by Method 1, see Remark 2.4 and compare with Example 3.6).

We can, nevertheless, for any complete extension E which cannot be obtained by Method 1, obtain it from the solutions of the equations generated for a larger network $\langle S', R' \rangle$ as shown in Example 2.5.

We shall say more about this in a later section.

Remark 2.6. Evaluation of the results so far for the syntactical probabilistic method.

1. We discovered a formal mathematical connection between the syntactical probabilistic approach (Method 1) and the Equational Eq_{inv} approach. Is this just a formal similarity or is there also a conceptual connection?

   The traditional view of an abstract argumentation frame $\langle S, R \rangle$, is that the arguments are abstract, some of them abstractly attack each other. We do not know the reason, but we seek complete extensions of arguments that can co-exist (i.e., being attack-free), and that protect themselves. The equational approach is an equational way of finding such extensions. Each solution $\mathbf{f}$ to the equations give rise to a complete extension. The numbers we get from such solutions $\mathbf{f}$ of the equational approach can be interpreted as giving the degree of being in the complete extension (associated with $\mathbf{f}$) or being out of it.

   Due to the mathematical similarity with the probability approach, these numbers are now interpreted as probabilities.

   To what extent is this justified? Can we do this at all?

   Let us recall the syntactical probabilistic method. We start with an abstract argumentation framework $\langle S, R \rangle$ and add the probability $P(x)$ for each $x \in S$. We can interpret $P(x)$ as the probability that x "is a player" to be considered (this is a vague statement which could mean anything but is sufficient for our purpose). The problem is how do we take into account the attack relation? Our choice was to require Eq. (E3). It is this choice that allowed the connection between the syntactical probabilistic approach and the Equational approach with Eq_{inv}.

   So our syntactical probabilistic approach should work as follows.

   Let P be the independent probability on each $x \in S$. This is an arbitrary number in $[0, 1]$. Such a P cannot be used for calculating extensions because it does not take into consideration the attack relation R. So modify P to a P' which does respect R via Eq. (E3).

   How do we modify P to find P'?

   Well, we can use a numerical iteration method. The details are not important here, the importance is in the idea, which can be applied to the traditional notion of extensions as well. Given $\langle S, R \rangle$ and an arbitrary desired assignment E of elements that are **in** (and consequently also determining elements that are **out**) for S, this E may not be legitimate in taking into account R, so we need to modify it to get the best proper extension E' nearest to E (cf. [2,6]).

So our syntactical probabilistic approach yielding a P satisfying Eq. (E3) can be interpreted as $\mathrm{Eq}_{\mathrm{inv}}$ extensions obtained from initial values which are probabilities (as opposed to, say, initial values being a result of voting) corrected via iteration procedures using R.

Alternatively, we can look at the $\mathrm{Eq}_{\mathrm{inv}}$ equations as a mathematical means of finding all those syntactical probabilities P which respect the attack relation R [via Eq. (E3)].

Or we can see the solutions of the $\mathrm{Eq}_{\mathrm{inv}}$ as giving probabilities for being included or excluded in the complete extension defined by these solutions (as opposed to the interpretation of the degree of being **in** or **out**).

2. The discussion in item 1. above hinged upon the choice we made to take account of R by respecting Eq. (E3). There are other alternatives for taking R into account. We can give direct, well-motivated definitions of how to propagate probabilities along attack arrows. This is similar to the well-known problem of how to propagate probabilities along proofs (provability support arrows, or modus ponens, etc). Such an analysis is required anyway for instantiated networks, for example in ASPIC+ style [10]). We shall deal with this in a subsequent paper.

3. The Semantical Approach (Method 2)

Let us now check what can be obtained if we use Method 2, i.e., giving probability to the models of the language. In this case the equation (for $\{y_i\} = Att(x)$) (E3) $P(x) = P(\bigwedge_i \neg y_i)$ still holds, but the $\neg y_i$ are not independent. So we cannot write Eq. (E3*) for them and get $\mathrm{Eq}_{\mathrm{inv}}$. Instead we need to use the schema $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$. We begin with a key lemma, which will enable us to compare later with the work of M. Thimm, see [13].

Lemma 3.1. *Let $\langle S, R\rangle$ be a network and let P be a probability measure on the space W of all models of the language whose set of atoms is S. For $x \in S$, let the following hold*

$$P(x) = P\left(\bigwedge_{i=1}^{n} \neg y_i\right)$$

where $Att(x) = \{y_1, \ldots, y_n\}$.

Then we have

1. $P(x) \leq P(\neg y_i), 1 \leq i \leq n$
2. $P(x) \geq 1 - \Sigma_{i=1}^{n} P(y_i)$

Proof. By induction on n.

1. If $x = \neg y$ then $P(x) = 1 - P(y)$ and the above holds.
2. Assume the above holds for m, show for $m+1$. Let $z = \bigvee_{i=1}^{m} y_i, y = y_{m+1}$. Then $x = \neg z \wedge \neg y$.

   We have by the induction hypothesis

   - $P(\neg z) \leq P(\neg y_i), i = 1, \ldots, m$

- $P(\neg z) \geq 1 - \Sigma_{i=1}^{m} P(y_i)$

Consider now:

$$\begin{aligned} P(\neg z \wedge \neg y) &= 1 - P(y \vee z) \\ &= 1 - (P(y) + P(z) - P(y \wedge z)) \\ &= 1 - P(y) - P(z) + P(y \wedge z) \\ &= 1 - P(y) - (P(z) - P(y \wedge z)) \end{aligned}$$

But $P(A \wedge B) \leq P(B)$ is always true.

So

$$P(\neg z \wedge \neg y) \leq 1 - P(y) = P(\neg y)$$

On the other hand, by our assumption

$$1 - P(z) = P(\neg z) \geq 1 - \Sigma_{i=1}^{m} P(y_i)$$

So

$$\begin{aligned} P(\neg z \wedge \neg y) &= 1 - P(y) - P(z) + P(y \wedge z) \\ &\quad (1 - P(z)) - P(y) + P(y \wedge z) \\ &\geq 1 - \Sigma P(y_i) - P(y) + P(y \wedge z) \\ &\geq 1 - \Sigma_{i=1}^{m+1} P(y_i) \end{aligned}$$

□

Remark 3.2. The converse of Lemma 3.1 does not hold, as we shall see in Example 3.5 below.

Let us look at some examples illustrating the use of Method 2.

Example 3.3. Consider the network in Fig. 4. This figure is taken from Thimm's "A probabilistic semantics for abstract argumentation" [13, Figure 1]. We include it here for two reasons:

1. To illustrate or probabilistic semantic approach.
2. To use it later to compare our work with Thimm's approach.

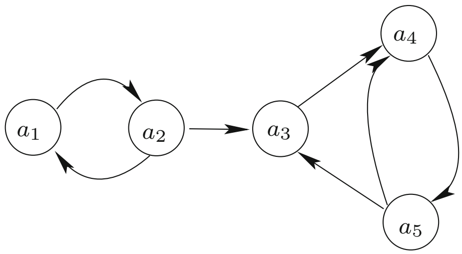

FIGURE 4. Figure 1 of "A probabilistic semantics for abstract argumentation" [13]

Let us apply Method 2 to it and assign probabilities to the models of the propositional language with the atoms $\{a_1, a_2, a_3, a_4, a_5\}$. We assign P as follows.

$$\begin{aligned} P(a_1 \wedge \neg a_2 \wedge a_3 \wedge \neg a_4 \wedge a_5) &= 0.3 \\ P(a_1 \wedge \neg a_2 \wedge \neg a_3 \wedge a_4 \wedge \neg a_5) &= 0.45 \\ P(\neg a_1 \wedge a_2 \wedge \neg a_3 \wedge \neg a_4 \wedge a_5) &= 0.1 \\ P(\neg a_1 \wedge a_2 \wedge \neg a_3 \wedge a_4 \wedge \neg a_5) &= 0.15 \\ P(\text{any other conjunctive model}) &= 0. \end{aligned}$$

Let us compute $P(a_i)$, for $i = 1, \dots, 5$.
We have

$$P(X) = \sum_{\varepsilon \Vdash X} P(\varepsilon).$$

We get

$$\begin{aligned} P(a_1) &= 0.3 + 0.45 = 0.75 \\ P(a_2) &= 0.1 + 0.15 = 0.25 \\ P(a_3) &= 0.3 \\ P(a_4) &= 0.45 + 0.15 = 0.6 \\ P(a_5) &= 0.3 + 0.1 = 0.4. \end{aligned}$$

To be a legitimate probabilistic model P must satisfy Eq. (E3) relating to the attack relation of Fig. 4. Namely we must have

$$P(X) = P\left(\bigwedge_{Y \in Att(X)} \neg Y\right) \tag{E3}$$

Therefore

$$\begin{aligned} P(a_1) &= P(\neg a_2) \\ P(a_2) &= P(\neg a_1) \\ P(a_3) &= P(\neg a_2 \wedge \neg a_5) \\ P(a_4) &= P(\neg a_3 \wedge \neg a_5) \\ P(a_5) &= P(\neg a_4) \end{aligned}$$

Let us calculate the P in the right hand side of the above equations.

$$P(\neg a_2) = 1 - 0.25 = 0.75$$

$$P(\neg a_1) = 1 - 0.75 = 0.25$$

$$P(\neg a_2 \wedge \neg a_5) = 0.45$$

$$P(\neg a_3 \wedge \neg a_5) = 0.45 + 0.15 = 0.6$$

$$P(\neg a_4) = 0.4$$

We see that

$$P(a_3) = 0.3 \neq P(\neg a_2 \wedge \neg a_5) = 0.45.$$

Therefore this distribution P is not legitimate according to our Method 2. It does not satisfy Eq. (E3) because

$$P(a_3) \neq P(\neg a_2 \wedge \neg a_5)$$

Therefore Lemma 3.1 does not apply and indeed, condition (2) of Lemma 3.1 does not hold for a_3. We have $P(a_3) = 0.3$ but $1 - P(a_2) - P(a_5) = 0.35$.

Example 3.4. Let us look at Fig. 5. This is also taken from Thimm's paper [13, Figure 2]. It shall be used later to compare our methods with Thimm's.

1. *We use Method 2.* Consider the following probability distribution on models

$$P(a_1 \wedge \neg a_2 \wedge \neg a_3) = 0.5$$

$$P(a_1 \wedge \neg a_2 \wedge a_3) = 0$$

$$P(a_1 \wedge a_2 \wedge \neg a_3) = 0$$

$$P(a_1 \wedge a_2 \wedge a_3) = 0$$

$$P(\neg a_1 \wedge a_2 \wedge a_3) = 0$$

$$P(\neg a_1 \wedge a_2 \wedge \neg a_3) = 0.5$$

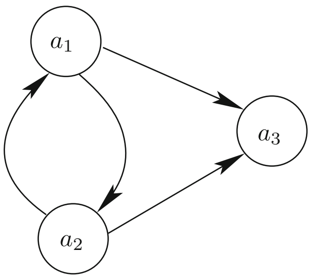

FIGURE 5. Figure 2 of "A probabilistic semantics for abstract argumentation" [13]

$$P(\neg a_1 \wedge \neg a_2 \wedge a_3) = 0$$
$$P(\neg a_1 \wedge \neg a_2 \wedge \neg a_3) = 0.$$

In this model we get

$$P(a_1) = 0.5$$
$$P(a_2) = 0.5$$
$$P(a_3) = 0$$

Let us check whether this probability distribution satisfies Eq. (E3), namely

$$P(X) = P\left(\bigwedge_{Y \in Att(X)} \neg Y\right) \tag{E3}$$

We need to have

$$P(a_1) = P(\neg a_2)$$
$$P(a_2) = P(\neg a_1)$$
$$P(a_3) = P(\neg a_2 \wedge \neg a_2)$$

Indeed

$$P(\neg a_1) = 1 - P(a_1) = 0.5$$
$$P(\neg a_2) = 1 - P(a_2) = 0.5$$
$$P(\neg a_1 \wedge \neg a_2) = 0.$$

Thus we have a legitimate model.

2. *We use Method 1.* Let us use $\mathrm{Eq}_{\mathrm{inv}}$ on this figure, namely we try and solve the equations

$$a_1 = 1 - a_2$$
$$a_2 = 1 - a_1$$
$$a_3 = (1 - a_1)(1 - a_2)$$

Let us use a parameter $0 \leq x \leq 1$ and let

$$a_1 = x,$$
$$a_2 = 1 - x,$$
$$a_3 = x(1 - x)$$

The probabilities we get with parameter x as well as for $x = 0.5$ are given below.

$$\begin{aligned}
P(a_1 \wedge a_2 \wedge a_3) &= x^2(1-x)^2 &&= \tfrac{1}{16}\\
P(a_1 \wedge a_2 \wedge \neg a_3) &= x(1-x)(1-x(1-x)) &&= \tfrac{3}{16}\\
P(a_1 \wedge \neg a_2 \wedge a_3) &= x^3(1-x) &&= \tfrac{1}{16}\\
P(a_1 \wedge \neg a_2 \wedge \neg a_3) &= x^2(1-x(1-x)) &&= \tfrac{3}{16}\\
P(\neg a_1 \wedge a_2 \wedge a_3) &= x(1-x)^3 &&= \tfrac{1}{16}\\
P(\neg a_1 \wedge a_2 \wedge \neg a_3) &= (1-x)^2(1-x(1-x)) &&= \tfrac{3}{16}\\
P(\neg a_1 \wedge \neg a_2 \wedge a_3) &= x^2(1-x)^2 &&= \tfrac{1}{16}\\
P(\neg a_1 \wedge \neg a_2 \wedge \neg a_3) &= x(1-x)(1-x(1-x)) &&= \tfrac{3}{16}
\end{aligned}$$

If we choose $x = 0.5$ we get $P(a_1) = P(a_2) = 0.5$ and $P(a_3) = \frac{1}{4}$.

Example 3.5. This example shows that the converse of Lemma 3.1 does not hold. Consider the network in Fig. 6.

Any legitimate probability assigned to models would be required to satisfy the following

$$P(a) = P(\neg a \wedge \neg b)$$

$$P(b) = P(\neg a \wedge \neg b)$$

Case 1. Try the following probability P_1.

$$P_1(a \wedge b) = P_1(a \wedge \neg b) = P_1(\neg a \wedge b) = P_1(\neg a \wedge \neg b) = 0.25.$$

Therefore

$$P_1(a) = 0.5$$

$$P_1(b) = 0.5$$

Note that we also have

$$P_1(a) = \frac{1}{2} \leq 1 - P_1(b) = \frac{1}{2}$$

$$P_1(a) = \frac{1}{2} \leq 1 - P_1(a) = \frac{1}{2}.$$

Similarly for $P_1(b)$ by symmetry.

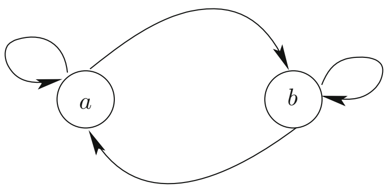

Figure 6. Network for Example 3.5

Also

$$P_1(a) = \frac{1}{2} \geq 1 - P_1(a) - P_1(b) = 1 - \frac{1}{2} - \frac{1}{2} = 0.$$

Thus the conditions of the conclusions of Lemma 3.1 hold. However the assumptions of Lemma 3.1 do not hold, because

$$P_1(a) = \frac{1}{2} \neq P_1(\neg a \wedge \neg b) = \frac{1}{4}.$$

Case 2. Let us check whether we can find a probability P_2 which is indeed acceptable to Method 2. Let us try with variables y, z and create equations and solve them:

$$P_2(a \wedge b) = y$$
$$P_2(\neg a \wedge b) = z.$$

Therefore $P_2(b) = y + z$.

$$P_2(\neg a \wedge \neg b) = y + z$$

and what is left is

$$P_2(a \wedge \neg b) = 1 - 2y - 2z$$

but we must also have

$$P_2(a) = P_2(\neg a \wedge \neg b)$$

and hence we must have

$$P_2(a) = 1 - 2y - 2z + y = P_2(\neg a \wedge \neg b) = y + z.$$

So we get the equation

$$1 - 2y - 3z = 0$$
$$2y + 3z = 1$$
$$y = \tfrac{(1-3z)}{2}$$

Since $0 \leq y, z \leq 1$ so z must be less than $\frac{1}{3}$.
Let us choose $z = 0.2$ and so $y = 0.2$.
We get, for example

$$P_2(a \wedge b) = 0.2$$
$$P_2(\neg a \wedge b) = 0.2$$
$$P_2(\neg a \wedge \neg b) = 0.4$$
$$P_2(a \wedge \neg b) = 0.2$$

We could also have chosen $z = \frac{1}{3}$ and $y = 0$. This would give P_3, where

$$P_3(a \wedge b) = 0$$

$$P_3(\neg a \wedge b) = \frac{1}{3}$$

$$P_3(\neg a \wedge \neg b) = \frac{1}{3}$$

$$P_3(a \wedge \neg b) = \frac{1}{3}$$

So we get

$$\begin{aligned} P_3(b) = P(a) &= \tfrac{1}{3} \\ P_3(\neg a \wedge \neg b) &= \tfrac{1}{3}. \end{aligned}$$

Example 3.6. Consider the network of Fig. 2. Let us try to find a probabilistic semantics for it according to Method 2. Assume we have

$$P(a \wedge b) = x_1$$

$$P(a \wedge \neg b) = x_2$$

$$P(\neg a \wedge b) = x_3$$

$$P(\neg a \wedge \neg b) = 1 - x_1 - x_2 - x_3.$$

We need to satisfy

$$P(a) = P(\neg a \wedge \neg b)$$

$$P(b) = P(\neg a)$$

This means we need to solve the following equations.

1. $x_1 + x_2 = 1 - x_1 - x_2 - x_3$
2. $x_1 + x_3 = 1 - x_1 - x_2$.

By adding $x_1 + x_2$ to both sides (1) can be written as

$$2(x_1 + x_2) = 1 - x_3,$$

and by swapping x_3 to the right and $-x_1 - x_2$ to the left (2) can be written as

$$2x_1 + x_2 = 1 - x_3.$$

Thus we get

3. $2x_1 + x_2 = 2x_1 + 2x_2$.

Therefore $x_2 = 0$.

There remains, therefore

4. $2x_1 = 1 - x_3$.

We can choose values for x_3.

Sample choice 1. $x_3 = 1$, so $x_1 = 0$.

We get $P_1(a \wedge b) = P(a \wedge \neg b) = P_1(\neg a \wedge \neg b) = 0$ and $P_1(\neg a \wedge b) = 1$.

This yields $P(a) = 0, P(b) = 1$. This is also the $\mathrm{Eq}_{\mathrm{inv}}$ solution to

$$b = 1 - a$$

$$a = (1-a)(1-b)$$

Sample choice 2. $x_3 = \frac{1}{2}$. So $x_1 = \frac{1}{4}$ and the probabilities are

$$P_2(a \wedge b) = \frac{1}{4}$$

$$P_2(a \wedge \neg b) = 0$$

$$P_2(\neg a \wedge b) = \frac{1}{2}$$

$$P_2(\neg a \wedge \neg b) = \frac{1}{4}.$$

P_2 is a Method 2 probability, which cannot be given by Method 1.

Sample choice 3. $x_3 = 0$. Then $x_1 = \frac{1}{2}$. We get

$$P_3(a \wedge b) = \frac{1}{2}$$

$$P_3(a \wedge \neg b) = 0$$

$$P_3(\neg a \wedge b) = 0$$

$$P_3(\neg a \wedge \neg b) = \frac{1}{2}.$$

Therefore $P_3(a) = P_3(b) = \frac{1}{2}$.

Lemma 3.7. *Let $\langle S, R\rangle$ be a network and let P be a semantic probability (Method 2) for $\langle S, R\rangle$. Let $x \in S$ and let $\{y_i\} = Att(x)$. Then*

1. *If for some $y_i, P(y_i) = 1$ then $P(x) = 0$.*
2. *If for all $y_i, P(y_i) = 0$ then $P(x) = 1$.*

Proof. Let us use Fig. 1 where $\{y_i\} = Att(x)$.

Case 1. Assume that $P(y_1) = 1$. We need to show that $P(x) = 0$. We have:

$$P(x) = P\left(\bigwedge_i \neg y_i\right) \tag{E3}$$

We also have

$$P(A) = \sum_{\varepsilon \Vdash A} P(\varepsilon) \tag{P1}$$

Therefore

$$P(x) = \sum_{\varepsilon \Vdash \bigwedge_i \neg y_i} P(\varepsilon)$$

$$P(x) = \sum_{\varepsilon \Vdash \neg y_1 \wedge \bigwedge_{j=1}^{n} \neg y_j} P(\varepsilon) \tag{i}$$

but

$$P(y_1) = \sum_{\varepsilon \Vdash y_1} P(\varepsilon) = 1$$

Therefore we have

$$\sum_{\varepsilon \Vdash \neg y_1} P(\varepsilon) = 0 \tag{ii}$$

From (i) and (ii) we get that $P(x) = 0$.

Case 2. We assume that for all i, $P(y_i) = 0$ and we need to show that $P(x) = 1$.

We have

$$P(x) = P\left(\bigwedge \neg y_i\right)$$

$$P(x) = 1 - P\left(\bigvee y_i\right) \tag{iii}$$

We also have

$$P\left(\bigvee y_i\right) = \Sigma_{\varepsilon \Vdash \bigvee y_i} P(\varepsilon) \tag{iv}$$

Suppose for some ε' such that $\varepsilon' \Vdash \bigvee y_i$ we have $P(\varepsilon') > 0$. But $\varepsilon' \Vdash \bigvee y_i$ implies $\varepsilon' \Vdash y_i$, for some i.

Say $i = 1$. Thus we have $\varepsilon' \Vdash y_1$ and $P(y_1) = 0$ and $P(\varepsilon') > 0$. This is impossible since

$$P(y_1) = \sum_{\varepsilon \Vdash y_1} P(\varepsilon) \tag{v}$$

Therefore for all ε such that $\varepsilon \Vdash \bigvee y_i$ we have that $P(\varepsilon) = 0$. Therefore by (iii) and (iv) we get

$$P(x) = 1.$$

□

Remark 3.8. Let $\langle S, R\rangle$ be a network and let P be a semantic probability for $\langle S, R\rangle$ (Method 2).

Let λ be defined as follows, for $x \in S$.

$$\lambda(x) = \begin{cases} \textbf{in}, & \text{if } P(x) = 1 \\ \textbf{out}, & \text{if } P(x) = 0 \\ \textbf{und}, & \text{if } 0 < P(x) < 1 \end{cases}$$

The perceptive reader might expect us to say that λ is a legitimate Caminada labelling, especially in view of Lemma 3.7. This is not the case as Example 3.9 shows.

Example 3.9. This example shows that in the probabilistic semantics it is possible to have $P(x) = 0$, while for all attackers y of x we have $0 < P(y) < 1$. Thus the nature of the probabilistic attack is different from the traditional Dung one. If $Att(x)$ is the set of all attackers of x and $P(\bigvee_{y \in Att(x)} y) = 1$, then, and only then $P(x) = 0$.

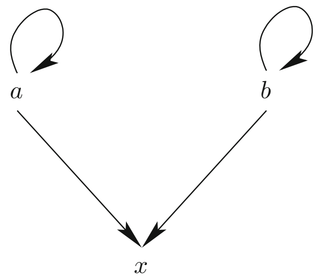

Figure 7. A network with Method 1 and Method 2 probabilities

Thus the attackers of x can attack with joint probability.
The example we give is the network of Fig. 7.
This has a Method 1 probability of $P_1(a) = \frac{1}{2}, P_1(b) = \frac{1}{4}$ and $P_1(x) = \frac{1}{4}$.
Thus for any model $\mathbf{m} = \pm a \wedge \pm b \wedge x$ we have

$$P_1(\mathbf{m}) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{4} = \frac{1}{16}$$

and for any model

$$\mathbf{m}' = \pm a \wedge \pm \wedge \neg x$$

we have

$$P_1(\mathbf{m}) = \frac{1}{2} \times \frac{1}{2} \times \frac{3}{4} = \frac{3}{16}.$$

Figure 7 also has a Method 2 probability model. We can have

$$P_2(a) = P_2(b) = \frac{1}{2}$$

$$P_2(x) = 0.$$

Let us check what values to give to the models. The models are:

$$m_1 = x \wedge a \wedge b$$

$$m_2 = x \wedge a \wedge \neg b$$

$$m_3 = x \wedge \neg a \wedge b$$

$$m_4 = x \wedge \neg a \wedge \neg b$$

$$m_5 = \neg x \wedge a \wedge b$$

$$m_6 = \neg x \wedge a \wedge \neg b$$

$$m_7 = \neg x \wedge \neg a \wedge b$$

$$m_8 = \neg x \wedge \neg a \wedge \neg b.$$

We want the following equations to be satisfied.

1. $P_2(x) = 0$. This means we need to let
$$P_2(m_i) = 0, i = 1, \ldots, 4.$$
2. $P_2(a) = \frac{1}{2}$. This means we need to let
$$P_2(m_5) + P_2(m_6) = \tfrac{1}{2}$$
$$P_2(m_7) + P_2(m_8) = \tfrac{1}{2}.$$
3. $P_2(b) = \frac{1}{2}$, yields the equations
$$P_2(m_5) + P_2(m_7) = \tfrac{1}{2}$$
$$P_2(m_6) + P_2(m_8) = \tfrac{1}{2}.$$
4. We also need to have the equation
$$0 = P_2(x) = P_2(\neg a \wedge \neg b)$$
Therefore $P_2(m_8) = 0$.

   We thus have the following equations left
   (a) $P_2(m_5) + P_2(m_6) = \frac{1}{2}$
   (b) $P_2(m_7) = \frac{1}{2}$
   (c) $P_2(m_5) + P_2(m_7) = \frac{1}{2}$
   (d) $P_2(m_6) = \frac{1}{2}$.

   From (b) and (c) we get $P_2(m_5) = 0$. This makes $P_2(m_6) = \frac{1}{2}$. Thus we get the following solution:
$$P_2(m_i) = 0, \text{ for } i = 1, 2, 3, 4, 5, 8$$
$$P_2(m_6) = P_2(m_7) = \tfrac{1}{2}.$$
Note that the Eq. (E3) hold for P_1 and P_2:
$$P(a) = 1 - P(a)$$
$$P(b) = 1 - P(b)$$
hold of both P_1 and P_2. As for $P(x) = P(\neg a \wedge \neg b)$ we check
$$\tfrac{1}{4} = P_1(x) = P_1(\neg a \wedge \neg b)$$
$$= P_1(\neg a) \times P_1(\neg b) = \tfrac{1}{4}.$$
For P_2 we have
$$0 = P_2(x) = P_2(\neg a \wedge \neg b)$$
$$P_2(\neg(a \vee b) = 1 - P_2(a \vee b)$$
$$P_2(a \vee b) = P_2(m_1) + P_2(m_2) + P_2(m_3) + P_2(m_5) + P_2(m_6)$$
$$+P_2(m_7) = 0 + 0 + 0 + \tfrac{1}{2} + \tfrac{1}{2} = 1.$$
Thus $P_2(\neg a \wedge \neg b) = 0$.

So P_1 and P_2 are legitimate probabilities on Fig. 7. P_1 is a Method 1 probability and P_2 is a Method 2 probability.

Definition 3.10. We now define the Gabbay–Rodrigues Probabilistic Labelling Π on a network $\langle S, R\rangle$. Π is a $\{\mathbf{in}, \mathbf{out}, \mathbf{und}\}$-labelling satisfying the following.

There exists a semantic probability P on $\langle S, R\rangle$ such that for all $x \in S$

1. $\Pi(x) = \mathbf{in}$, if $P(\bigvee Att(x)) = 0$
2. $\Pi(x) = \mathbf{out}$, if $P(\bigvee Att(x)) = 1$
3. $\Pi(x) = \mathbf{und}$, if $0 < P(\bigvee Att(x)) < 1$

Example 3.11. This example is due to M. Thimm, oral communication, 24th October 2014. Consider Fig. 8.

This figure contains Fig. 7 and its mirror image. We saw that in Fig. 7 (as well as in this Fig. 8) any probability on the figures must yield

$$P(a) = P(b) = \frac{1}{2}.$$

Figure 7 allowed for two possibilities for x. $P_1(x) = \frac{1}{4}$ and $P_2(x) = 0$. Let us try P for our Fig. 8 with

$$P(x_1) = \frac{1}{4} \text{ and } P(x_2) = 0.$$

This is not possible because we must have

$$P(x_i) = P(\neg a \wedge \neg b).$$

So $P(x_1)$ must be equal to $P(x_2)$.

This example will show in the comparison with the literature section that our probability semantics is different from that of M. Thimm in [13].

See also Example 3.5.

Theorem 3.12. *Let $\langle S, R\rangle$ be a network and let λ be a legitimate Caminada labelling on S, giving rise to a complete extension. Then there exists a probability P_λ on the models (Method 2 probabilistic semantics) such that for all $x \in S$:*

- $P_\lambda(x) = 1$, *if* $\lambda(x) = \mathbf{in}$
- $P_\lambda(x) = 0$, *if* $\lambda(x) = \mathbf{out}$
- $P_\lambda(x) = \frac{1}{2}$, *if* $\lambda(x) = \mathbf{und}$.

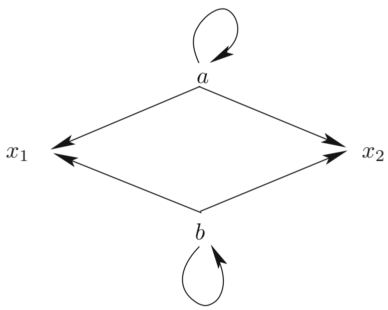

FIGURE 8. Mirrored network of Fig. 7

Proof. (We use an idea from M. Thimm [13])

Let $S = \{s_1, \ldots, s_k\}$. Then when we regard the elements of S as atomic propositions in classical propositional logic, there are 2^k models based on S. Each of these models gives values 0 (false) or 1 (true) to each atomic proposition. Each such a model can be represented by a conjunction of the form $\alpha = \bigwedge_i \pm s_i$. α represents the model which gives value 1 to s_i if $+s_i$ appears in α and gives value 0 to s_i if $-s_i$ appears in α. Given a model we can construct the respective α for it. Let

$$\alpha_1 = \bigwedge_{\lambda(s)=\mathbf{in}} s; \quad \alpha_0 = \bigwedge_{\lambda(s)=\mathbf{out}} \neg s; \quad \alpha_{\frac{1}{2}} = \bigwedge_{\lambda(s)=\mathbf{und}} s; \text{ and } \beta_{\frac{1}{2}} = \bigwedge_{\lambda(s)=\mathbf{und}} \neg s.$$

We now define a Method 2 probability P_λ on the models.

1. $P_\lambda(\alpha_1 \wedge \alpha_0 \wedge \alpha_{\frac{1}{2}}) = \frac{1}{2}$
2. $P_\lambda(\alpha_1 \wedge \alpha_0 \wedge \beta_{\frac{1}{2}}) = \frac{1}{2}$
3. $P_\lambda(m) = 0$, for any other model, m different from the above.

Clearly P_λ is a probability. We examine its properties

(i) Let x be such that $\lambda(x) = \mathbf{in}$.

Then

$$P_\lambda(x) = \sum_{m \Vdash x} P_\lambda(m).$$

Only (1) and (2) can contribute to $P_\lambda(x)$, so the value is 1.

(ii) Let $\lambda(x) = \mathbf{out}$.

The only two models that can contribute to $P_\lambda(x)$ are in (1) and (2) above, but they prove $\neg x$. So $P_\lambda(x) = 0$.

(iii) Let $P_\lambda(x) = \mathbf{und}$.

Then clearly $P_\lambda(x)$ gets a contribution from (1) only. We get $P_\lambda(x) = \frac{1}{2}$.

We now need to verify that P_λ actually satisfies the equations of (E3).

Let $x \in S$ and let y_i be its attackers. We want to show that

$$P_\lambda(x) = P_\lambda\left(\bigwedge_i \neg y_i\right)$$

or

$$P_\lambda(x) = 1 - P_\lambda\left(\bigvee_i y_i\right).$$

(iv) Assume $P_\lambda(x) = 1$. Then $P_\lambda(x)$ gets contributions from both (1) and (2). The only option is that then $\lambda(x) =$ in, and so all attackers of y_i of x are out, so $\alpha_0 \Vdash \bigwedge \neg y_i$ and so $P_\lambda(\bigwedge_i \neg y_i) = 1$, because it gets contributions from both (1) and (2).

(v) Assume $P_\lambda(x) = 0$.

Thus neither (1) nor (2) contribute to $P_\lambda(x)$. Therefore $\alpha_0 \Vdash x$ and so $\lambda(x) =$ out and so for some attacker $y_i, \lambda(y_i) =$ in and so $\alpha_1 \Vdash y_i$ and so $P_\lambda(\bigwedge_i \neg y_i)$ cannot get any contribution either from (1) or from (2) and so $P_\lambda(\bigwedge_i \neg y_i) = 0$.

(vi) Assume that $P_\lambda(x) = \frac{1}{2}$.

So $P_\lambda(x)$ can get a contribution either from (1) or from (2), but not from both. So $\lambda(x)$ must be undecided.

So the attackers y_i of x are either **out** (with $P_\lambda(y_i) = 0$)) or **und** (with $P_\lambda(y_i) = \frac{1}{2}$), and we have that at least one attacker y of x is **und**.

Let y_i^0 be the attackers that are out and let $y_j^{\frac{1}{2}}$ be the undecided attackers. Consider

$$e = \bigwedge_i \neg y_i^0 \wedge \bigwedge_j \neg y_j^{\frac{1}{2}}.$$

The only model which can both contribute to $P_\lambda(e)$ is $\alpha_1 \wedge \alpha_0 \wedge \beta_{\frac{1}{2}}$ and thus $P_\lambda(e) = \frac{1}{2}$.

Thus from (iv), (v) and (vi) we get that (E3) holds for P_λ. □

Remark 3.13. Note that the P_λ of Theorem 3.12 is strictly a Method 2 probability. For example we saw that the extension $a = b =$ und of the network of Fig. 2 cannot be obtained by any Method 1 probability. The next section will see how far we can go with Method 1.

Summary of the results so far for the semantical probabilistic Method 2. We saw that Dung's traditional complete extensions strictly contain the probabilistic Method 1 extensions and is strictly contained in the probabilistic Method 2 extensions.

4. Approximating the Semantic Probability by Syntactic Probability

We have seen in Theorem 3.12 that the Method 2 probabilistic semantics can give us all the traditional Dung complete extensions. This result, together with the probabilistic semantics P_2 of Example 3.9 would show that Method 2 semantics is stronger than traditional Dung complete extensions semantics.

This section examines how far we can stretch the applicability of the syntactical probability approach (Method 1). We know from the "all-undecided" extension for the network in Fig. 2 that there are cases where we cannot give Method 1 probability. We ask in this section, can we approximate such extensions by Method 1 probabilities?

We find that the answer is yes.

Let $\langle S, R\rangle$ be a network. Let λ be a legitimate Caminada labelling giving rise to a complete extension $E = E_\lambda$. If the extension is a preferred extension, then there exists a solution f to the Eq_{inv} equations which yield λ and f is actually a Method 1 (and here also a Method 2) probabilistic semantics for $\langle S, R\rangle$. The question remains as to what happens in the case where λ is not a preferred extension. In this case we are not sure whether λ can be realised by a solution f of the

$\mathrm{Eq}_{\mathrm{inv}}$ equations. In fact there are examples of networks where no such f exists. We know from Theorem 3.12 that there exists a probability function P_λ on models that would yield λ according to Definition 3.10. We seek an $\mathrm{Eq}_{\mathrm{inv}}$ function which approximates this probability.

We shall use the ideas of Example 2.5.

Remark 4.1. We need to use some special networks.

1. Consider Fig. 9, which we shall call U_n. $n = 1, 2, 3, \ldots$.

   The $\mathrm{Eq}_{\mathrm{inv}}$ equations solve for this figure as $u_i = \frac{1}{2}, i = 1, \ldots, n$.

   $$u = \frac{1}{2^n}$$

   Thus if u attacks any node x, its "impact" on x is the multiplicative value $1 - \frac{1}{2^n}$. For n very large, the attack is almost negligible.
2. Let $\langle S, R\rangle$ be any network. Let u be a node not in S. If we add u to S and let it attack all elements of S, we can assume in view of (1) above that the $\mathrm{Eq}_{\mathrm{inv}}$ value of u is $\frac{1}{2^n}$. Figure 10 depicts this scenario.

   We suppress $\{u_1, \ldots, u_n\}$ and just record that $u = \frac{1}{2^n}$.

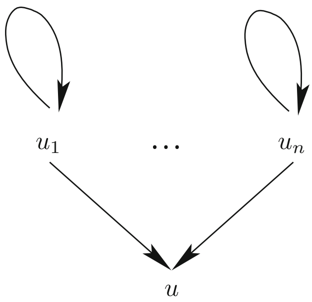

FIGURE 9. Multiple attacks by undecided nodes

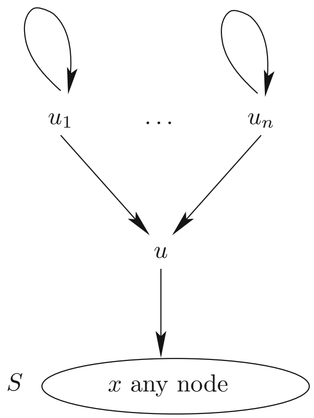

FIGURE 10. Scenario depicted in Remark 4.1

Construction 4.2. *Let $\langle S, R\rangle$ be given and let λ be a legitimate Caminada labelling giving rise to a non-preferred extension.*

Let $u \notin S$ be a new point and assume in view of Remark 4.1 that the value of u is very very small. Let

$$S' = S \cup \{u\}$$

and let

$$R' = R \cup \{(u, v) | \lambda(v) = \mathbf{und}\}.$$

Let $\lambda' = \lambda \cup \{(u, \mathbf{und})\}$.

Let $Att(x)$ be the set of all attackers of x in $\langle S, R\rangle$ and let $Att'(x)$ be the set of all attackers of x in $\langle S', R'\rangle$.

We have if $\lambda'(x) \in \{\mathbf{in}, \mathbf{out}\}$, then $u \notin Att'(x)$.

If $\lambda'(x) = \mathbf{und}$, then $y \in Att'(u)$.

Consider the following set of equations on $\langle S', R'\rangle$.

$$x = 1, \ \textit{if } \lambda'(x) = \mathbf{in} \tag{EQ1}$$

$$x = 0, \ \textit{if } \lambda'(x) = \mathbf{out} \tag{EQ0}$$

$$x = \Pi(1 - y)_{y \in Att'(x) in \ \langle S', R'\rangle}, \ \textit{if } \lambda'(x) = \mathbf{und} \tag{EQU}$$

This set of equations has a solution **f**.

We claim the following

1. *$\lambda(f)$ is a complete extension*
2. *$\lambda(f) = \lambda'$*

It is clear that $\lambda(f)(x) = \lambda'(x)$, for $\lambda'(x) \in \{\mathbf{in}, \mathbf{out}\}$. Does $\lambda(f)$ agree with λ' on undecided points of λ'? The answer is that it must be so, because λ' is a preferred extension. So $\lambda(f)$ cannot be an extension with more zeros and ones than λ'.

Remark 4.3. The perceptive reader might ask why do we use those particular equations in Construction 4.2 (page 24)? The answer can be seen from Fig. 11.

Consider $\lambda(a) = \mathbf{in}$, $\lambda(b) = \mathbf{out}$, $\lambda(c) = \lambda(d) = \mathbf{und}$.

We create Fig. 12.

We take the equation

$$a = 1, b = 0$$

$$c = (1 - d)(1 - u)$$

$$d = (1 - c)(1 - u)$$

$$u = 1 - u.$$

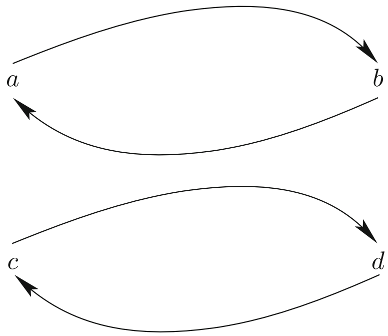

Figure 11. A network with two cycles

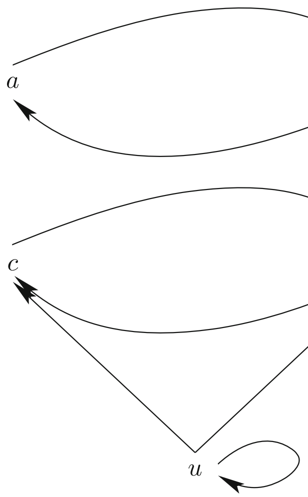

Figure 12. A self-attacking node attacking one of the cycles in the network of Fig. 11

The solution for the equations for c, d and u are

$$u = \frac{1}{2}$$

$$c = d = \frac{1}{3}$$

We have to insist on $a = 1, b = 0$. If we do not insist and write the usual equations

$$a = 1 - b$$

$$b = 1 - a,$$

we might get a different solution, e.g.

$$b = 1, a = 0.$$

This not the original λ.

Remark 4.4. This remark motivates and proves the next Theorem 4.5. We need some notation. Let Q be a set of atoms. By the models of Q (based on Q) we mean all conjunction normal forms of atoms from Q or their negations. So, for example, if $Q = \{a, b, c\}$, we get 8 models, namely

$$m_1 = a \wedge b \wedge c$$
$$\vdots$$
$$m_8 = \neg a \wedge \neg b \wedge \neg c.$$

If we have atoms

$$Q_1 = \{a_i\}, Q_2 = \{b_j\}, Q_3 = \{c_k\}$$

where Q_i are pairwise disjoint we can write the models of $Q_1 \cup Q_2 \cup Q_3$ in the form

$$\alpha \wedge \beta \wedge \gamma$$

where α is a model of Q_1, β of Q_2 and γ of Q_3.

For example

$$\alpha_1 \wedge \beta_1 \wedge \gamma_1 = (a_1 \wedge a_2 \wedge \cdots) \wedge (\neg b_1 \wedge b_2 \wedge \cdots) \wedge (c_2 \wedge \cdots).$$

Now let $\langle S, R\rangle$ and λ be as in Construction 4.2. Remember we assume that the value of u is very very small, and so the attack value $(1 - u)$ is very close to 1. Consider λ' and f and $\lambda(f)$ again as in Construction 4.2. f is a solution of $\mathrm{Eq}_{\mathrm{inv}}$ Eqs. (EQ1), (EQ0) and (EQU). Therefore any model of S', say $\alpha = \pm s_1 \wedge \pm s_2 \wedge \pm \ldots \wedge \pm s_k \wedge \pm u$ where $S = \{s_1, \ldots, s_k\}$ will have its probability semantics as

$$P_f(\alpha = \Pi_{i=1}^{k} f(\pm s_k))) \times f(\pm u) \qquad (*)$$

where

$$f(+s) = f(s)$$
$$f(-s) = 1 - f(s).$$

In particular, we have the following:

1. Let $E^+ = \{e_1^+, \ldots\}$ be the subset of S such that $\lambda(e_i^+) = \mathbf{in}$. Let $E^- = \{e_j^-\}$ be the subset of S such that $\lambda(e_j^-) = \mathbf{out}$. Let $E_{\mathrm{und}} = \{b_k\}$ be the set of all nodes in S such that $\lambda(b_k) = \mathbf{und}$.

   We therefore have that any model δ of S' has the form

$$\begin{aligned}\delta &= \textstyle\bigwedge_i \pm e_i^+ \wedge \bigwedge_i \pm e_j^- \wedge \bigwedge_k \pm b_k \wedge \pm u \\ &= \alpha \wedge \beta \pm u\end{aligned}$$

   where α is a model of $E^+ \cup E^-$ and β is a model of E_{und}.

   Let $\alpha_{1,0}$ be the particular conjunction

$$\alpha_{1,0} = \bigwedge_i e_i^+ \wedge \bigwedge_j \neg e_j^-.$$

Let β be any model of E_{und}. Consider $P_f(\delta), \delta = \alpha \wedge \beta \wedge \pm u$. Then by (*) we have that

$$P_f(\delta) = 0, \text{ if } \alpha \neq \alpha_{1,0}. \qquad (**)$$

Since P_f is a probability, we have for any $s \in S'$

$$P_f(s) = P_f\left(\bigwedge_{y \in Att'(s)} \neg y\right).$$

Note that for $s \in S, s \neq u$ such that $\lambda(s) \in \{\mathbf{in}, \mathbf{out}\}$, u does not attack s, and so we have

$$\begin{aligned} P_f(s) &= P(f)\left(\bigwedge\nolimits_{y \in Att(s)} \neg y\right) \\ &= \Pi_{y \in Att(s)}(1 - f(y)) \end{aligned} \qquad (\sharp 1)$$

For u we have that u is very small and so $P_f(u) = \frac{1}{2^n}$.

For $s \in S$ such that $\lambda(s) = \mathbf{und}$, we have that u attacks s and so

$$\begin{aligned} P_f(s) &= P_f\left(\bigwedge\nolimits_{y \in Att'(s)} \neg y\right) \\ &= \left(\Pi_{y \in Att(s)}(1 - f(y)\right) \times \left(1 - \tfrac{1}{2^n}\right) \end{aligned} \qquad (\sharp 2)$$

The $(1 - \frac{1}{2^n})$ is the attack of u.

We ask what are the attackers of $s \in E_{\text{und}}$? They cannot be nodes x such that $\lambda(x) = \text{in}$, because then s would be out. So the value of $f(y)$, (for $y \in Att(s)$) is either 0 or a value in $(0, 1)$.

So we can continue and write

$$P_f(s) = \left(1 - \frac{1}{2^n}\right) \Pi_{\substack{y \in Att(s) \\ \lambda(y) = \mathbf{und}}} (1 - f(y)) \qquad (\sharp 3)$$

Note that $0 < P_f(s) < 1$, because all the $f(y)$, for $\lambda(y) = \mathbf{und}$, satisfy $0 < f(y) < 1$.

We also have

$$\sum_{\text{all models } m} P_f(m) = 1. \qquad (\sharp 4)$$

Since (**) holds, we need consider only models m of the form $\alpha_{1,0} \wedge \beta \wedge \pm u$.

We can write

$$1 = \sum_{\beta \wedge \pm u} P_f(\alpha_{1,0} \wedge \beta \wedge \pm u) \qquad (\sharp 5)$$

where β is a model of E_{und}. Let us analyse ($\sharp$5) a bit more.

Assume $\beta = \bigwedge_k \pm b_k$.

So

$$P_f(\alpha_{0,1} \wedge \beta \wedge u) + P_f(\alpha_{0,1} \wedge \beta \wedge \neg u) = \Pi_k f(\pm b_k). \qquad (\sharp 6)$$

We thus get that:

$$\sum_{\beta} \Pi_k f(\pm b_k) = 1. \qquad (\sharp 7)$$

($\sharp$7) says something very interesting. It says that f restricted to E_{und} gives a proper probability distribution on the models of E_{und}.

This combined with ($\sharp$3) gives us the following result.

Consider $(E_{\text{und}}, R_{\text{und}})$ where $R_{\text{und}} = R \upharpoonright E_{\text{und}}$. Then $f \upharpoonright E_{\text{und}}$ is a proper probability distribution on $(E_{\text{und}}, R_{\text{und}})$.

Does it satisfy the proper equations?

Let $s \in E_{\text{und}}$. Do we have

$$P_{\text{und}}(s) \stackrel{?}{=} P_{\text{und}}\left(\bigwedge_{\substack{y \in E_{\text{und}} \\ yRx}} \neg y\right)$$

Let us check.

The real equation is

$$P_{\text{und}}(s) = P_{\text{und}}\left(\bigwedge_{\substack{y' \in E_{\text{und}} \\ [1ex]yRx}} \neg y\right) \times (1-u) \qquad (\sharp 8)$$

Since u is very small, we have a very good approximation.[5]

We can now define a probability P on $\langle S, R\rangle$. Let $m = \alpha \wedge \beta$ be a model, where α is a model for $E^+ \cup E^-$ and β is a model for E_{und}.

Then define P as follows

$$P(\alpha \wedge \beta) = 0, \text{ if } \alpha = \neg\alpha_{1,0}$$

$$P(\alpha \wedge \beta) = P_{\text{und}}(\beta), \text{ if } \alpha = \alpha_{1,0}$$

We need to show that approximately

$$P(s) = P\left(\bigwedge_{y \in Att(s)} \neg y\right)$$

If $s \in E^+ \cup E^-$ this follows from ($\sharp$1).

If $s \in E_{\text{und}}$, this follows from ($\sharp$3) and ($\sharp$8).

Note that since the f involved came from Eq_{inv} equations, P satisfies the following on $\langle S, R\rangle$.

$$\begin{aligned} P(s) &= 0, \text{ if some } y \in Att(s) P(y) = 1 \\ P(s) &= 1, \text{ if for all } y \in Att(s), P(y) = 0 \\ P(s) &= \text{ undecided, otherwise.} \end{aligned} \qquad (\sharp 9)$$

[5] The perceptive reader might ask what happens if we let u converge to 0? The answer is that we get a proper Eq_{inv} extension. However, this may be an all undecided extension (which is what we do want), or it may be a complete extension properly containing all the undecided extensions (which is not what we want!).

We may decide to do what physicists do to their equations. Write the equations in full and simply neglect any item containing higher order u, i.e., u^2, u^3, etc. This is reasonable when the value of each node is small.

Theorem 4.5.

1. *Let* $\langle S, R\rangle$ *be a network and let* λ *be a legitimate Caminada labelling on* S*. Then there exists a Method 1 probability distribution* P_λ*, which almost satisfies Eq.* (E3)*, namely for every* ε*, there exists a Method 1 probability* P_λ *depending on* ε*, such that for every* x *and its attackers* y_i*, we have* $|P_\lambda(x) - P_\lambda(\wedge\neg y_i)| < \varepsilon$*, such that*
$$\lambda(x) = \textbf{in},\ \textit{if } P_\lambda(x) = 1$$
$$\lambda(x) = \textbf{out},\ \textit{if } P_\lambda(x) = 0$$
$$\lambda(x) = \textbf{und},\ \textit{if } 0 < P_\lambda(x) < 1.$$
2. P *is obtained as follows*

   Case 1. λ *is a preferred extension. Then let* f *be a solution of* Eq_{inv} *for* $\langle S, R\rangle$*. Let* $P_\lambda = f$*.*

   Case 2. λ *is not a preferred extension.*

   Let $E^\lambda_{\text{und}} = \{x|\lambda(x) = \text{und}\}$*. Consider* $\langle S', R'\rangle$*, where* $S' = E^\lambda_{\text{und}} \cup \{u\}$*, where* u *is a new point not in* S *with value almost 0.*
$$R' = R \restriction E^\lambda_{\text{und}} \cup \{u\} \times E^\lambda_{\text{und}}.$$
   Then $\langle S', R'\rangle$ *has only one extension (all undecided). Let* f' *be a solution to* Eq_{inv} *on* $\langle S', R'\rangle$*. We now define* P_λ *on* $\langle S, R\rangle$*.*

   Let $\alpha_{1,0} = \bigwedge_{\lambda(x)=\ in} x \wedge \bigwedge_{\lambda(y)=\ out} \neg y$*.*

   Let $m = \alpha \wedge \beta$ *be an arbitrary model of* S*, where* α *is a model of* $\{x|\lambda(x) \in \{\textbf{in}, \textbf{out}\}$ *and* β *is a model of* E^λ_{und}*. Define* $P_\lambda(\alpha \wedge \beta)$ *to be*
$$P_\lambda(\alpha \wedge \beta = 0 \textit{ if } \alpha \neq \alpha_{1,0}$$
$$P_\lambda(\alpha_{1,0} \wedge \beta) = f'(\beta)$$
$$\textit{where } \beta = \bigwedge\nolimits_{s\in E^\lambda_{\text{und}}} \pm s$$
$$\textit{and } f(\beta) = \Pi_{\pm s\ in\ \beta} f(\pm s).$$

Proof. Follows from the considerations of Remark 4.4. □

Example 4.6. Let us show how Theorem 4.5 works by doing a few examples.

1. Consider the network of Fig. 11 and the extension λ mentioned there, namely $\lambda(a) = \textbf{in}$, $\lambda(b) = \textbf{out}$, $\lambda(c) = \lambda(d) = \textbf{und}$.

   Following our algorithms we look at the $\{c, d, u\}$ part of Fig. 12 and solve the equations. We get $u = \frac{1}{2}, c = d = \frac{1}{3}$.

   The probability P_λ will be as follows:
$$P_\lambda(\alpha \wedge \beta) = 0$$
$$\text{if } \alpha \neq a \wedge \neg b.$$
   Now look at
$$P_\lambda(a \wedge \neg b \wedge c \wedge d) = \tfrac{1}{3} \times \tfrac{1}{3} = \tfrac{1}{9}$$
$$P_\lambda(a \wedge \neg b \wedge c \wedge \neg d) = \tfrac{1}{3} \times \tfrac{2}{3} = \tfrac{2}{9}$$
$$P_\lambda(a \wedge \neg b \wedge \neg c \wedge d) = \tfrac{2}{3} \times \tfrac{1}{3} = \tfrac{2}{9}$$
$$P_\lambda(a \wedge \neg b \wedge \neg c \wedge \neg d) = \tfrac{2}{3} \times \tfrac{2}{3} = \tfrac{4}{9}.$$

2. Let us look at Fig. 13.

   With $\lambda(a) = \mathbf{in}$, $\lambda(b) = \mathbf{out}$, $\lambda(c) = \lambda(d) = \mathbf{und}$.

   The $\{c, d\}$ part is Fig. 2. Here we solve the equations on the $\{c, d, u\}$ part associated with $\{c, d\}$, which is the same as Fig. 3. The solution is found in Example 2.5, with $u = \frac{1}{2}$.

   We get $u = \frac{1}{2}; c = 0.36, 1 - c = 0.764, d = 0.382, 1 - d = 0.618$. The probability P_λ of this case is $P_\lambda(\alpha \wedge \beta) = 0$, if $\alpha \neq a \wedge \neg b$.

$$P_\lambda(a \wedge \neg b \wedge c \wedge d) = 0.236 \times 0.382 = 0.09$$

$$P_\lambda(a \wedge \neg b \wedge c \wedge \neg d) = 0.236 \times 0618 = 0.146$$

$$P_\lambda(a \wedge \neg b \wedge \neg c \wedge d) = 0.764 \times 0.382 = 0.292$$

$$P_\lambda(a \wedge \neg b \wedge \neg c \wedge \neg d) = 0.764 \times 0.618 = 0.472.$$

   Indeed

$$0.09 + 0.146 + 0.292 + 0.472 = 1.000.$$

We now discuss imposing probability on instantiated networks such as ASPIC+. We begin with simple instantiations into classical propositional logic.

Definition 4.7. 1. An abstract instantiated network (into classical propositional logic) has the form $\mathcal{A} = \langle S, R, I \rangle$, where $\langle S, R \rangle$ is an abstract argumentation network and I is a mapping associating with each $x \in S$, a well-formed formula $I(x) = \varphi_x$ of classical propositional logic.

2. For any $\mathcal{A}$ as in 4.7, we associate the theory $\Delta_\mathcal{A} = \{\varphi_x \leftrightarrow \wedge_{(y,x) \in R} \neg \varphi_y \mid x \in S\}$.
3. A semantic probability model P on $\mathcal{A}$ is a probability distribution on the models based on S such that for all $x \in S$, we have:

$$P(\varphi_x) = P(\wedge_{(y,x) \in R} \neg \varphi_y)$$

Example 4.8. Consider Fig. 14 where part (b) is an instatiation of part (a) with $I(x) = a_1 \vee a_2$ and $I(a_3) = a_3$. The equations any probability assignment

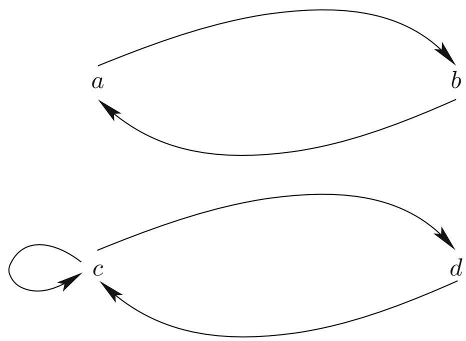

FIGURE 13. Augmented network of Fig. 2 with node a as c and b as d and an extra cycle

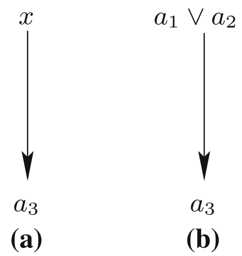

FIGURE 14. **a** A network and **b** one of its instantiations with $x = a_1 \vee a_2$

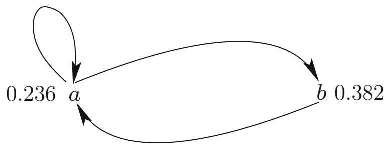

FIGURE 15. The $\mathrm{Eq}_{\mathrm{inv}}$ solution to the networks of Figs. 2 and 3

needs to satisfy are

$$\begin{aligned} P(a_1 \vee a_2) &= 1 \\ P(a_3) &= P(\neg(a_1 \vee a_2)) \\ &= P(\neg a_1 \wedge \neg a_2) \\ &= 0. \end{aligned}$$

If we let $P(a_1) = x$, $P(a_2) = 1 - x$, $P(a_3) = 0$, with $x \in [0, 1]$, then P satisfies the equations. Compare with Example 3.4.

5. Comparison with the Literature

There are several probabilistic argumentation papers around. This is a hot topic in 2014. We highlight two main points of view. The external and the internal views.

Let $\langle S, R\rangle$ be a network and let $\mathbf{f}$ be a function from S to $[0, 1]$. We can regard $\mathbf{f}$ as giving a probability number to each $x \in S$. The internal probability is where the above numbers signify the value of the argument. Its truth, its reliability, its probability of being effective, etc., or whatever measure we attach to it as an argument. Figure 15 represents in this case the $\mathrm{Eq}_{\mathrm{inv}}$ solution (and hence probability) of the network of Figs. 2 and 3. The external view is to think of $\mathbf{f}(x)$ as the probability of the predicate "$x \in S$". That is, the probability that the argument x is present in S. Consider again Fig. 15.

The probability that a is in the network is 0.236 and the probability that b is in the network is 0.382. Therefore, the probability that the network contains both $\{a, b\}$ is $0.236 \times 0.388 = 0.09$. The probability that the network contains only a is $0.236\times(1-0.382) = 0.1458$. The probability that the network contains only b is $0.382 \times (1 - 0.236) = 0.292$ and the probability that the network is empty is $(1 - 0.236) \times (1 - 0.382) = 0.472$. It is clear why we are calling this view an external probability view. It imposes probability externally expressing uncertainty on what the network graph is. This is done either by giving the probability to points or more generally by giving probability directly to subsets G of S, expressing the probability that the graph is really that subset of S with R restricted to G. This external view has value in dialogue argumentation or negotiation when we try to estimate what network our opponent is reasoning with. The problem with this external view is how to connect with the attack relation. Note that mathematically in the external view we have probabilities on points in S or probabilities on subsets of S, which are the same options as in our internal view, but the understanding of them is different. We in the internal view considered the subset as a classical model, while the external view considers it as a subnetwork. When we use the internal view, we can connect it with the attack relation via the equational approach [Eq. (E3)], but how would the external view connect with the attack relation? We can ask, for example, how to get a value for a single point to be "in" an extension? Intuitively, looking back at Fig. 15, we can say the point a for example is "in" in case the network is $\{a\}$ and is also "in" in one of the three extensions in case the network is $\{a, b\}$. So we might take the "in" value to be $0.1458+0.09/3 = 0.1458+0.03 = 0.1758$. The connection with the attack relation can be done perhaps through the probabilities for admissible sets, since being admissible is connected with the attack relation. There are problems, however, with this approach.

Hunter [7] was trying to lay some foundations for this view, following the papers [3,9]. See also a good summary in Hunter [8]. Hunter was trying to find a connection between the external probability view and some reasonable values we can give to admissible subsets. He proposes restrictions on the probability function on S. We are not going to discuss or reproduce Hunter's arguments here. It suffices to say that possibly a subsequent paper of ours will critically examine the external view and compare with the internal view.

Let us now compare our work with that of Thimm, [13], whose approach is also internal. We quote from [13]:

> In this paper we use another interpretation for probability, that of *subjective probability* [11]. There, a probability $P(X)$ for some $X \in \mathcal{X}$ denotes the *degree of belief* we put into X. Then a probability function P can be seen as an epistemic state of some agent that has uncertain beliefs with respect to $\mathcal{X}$. In probabilistic reasoning [11,12], this interpretation of probability is widely used to model uncertain knowledge representation and reasoning.
>
> In the following, we consider probability functions on sets of arguments of an abstract argumentation frameworks. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be some fixed abstract argumentation

framework and let $\mathcal{E} = 2^{\mathsf{Arg}}$ be the set of all sets of arguments. Let now $\mathcal{P}_{\mathsf{AF}}$ be the set of probability functions of the form $P : 2^{\mathcal{E}} \rightarrow [0,1]$. A probability function $P \in \mathcal{P}_{\mathsf{AF}}$ assigns to each set of possible extensions of AF a probability, i.e. $P(e)$ for $e \in \mathcal{E}$ is the probability that e is an extension and $P(E)$ for $E \subseteq \mathcal{E}$ is the probability that any of the sets in E is an extension. In particular, note the difference between e.g. $P(\{\mathcal{A},\mathcal{B}\}) = P(\{\{\mathcal{A},\mathcal{B}\}\})$ and $P(\{\{\mathcal{A}\},\{\mathcal{B}\}\})$ for arguments $\mathcal{A},\mathcal{B}$. While the former denotes the probability that $\{\mathcal{A},\mathcal{B}\}$ is an extension the latter denotes the probability that $\{\mathcal{A}\}$ or $\{\mathcal{B}\}$ is an extension. In general, it holds $P(\{\mathcal{A},\mathcal{B}\}) \neq P(\{\{\mathcal{A}\},\{\mathcal{B}\}\})$.

For $P \in \mathcal{P}_{\mathsf{AF}}$ and $\mathcal{A} \in \mathsf{Arg}$ we abbreviate

$$P(\mathcal{A}) = \sum_{\mathcal{A} \in e \subseteq \mathsf{Arg}} P(e).$$

Given some probability function P, the probability $P(\mathcal{A})$ represents the degree of belief that $\mathcal{A}$ is in an extension (according to P), i.e. $P(\mathcal{A})$ is the sum of the probabilities of all possible extensions that contain $\mathcal{A}$. The set $\mathcal{P}_{\mathsf{AF}}$ contains all possible views one can take on the arguments of an abstract argumentation framework AF.

Example 4. We continue Ex. 1. (Comment by Gabbay and Rodrigues: This is the network of our Fig. 4.) Consider the function $P \in \mathcal{P}_{\mathsf{AF}}$ defined via $P(\{\mathcal{A}_1,\mathcal{A}_3,\mathcal{A}_5\}) = 0.3, P(\{\mathcal{A}_1, \mathcal{A}_4\}) = 0.45, P(\{\mathcal{A}_5,\mathcal{A}_2\}) = 0.1, P(\{\mathcal{A}_2,\mathcal{A}_4\}) = 0.15$, and $P(3) = 0$ for all remaining $e \in \mathcal{E}$. Due to Prop. 1 the function P is well-defined as e.g.

$$\begin{aligned}&P(\{\{\mathcal{A}_5,\mathcal{A}_2\},\{\mathcal{A}_2,\mathcal{A}_4\},\{\mathcal{A}_3\}\})\\&\quad= P(\{\mathcal{A}_5,\mathcal{A}_2\}) + P(\{\mathcal{A}_2,\mathcal{A}_4\}) + P(\{\mathcal{A}_3\})\\&\quad= 0.1 + 0.15 + 0 = 0.25.\end{aligned}$$

Therefore, P is a probability function according to Def. 3. According to P the probabilities to reach argument of AF compute to $P(\mathcal{A}_1) = 0.75, P(\mathcal{A}_2) = 0.25, P(\mathcal{A}_3 = 0.3, P(\mathcal{A}_4) = 0.6$, and $P(\mathcal{A}_5) = 0.4$.

In the following, we are only interested in those probability functions of $\mathcal{P}_{\mathsf{AF}}$ that agree with our intuition on the interrelationships of arguments and attack. For example, if an argument $\mathcal{A}$ is not attacked we should completely believe in its validity if no further information is available. We propose the following notion of *justifiability* to describe this intuition.

Definition 4. A probability function $P \in \mathcal{P}_{\mathsf{AF}}$ is called *p-justifiable* wrt. AF, denoted by $P \Vdash_{\mathcal{J}} \mathsf{AF}$, if it satisfies for all $\mathcal{A} \in \mathsf{Arg}$.

1. $P(\mathcal{A}) \leq 1 - P(\mathcal{B})$ for all $\mathcal{B}, \in \mathsf{Arg}$ with $\mathcal{B} \rightarrow \mathcal{A}$ and

2. $P(\mathcal{A}) \geq 1 - \sum_{\mathcal{B}\in\mathcal{F}} P(\mathcal{B})$ where $\mathcal{F} = \{\mathcal{B}|\mathcal{B} \rightarrow \mathcal{A}\}$.

Let $P^{\mathcal{J}}_{\mathsf{AF}}$ be the set of all p-justifiable probability functions wrt. AF.

The notion of p-justifiability generalizes the concept of complete semantics to the probabilistic setting. Property 1.) says that the degree of belief we assign to an argument $\mathcal{A}$ is bounded from above by the complement to 1 of the degrees of belief we put into the attackers of $\mathcal{A}$. As a special case, note that if we completely believe in an attacker of $\mathcal{A}$, i.e., $P(\mathcal{B}) = 1$ for some $\mathcal{B}$ with $\mathcal{B} \rightarrow \mathcal{A}$, then it follows $P(\mathcal{A}) = 0$. This corresponds to property 1.) of a complete labelling (see Section 2). Property 2.) of Def. 4 says that the degree of belief we assign to an argument $\mathcal{A}$ is bounded from below by the inverse of the sum of the degrees of belief we put into the attacks of $\mathcal{A}$. As a special case, note that if we completely disbelieve in all attackers of $\mathcal{A}$, i.e. $P(\mathcal{B}) = 0$ for all $\mathcal{B}$ with $\mathcal{B} \rightarrow \mathcal{A}$, then it follows $P(\mathcal{A}) = 1$. This corresponds to property 2.) of a complete labeling, see Section 2. The following proposition establishes the probabilistic analogue of the third property of a complete labelling.

Proposition 2. *Let P be p-justifiable and $\mathcal{A} \in \mathsf{Arg}$. If $P(\mathcal{A}) \in (0, 1)$ then*

1. *there is no $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \rightarrow \mathcal{A}$ and $P(\mathcal{B}) = 1$ and*
2. *there is a $\mathcal{B}' \in \mathsf{Arg}$ with $\mathcal{B}' \rightarrow \mathcal{A}$ and $P(\mathcal{B}') > 0$.*

From our point of view, Thimm's approach is a variant of our semantic Method 2 approach without the strong Eq. (E3) but the weaker Definition 4 of Thimm. Thus Thimm will allow for different values for nodes x_1 and x_2 in our Fig. 8, while we would not (see Example 3.5).

Although Thimm's approach is mathematically close to us, conceptually we are far apart. Thimm motivates his approach as a degree of belief in a subset $E \subseteq S$, considering E as an extension. We consider E as representing a classical model m of the classical propositional logic with atoms S

$$m = \bigwedge_{s\in E} s \wedge \bigwedge_{s\notin E} \neg s$$

and assign probability to it and then we export this probability to argumentation via the equational approach, Eq. (E3).

This is an instance of our methodology of "Logic by Translation", From our point of view, Eq. (E3) are essential, conceptual and non-technical. For Thimm, the inequalities of his Definition 4 appear to be technical to enable the probabilities to work of ground extension.

Our point of view also leads us to the $\mathrm{Eq}_{\mathrm{inv}}$ Method 1 probabilities and to the approximation results of Sect. 4.

In Thimm's conceptual approach, this way of thinking does not even arise.

To summarise, this paper presented an internal view of probabilistic argumentation. There is a need for two subsequent research papers

1. The external view done coherently and its connection to the internal view.
2. A conditional probability view and its connection with Bayesian Networks views as Argumentation Networks.

6. Conclusions

This section explains and sets our approach in a general generic context.

Suppose we are given a system $\mathbb{S}$ such as an argumentation system $\langle S, R\rangle$ and we want to add to it some aspect $\mathbb{A}$.

There is a generic way to add any new feature to a system. It involves (1) identifying the basic units which build up the system and (2) introducing the new feature to each of these basic units. In the case where the system is argumentation and the feature is probabilistic we have the following: the basic units are (a) the nature of the arguments involved; (b) the membership relation in the set S of arguments;[6] (c) the attack relation; and (d) the choice of extensions.

Generically to add a new aspect (probabilistic, or fuzzy, or temporal, etc) to an argumentation network $\langle S, R\rangle$ can be done by adding this feature to each component. (a) We make the effective strength of the argument probabilistic; (b) we give probability to whether an argument is included in S;[7] (c) we make the attack relation probabilistic; and (d) we put probability on the extensions.

These features interact and need to be chosen with care and coordination. We need a methodological approach to make our choices. One such methodology is what we called âĂIJlogic by translationâĂİ.

We meaningfully translate the argumentation system into classical logic which does have probabilistic models and then let probabilistic classical logic endow the probability on the argumentation system. As we mentioned, this of course depends on how we translate.

We gave in this paper an object-level translation. The arguments of S became atoms of classical propositional logic, we then used probability on the models of classical logic and used the attack relation R to express equational restrictions on the probabilities. In this kind of translation, the attack relation did not become probabilistic.

We could have used a meta-level translation into classical predicate logic, using a binary relation R for expressing in classical logic the attack relation and using unary predicates to express that an argument x is “in”, x is “out”, etc., with suitable coordinating axioms. In this case all predicates would have

---

[6] Note that the set S itself may not be fully or accurately known, especially modelling an opponent in dialogue systems.

[7] (a) and (b) are distinct, because **a.** represents how effective an argument is, whereas (b) is the decision of whether or not to include an argument for consideration. An argument may be deemed very effective but not included for consideration for completely different reasons.

become probabilistic including the attack relation R. As far as we know nobody has done this to R.

In this context of possible options what we have done is one systematic approach and we compared it with other approaches. It should be noted that we could have followed the same steps to get fuzzy argumentation networks; temporal argumentation networks; or indeed any other feature available for classical propositional logic.

References

[1] Caminada, M., Gabbay, D.: A logical account of formal argumentation. Studia Logica, 109–145 (2012)

[2] Caminada, M., Pigozzi, G.: On judgment aggregation in abstract argumentation. Auton. Agents Multi Agent Syst. **22**(1), 64–102 (2011)

[3] Dung, P.M., Thang, P.: Towards (probabilistic) argumentation for jury-based depute resolution. In: Verheij, B., Szeider, S., Woltran, S. (eds.) Proceedings of COMMA III, Frontiers in Artificial Intelligence and Applications, pp. 171–182. IOS Press (2012)

[4] Gabbay, D.: Logics for Artificial Intelligence and Information Technology. College Publications, (2007)

[5] Gabbay, D.: Equational approach to argumentation networks. Argum. Comput. 87–142 (2012)

[6] Gabbay, D.M. Rodrigues, O.: A self-correcting iteration schema for argumentation networks. In: Parsons, S., Oren, N., Reed, C., Cerutti, F. (eds.) Proceedings of COMMA V, Frontiers in Artificial Intelligence and Applications, pp. 377 – 384. IOS Press (2014). doi:10.3233/978-1-61499-436-7-377

[7] Hunter, A.: Some foundations for probabilistic abstract argumentation. In: Verheij, B., Szeider, S., Woltran, S. (eds.) Proceedings of COMMA IV, Frontiers in Artificial Intelligence and Applications, pp. 117–128. IOS Press (2012)

[8] Hunter, A.: A probabilistic approach to modelling uncertain logical arguments. Int. J. Approx. Reason. **54**, 47–81 (2013)

[9] Li, H., Oren, N., Norman, T.: Probabilistic argumentation frameworks. In: Proceedings of the First International Workshop on the Theory and Applications of Formal Argumentation (TAFA'11), vol. 7132 of Lecture Notes in Computer Science. Springer, (2012)

[10] Modgil, S., Prakken, H.: the ASPIC+ framework for structured argumentation: A tutorial. Argum. Comput. **5**(1), 31–62 (2014)

[11] Paris, J.B.: The Uncertain Reasoner's Companion. A Mathematical Perspective. Cambridge University Press (2006)

[12] Pearl, J.: Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference. Morgan Kaufmann (1998)

[13] Thimm, M.: A probabilistic semantics for abstract argumentation. In: Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12) (2012)

D. M. Gabbay and O. Rodrigues
Department of Informatics
King's College London
The Strand
London WC2R 2LS
UK
e-mail: odinaldo.rodrigues@kcl.ac.uk

D. M. Gabbay
Bar Ilan University
Ramat Gan
Israel

D. M. Gabbay
University of Luxembourg
Luxembourg
Luxembourg
e-mail: dov.gabbay@kcl.ac.uk

Received: March 18, 2015.
Accepted: March 25, 2015.