Towards a probabilistic Dung-style argumentation system*

Tjitze Rienstra

University of Luxembourg / LIRMM, University of Montpellier 2, France

Abstract. In this paper we present a generalization of the ASPIC argumentation system, where a rule can be associated with a probability. This probability expresses uncertainty about whether or not the particular rule is active. We call this system p-ASPIC. The uncertainty about rules carries over to uncertainty about whether or not a particular argument is active. We generalize the usual Dung-style abstract argumentation semantics, in order to deal with this uncertainty. Our work should be understood as an initial step in the direction of probabilistic, instantiated, Dung-style argumentation, and we discuss a number of directions for future work.

1 Introduction

Dung's famous '95 paper [9] introduced a model of non-monotonic reasoning that, given a non-monotonic formalism, is based on the three-step procedure of (1) argumentation framework generation, (2) evaluation of arguments and (3) extraction of conclusions. That is, from a knowledge base, an argumentation framework is generated, consisting of arguments and attacks; this framework is evaluated in order to determine which arguments are acceptable; and from a set of acceptable arguments, we can extract conclusions that follow from the initial knowledge base. The main reason of the model's appeal is the fact that the complicated task of determining which conclusions can be non-monotonically derived from a knowledge base is reduced to the relatively simple task of selecting arguments from a framework. That is, on the *instantiated level*, we specify how to generate a framework, and on the *abstract level*, we deal only with the arguments and attacks, without looking at the actual content of the arguments. This *Dung-style argumentation* model has been used with a number of non-monotonic formalisms. One of the most general ones is the ASPIC argumentation system [15].

In this paper we look at a problem that has received little attention so far, namely the problem of reasoning under uncertainty using the Dung-style argumentation model. In some respect, non-monotonic reasoning is already reasoning under uncertainty. That is, we want to derive conclusions under incomplete or uncertain knowledge. This knowledge comes in the form of defeasible rules, which we do not want to accept as strict rules, or in the form of assumptions, which

* AT2012, 15-16 October 2012, Dubrovnik, Croatia.

we do not want to accept as facts. This defeasibility, sometimes made more expressive, as in the ASPIC system, by using of preferences over rules (see also [1, 13]) is a form of qualitative uncertainty. However, uncertainty is often expressed quantitatively, using probability theory. It is not clear how to deal with this kind of uncertainty, within the Dung-style argumentation model, and this is the problem that we address in this paper.

On the abstract level, we extend the notion of a framework to that of a probabilistic framework. This is an argumentation framework supplemented with a probability distribution over sets of arguments. Furthermore, we generalize the usual notion of an acceptance function, in order to evaluate the arguments of a probabilistic framework, and we show how it can be used to determine, given a probabilistic framework, the probability of skeptical and credulous acceptance of an argument.

On the instantiated level, we take as a basis a simplified version of the ASPIC argumentation system. This simplified version has also been considered in [6, 7], and it consists of the fragment of ASPIC dealing only with strict rules, defeasible rules, rebutting attack and undercutting attack. We call this system *ASPIC-Lite*. The approach that we take is simple: we associate with each strict and defeasible rule a probability. This probability is associated with the event that the rule is active, and may be applied in an argument. We finally show how to generate a probabilistic framework given a p-ASPIC theory.

The work presented in this paper should be understood as an initial step in the direction of probabilistic, instantiated, Dung-style argumentation. There are many open issues and possible directions for future work.

The outline of this paper is as follows. In section 2 we focus on the abstract level, and present first the familiar concepts of Dung-style argumentation: argumentation frameworks, labelings and acceptance functions. Then we present our probabilistic generalizations of these concepts, namely that of probabilistic frameworks, along with a method of determining the probability that arguments are skeptically or credulously accepted. In section 3 we turn to the instantiated setting, and we present a generalization of the ASPIC-Lite argumentation system, which we call p-ASPIC. This section defines how we can generate an instance of a probabilistic framework. In section 4.1, we turn again to the abstract level, where we demonstrate the evaluation of the arguments of a probabilistic framework, generated from a p-ASPIC theory. Finally, before we present our conclusions in section 7, we discuss in section 5 a number of open problems and in section 6 we give an overview of related work.

2 Formal preliminaries

2.1 Dung-style abstract argumentation theory

Dung-style argumentation theory centers on the concept of an *argumentation framework* (in short *framework*). A framework consists of a set of *arguments* and a binary *attack* relation, representing conflicts among arguments.

Definition 1. *An* argumentation framework F *is a pair* $(Args, Att)$*, where* $Args$ *is a finite set of* arguments*, and where* $Att \subseteq Args \times Args$ *is called the* attack relation*.*

Given a framework, we are interested in which arguments we can simultaneously accept. The answer is given in the form of a *labeling*, that represents an evaluation of the arguments of the framework.

Definition 2. *Given a framework* $F = (Args, Att)$*, a* labeling *is a function* $L : Args \to V$*, where* $V = \{I, U, O\}$*. We denote the set of all labelings of* F *by* $\mathbb{L}_F^{all}$*. If* L *is a labeling of* F *then* $L^{-1} : V \to 2^{Args}$*, where* 2^{Args} *denotes the power set of* $Args$*, is defined by* $L^{-1}(v) = \{a \in Args \mid L(a) = v\}$*.*

An argument labeled I, O or U (for *in, out* or *undecided*) is, respectively, accepted, rejected or neither accepted nor rejected. A function that takes as input a framework and returns a set of labelings, is called an *acceptance function*:

Definition 3. *An* acceptance function *is a function* $\mathbb{A}$ *that returns, for any framework* F*, a set* $\mathbb{A}(F) \subseteq \mathbb{L}_F^{all}$*.*

Different types of acceptance functions can be defined, each corresponding to an argumentation semantics that embodies a particular set of rationality criteria. The properties of *conflict-freeness* and *admissibility* are usually considered to be the minimum requirements:

Definition 4. *Given a framework* $F = (Args, Att)$*, we say that a labeling* $L \in \mathbb{L}_F^{all}$ *is* conflict-free[1] *iff:* $\nexists(a, b) \in Att$ *s.t.* $L(a) = L(b) = I$*; and* admissible *iff:* $\forall a \in Args$*,* $L(a) = I$ *implies* $\forall(b, a) \in Att, L(b) = O$ *and* $L(a) = O$ *implies* $\exists(b, a) \in Att$ *s.t.* $L(b) = I$*. We denote the set of conflict-free and admissible labelings of* F *by* $\mathbb{L}_F^{cf}$ *and* $\mathbb{A}_F^{ad}$*, respectively.*

Another property is *completeness*. It states that an argument is labeled I if and only if all attackers are labeled O, and that an argument is labeled O if and only if there is an attacker labeled I.

Definition 5. *Given a framework* $F = (Args, Att)$*, we say that a labeling* L *is* complete *if and only if:*

1. $L(a) = I$ *if and only if* $\forall(b, a) \in Att$*,* $L(b) = O$*.*
2. $L(a) = O$ *if and only if* $\exists(b, a) \in Att$*,* $L(b) = I$*.*

Note that complete labelings are also conflict free and admissible. Among the complete labelings, we can select those that satisfy certain properties, such as having a maximal set of arguments labeled I or U, or having no argument labeled U. The main types of acceptance functions are now defined as follows:

[1] Different definitions of conflict-free labelings exist. In [3], a conflict-free labeling is defined by the condition that, $\forall a \in Args$, $L(a) = I$ implies $\nexists(b, a) \in Att$ s.t. $L(b) = I$, and $L(a) = O$ implies $\exists(b, a) \in Att$ s.t. $L(a) = I$. Our definition is the same as in [8].

Definition 6. *The* complete acceptance function*, denoted by* $\mathbb{A}^{co}$*, is defined to return all complete labelings of* F*. The* preferred, grounded *and* stable *acceptance functions, denoted by* $\mathbb{A}^{pr}, \mathbb{A}^{gr}$ *and* $\mathbb{A}^{st}$*, respectively, are defined as follows.*

- $\mathbb{A}^{pr}(F) = \{L \in \mathbb{A}^{co}_F \mid \nexists K \in \mathbb{A}^{co}(F), K^{-1}(I) \supset L^{-1}(I)\}$
- $\mathbb{A}^{gr}(F) = \{L \in \mathbb{A}^{co}_F \mid \nexists K \in \mathbb{A}^{co}(F), K^{-1}(U) \supset L^{-1}(U)\}$
- $\mathbb{A}^{st}(F) = \{L \in \mathbb{A}^{co}_F \mid L^{-1}(U) = \emptyset\}$

Given a framework $(Args, Att)$, a set of arguments $B \subseteq Args$ and a semantics sem, two basic types of queries that can be answered are:

- *Skeptical acceptance:* Is, for all labelings returned by $\mathbb{A}^{sem}(F)$, at least one argument $a \in B$ labeled I?
- *Credulous acceptance:* Is, for at least one labeling returned by $\mathbb{A}^{sem}(F)$, at least one argument $a \in B$ labeled I?

Note that we consider acceptance of sets rather than single arguments because we may have that, in the instantiated setting presented in section 3, multiple arguments have the same conclusion. We are then interested in skeptical and credulous acceptance of a conclusion, rather than an argument, and we need to look at the set of all arguments with this conclusion.

2.2 Probabilistic Dung-style argumentation theory

In the previous section, we introduced the basic concepts of the theory of abstract argumentation theory. We now consider a probabilistic generalization of these concepts. Given a framework $(Args, Att)$, there may be uncertainty about whether or not an argument $a \in Args$ is active. This uncertainty may arise, for example, from:

- **Uncertainty of evidence.** Individual pieces of evidence, on which an argument is based, may be uncertain. This uncertainty carries over to the argument. So the probability that the argument is active is the *probability that the evidence is true.*
- **Opponent modeling.** If we use a framework to model the knowledge of an opponent (e.g. in the setting of an argumentation game), we may be uncertain about which arguments the opponent is aware of. So the probability that the argument is active is the *probability that the opponent is aware of the argument.*

To represent this kind of uncertainty, we introduce the concept of a *probabilistic framework*:

Definition 7. *A* probabilistic framework *is a pair* $PF = (F, P)$*, where* $F = (Args, Att)$ *is a framework and* $P : 2^{Args} \rightarrow [0, 1]$ *is a probability distribution over sets of arguments, called* framework states*, such that* $\sum_{C \in 2^{Args}} P(C) = 1$.

A framework state C corresponds to the event that all (and only) the arguments in C are active. We do not know which framework state $C \subseteq 2^{Args}$ is active; we only know its probability $P(C)$. We say that an argument a is active (resp. inactive) in C if $a \in C$ (resp. $a \notin C$).

Example 1. Consider the probabilistic framework (F, P) with $F = (Args, Att)$, where $Args = \{a, b\}$, $R = \{(a, b), (b, a)\}$. Thus, we have four framework states: $C_1 = \emptyset, C_2 = \{a\}, C_3 = \{b\}$ and $C_4 = \{a, b\}$. Suppose the probability that a is active is 0.4 and that b is active is 0.5, and that these events are independent. The probability distribution over framework states is then defined by $P(C_1) = P(C_3) = 0.3$ and $P(C_2) = P(C_4) = 0.2$.

The way we evaluate the arguments of the framework F, given a state C, is straightforward. We apply the criterium of completeness with the proviso that all inactive arguments are labeled O. More precisely, given a state C, an argument is labeled I if and only if it is active and all attackers are labeled O or is labeled O if and only if it is inactive or there is an attacker labeled I. We thus end up with a notion of *completeness given a framework state C:*

Definition 8. *Given a framework $(Args, Att)$ and a framework state $C \subseteq Args$, we say that a labeling $L \in \mathbb{L}^{all}$ is* complete given the framework state C *iff:*

1. *$L(a) = I$ if and only if $a \in C$ and $\forall (b, a) \in Att$, $L(b) = O$.*
2. *$L(a) = O$ if and only if either $a \notin C$ or $\exists (b, a) \in Att$, $L(b) = I$.*

We can now define the *conditionally* complete, preferred, grounded and stable acceptance functions, that take as input a framework *and* a framework state:

Definition 9. *Given a framework F, the* conditionally *complete, preferred, grounded and stable acceptance functions*[2] $\mathbb{CA}_F^{co}, \mathbb{CA}_F^{pr}, \mathbb{CA}_F^{gr}$ *and* $\mathbb{CA}_F^{st}$ *are defined by:*

- $\mathbb{CA}^{co}(F, C) = \{L \in \mathbb{L}_F^{all} \mid L$ *is complete given* $C\}$.
- $\mathbb{CA}^{pr}(F, C) = \{L \in \mathbb{CA}_F^{co}(C) \mid \nexists K \in \mathbb{CA}_F^{co}(C), K^{-1}(I) \supset L^{-1}(I)\}$
- $\mathbb{CA}^{gr}(F, C) = \{L \in \mathbb{CA}_F^{co}(C) \mid \nexists K \in \mathbb{CA}_F^{co}(C), K^{-1}(U) \supset L^{-1}(U)\}$
- $\mathbb{CA}^{st}(F, C) = \{L \in \mathbb{CA}_F^{co}(C) \mid L^{-1}(U) = \emptyset\}$

Now we generalize the two basic types of queries. Given a probabilistic framework $((Args, Att), P)$, a semantics *sem* and a set $B \subseteq Args$, we can ask:

- What is the probability that B is skeptically accepted?
- What is the probability that B is credulously accepted?

Loosely speaking, the probabilities of skeptical and credulous acceptance can be understood as lower and upper bound probabilities, where the differences between the two arise from conflicts among arguments. The probabilities are calculated as follows.

[2] The concept of a conditional acceptance function presented here is similar to the one presented in [4], with some difference in notation. More precisely, our $\mathbb{CA}^x(F, C)$ is equivalent in [4] to $\mathbb{CA}_F^x([L]_F)$, where $L^1(I) = C$ and $L^1(O) = Args \setminus C$.

Definition 10. *Let $PF = (F, P)$ be a probabilistic framework, $F = (Args, Att)$, $B \subseteq Args$ a set of arguments and sem a semantics. The probability of skeptical acceptance, denoted $P^{sem}_{skep}(B)$, is defined by:*

$$P^{sem}_{skep}(B) = \sum_{\{C \in 2^{Args} | \forall L \in \mathbb{CA}^{sem}(F,C), \exists a \in B, L(a)=I\}} P(C)$$

The probability of credulous acceptance, denoted $P^{sem}_{cred}(B)$, is defined by:

$$P^{sem}_{cred}(B) = \sum_{\{C \in 2^{Args} | \exists L \in \mathbb{CA}^{sem}(F,C), \exists a \in B, L(a)=I\}} P(C)$$

In words, the definitions above amount to: the probability of skeptical (resp. credulous) acceptance of B is the sum probability of all framework states in which B is skeptically (resp. credulously) accepted.

Example 2. (Continued from 1) First, if we look at the framework F, we can see that both $\{a\}$ and $\{b\}$ are credulously accepted but not skeptically accepted, under the complete semantics. Let us now look at the probability of credulous and skeptical acceptance under the complete semantics, given the probabilistic framework (F, P). We have that the set $\{a\}$ is skeptically accepted only given framework state C_2 and the set $\{b\}$ only given framework state C_3. Hence $P^{co}_{skep}(\{a\}) = P(C_2) = 0.2$ and $P^{co}_{skep}(\{a\}) = P(C_3) = 0.3$. The set $\{a\}$ is credulously accepted in the framework states C_2 and C_4 and $\{b\}$ in C_3 and C_4. Hence $P^{co}_{cred}(\{a\}) = P(C_2) + P(C_4) = 0.4$ and $P^{co}_{cred}(\{b\}) = P(C_3) + P(C_4) = 0.5$.

We are now equipped to deal with uncertainty of arguments on the abstract level, using the concept of a probabilistic argumentation framework. In the following section, we present an instantiation that generates a probabilistic framework, called p-ASPIC. That is, we define a logical language to specify a p-ASPIC theory, which includes rules associated with probabilities, and we specify how we can construct a probabilistic framework from such a theory. After that we return to the issue of evaluating the arguments of a probabilistic framework, where we apply the concepts introduced here, to a probabilistic framework generated by the p-ASPIC system.

3 p-ASPIC: ASPIC with probabilities

We present here a generalization of the ASPIC-Lite system that supports rules associated with a probability. We call it p-ASPIC. Let $\mathcal{L}$ be a logical language which is closed under negation and let $- : \mathcal{L} \to \mathcal{L}$ be defined by $-\phi = \psi$, if $\phi = \neg\psi$, and $-\phi = \neg\phi$, otherwise. In the ASPIC-Lite system, knowledge is represented using two types of rules: *strict rules*, of which the conclusion *always* holds, if the premises hold, and defeasible rules, of which the conclusion holds defeasibly, if the premises hold. Here, we associate the rules with a probability, and we call them *p-rules*.

Definition 11. *A* p-rule *is a rule of the form* $\Phi \rightarrow_p^x \psi$*, where* $\Phi \subseteq \mathcal{L}$ *is the* antecedent *and* ψ *the* consequent*,* $x \in \{s, d\}$ *the* type*, and* $p \in (0, 1]$ *is the* probability *of the rule. Given a rule* $\Phi \rightarrow_p^x \psi$ *we say that it is a* strict p-rule *if* $x = s$ *and a* defeasible p-rule *if* $x = d$*.*

Throughout this text, we either write $\Phi \rightarrow_p^x \psi$, where x ranges over any of the two types of rules, or $\Phi \rightarrow_p^s \psi$ and $\Phi \rightarrow_p^d \psi$, to denote a strict and defeasible p-rule in particular. Our notation is different from that in [6, 7], where strict rules are written using $\rightarrow$ and defeasible rules using $\Rightarrow$. We adopt the type parameter for notational convenience.

The probability p, associated with a p-rule $\Phi \rightarrow_p^x \psi$, is interpreted as the probability that the rule is active. As we hinted at before, such probabilities may arise from:

- **Uncertainty of evidence.** It may be uncertain whether or not a fact or rule is a valid constituent of an argument. So the probability that a rule is active is the *probability that the rule is valid*.
- **Opponent modeling.** If we use a framework to model the knowledge of an opponent (e.g. in the setting of an argumentation game), we may be uncertain about which rules the opponent is aware of. So the probability that the argument is active is the *probability that the opponent is aware of the rule.*

Note that we assume that there are no rules with zero probability. For a rule that is certain, i.e. has a probability of 1, we omit the probability and simply write $\Phi \rightarrow^x \psi$. A *p-theory* is a set of p-rules:

Definition 12. *A* p-theory *is a set* R *of p-rules. We denote by* $D(R)$ *the set of defeasible rules in* R *and by* $S(R)$ *the set of strict rules in* R*.*

In the ASPIC-Lite system, a theory is assumed to be closed under transposition of strict rules and to be consistent. These requirements are necessary to enforce certain coherence requirements on the output of the ASPIC-Lite system (we refer to [5] for details). We define these properties below. Again,we deviate slightly from [6, 7]. We do not assume R to be closed under transposition of strict rules, and R to be consistent. Rather, we require that the *closure* of R under transposition of strict rules is consistent. Furthermore, we require another property, not considered in the normal ASPIC-Lite system, namely *non-redundancy* of strict rules. This means that R is minimal, in that R does not contain transpositions of other rules in R. Formally:

Definition 13. *Let* $P \subseteq \mathcal{L}$ *and* R *a p-theory. The* closure of P under the strict rules in R*, denoted* $Cl_R(P)$*, is the smallest set satisfying:*

- $P \subseteq Cl_R(P)$,
- *If* $\{\phi_1, \ldots, \phi_n\} \rightarrow_p^s \psi \in R$ *and* $\phi_1, \ldots, \phi_n \in Cl_R(P)$ *then* $\psi \in Cl_R(P)$.

Definition 14. *A strict rule* $\Phi' \rightarrow_p^s \psi'$ *is a* transposition *of a strict rule* $\Phi \rightarrow_p^s \psi$ *if and only if* $\exists \phi \in \Phi$ *s.t.* $\Phi' = (\Phi \setminus \{\phi\}) \cup \{-\psi\}$ *and* $\psi' = -\phi$. *Let* R *be a p-theory. The* closure of R under transposition of strict rules*, denoted* $Tr(R)$*, is the smallest set satisfying:*

- $R \subseteq Tr(R)$
- *If* $\Phi \rightarrow_p^s \psi \in Tr(R)$ *and* $\Phi' \rightarrow_p^s \psi'$ *is a transposition of* $\Phi \rightarrow_p^s \psi$ *then* $\Phi' \rightarrow_p^s \psi' \in Tr(R)$.

Definition 15. *Let* $P \subseteq \mathcal{L}$. *We say that* P is consistent *if and only if there is no* $\phi \in \mathcal{L}$ *such that* $\phi, -\phi \in P$. *Let* R *be a p-theory. We say that* R is consistent *if and only if there is no consistent* $P \subseteq \mathcal{L}$ *and* $\phi \in \mathcal{L}$ *such that* $\phi, -\phi \in Cl_{Tr(R)}(P)$.

Definition 16. *Let* R *be a p-theory. A strict rule* $r \in S(R)$ *is* redundant in R *if and only if* $Tr(R) = Tr(R \setminus \{r\})$. *We say that* R *is non-redundant if and only if* $\forall r \in S(R)$, r *is not redundant.*

Throughout this paper, we assume that R is consistent and non-redundant and, furthermore, that there are no two rules $\Phi \rightarrow_p^x \psi, \Phi \rightarrow_{p'}^x \psi \in R$ such that $p \neq p$, i.e. there are no two rules with equivalent antecedents/premises, but a different p-value. In the rest of this paper, we use the following running example.

Example 3. Consider the p-theory $R = \{r_1, \ldots, r_{10}\}$ with:

$r_1 = \emptyset \rightarrow^s aw$	$r_5 = \{aw\} \rightarrow^d aj$	$r_9 = \{aj, bj, cj\} \rightarrow^s \neg dj$
$r_2 = \emptyset \rightarrow_{0.5}^s bw$	$r_6 = \{bw\} \rightarrow^d bj$	$r_{10} = \{cj, dj\} \rightarrow^s \neg \lceil r_5 \rceil$
$r_3 = \emptyset \rightarrow_{0.6}^s cw$	$r_7 = \{cw\} \rightarrow^d cj$	
$r_4 = \emptyset \rightarrow_{0.7}^s dw$	$r_8 = \{dw\} \rightarrow^d dj$	

The example can be interpreted as follows: Four friends, let us call them Anne, Bob, Chris and David have expressed their desire to join you on a road trip. The strict rules $r_1, \ldots, r_4$ stand for 'Anne wants to join', ..., 'David wants to join'. You are not certain whether Bob, Chris and David want to join. This is expressed by the probabilities associated with r_2, r_3 and r_4. The defeasible rules $r_5, \ldots, r_8$ express the principle that if your friends want *want* to join you then, if possible, they *can* join you. However, your car seats at most three passengers, which is expressed by the rule r_9. (Also consider the transpositions of r_9, e.g. $\{aj, bj, dj\} \rightarrow^s \neg cj$.) Furthermore, both Chris and David are in love with Anne, which could lead to unpleasant situations in the cramped car. You decide, therefore, that if Chris and David both join, then Anne cannot join, which is expressed by r_{10}. (Here, $\lceil \ldots \rceil$ stands for the objectification operator, which maps defeasible rules to elements of the language. We need this to define *undercut*, which we introduce below.)

An ASPIC-Lite theory can be used to instantiate a Dung-style argumentation framework. This framework is then an abstract representation of the reasoning problem posed by the theory. To do this, the ASPIC system specifies how arguments are constructed and how the attack relation is defined. Moreover, the probabilities associated with the rules of a p-theory, allow us to define a probability distribution over sets of arguments. We can thus specify how to instantiate a probabilistic framework. Before turning to this aspect, we start out with arguments and attacks. An argument is a chaining of rules, where the premises of every rule follow from the conclusions of preceding rules. Apart from notation, we define argument generation as usual in the ASPIC-Lite system:

Definition 17. *An* argument *a is a pair (B, ϕ), where B is a set of p-rules and $\phi \in \mathcal{L}$ is the* conclusion. *The set of arguments* generated by *a p-theory R, denoted $Args_R$, is inductively defined as follows:*

– *If $\{\phi_1, \ldots, \phi_n\} \rightarrow_p^x \psi \in Tr(R)$ and $(\Phi_1, \phi_1), \ldots, (\Phi_n, \phi_n) \in Args_R$ then $(\Phi_1 \cup \ldots \cup \Phi_n \cup \{\{\phi_1, \ldots, \phi_n\} \rightarrow_p^x \psi\}, \psi) \in Args_R$.*

Given a set of arguments A, we denote by $R(A)$ the set of rules appearing in the arguments in A and by $S(A)$ and $D(A)$ the set of strict and defeasible rules, respectively, appearing in A.

Notice that the base case of the inductive definition of $Args_R$ is the case where a rule has an empty set of premises.

Example 4. (Continued from 3) The set of arguments generated by the p-theory of example 3 is the set $Args_R = \{a_1, \ldots, a_{13}\}$ with:

$a_1 = (\{r_1\}, aw)$
$a_2 = (\{r_2\}, bw)$
$a_3 = (\{r_3\}, cw)$
$a_4 = (\{r_4\}, dw)$
$a_5 = (\{r_1, r_5\}, aj)$
$a_6 = (\{r_2, r_6\}, bj)$
$a_7 = (\{r_3, r_7\}, cj)$
$a_8 = (\{r_4, r_8\}, dj)$
$a_9 = (\{r_2, r_3, r_4, r_6, r_7, r_8, r_9\}, \neg aj)$
$a_{10} = (\{r_1, r_3, r_4, r_5, r_7, r_8, r_9\}, \neg bj)$
$a_{11} = (\{r_1, r_2, r_4, r_5, r_6, r_8, r_9\}, \neg cj)$
$a_{12} = (\{r_1, r_2, r_3, r_5, r_6, r_7, r_9\}, \neg dj)$
$a_{13} = (\{r_3, r_7, r_4, r_8, r_{10}\}, \neg\lceil r_5 \rceil)$

Two kinds of attack are defined by the ASPIC-Lite system. The first is defined by a conclusion of the attacking argument contradicting a defeasible conclusion of the attacked argument. This is called *rebut*. For the second type, called *undercut*, it is assumed, in the ASPIC-Lite system, that there is an *objectification operator* $\lceil \ldots \rceil$ that maps defeasible rules to elements of $\mathcal{L}$. Formally:

Definition 18. *Given a p-theory R and the generated arguments $Args_R$, we say:*

– *An argument $(B, \phi) \in Args_R$* rebuts *$(B', \phi') \in Args_R$ if and only if there is a $\Phi \rightarrow_p^d \psi \in D(B')$ and $\psi = -\phi$.*
– *An argument $(B, \phi) \in Args_R$* undercuts *$(B', \phi') \in Args_R$ if and only if there is a $\Phi \rightarrow_p^d \psi \in D(B')$ and $\phi = \neg\lceil \Phi \rightarrow_p^d \psi \rceil$.*

The attack relation *$Att_R \subseteq Args_R \times Args_R$ is defined by $((B, \phi), (B', \phi')) \in Att_R$ if and only if (B, ϕ) rebuts or undercuts (B', ϕ').*

Example 5. (Continued from 4) Some attacks in our running example are: a_9 rebuts a_5, a_{10}, a_{11} and a_{12}; a_{13} undercuts a_{10}, a_{11}, a_{12} and a_{13}; and both a_{11} and a_{12} rebut a_{13}.

Given a p-theory we can now generate a framework:

Definition 19. *Given a p-theory R, the* framework generated by R *is the framework* $F_R = (Args_R, Att_R)$.

Example 6. (Continued from 4) Figure 1 shows the argumentation framework $(Args_R, Att_R)$ generated by the p-theory of our running example. The nodes represent arguments in $Args_R$ and are labeled with the argument name and the conclusion. The edges represent the attack relation Att_R.

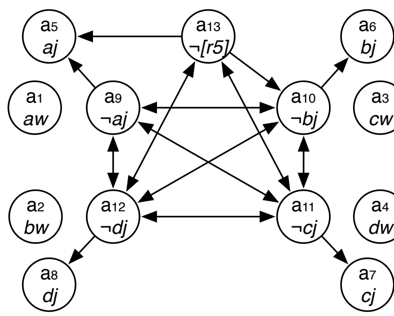

Fig. 1. The framework of our running example.

S	p_S
$S_0 = R_{certain}$	.06
$S_1 = R_{certain} \cup \{r_2\}$	.14
$S_2 = R_{certain} \cup \{r_3\}$	.09
$S_3 = R_{certain} \cup \{r_2, r_3\}$	.21
$S_4 = R_{certain} \cup \{r_4\}$	.06
$S_5 = R_{certain} \cup \{r_2, r_4\}$	.14
$S_6 = R_{certain} \cup \{r_3, r_4\}$	.09
$S_7 = R_{certain} \cup \{r_2, r_3, r_4\}$	.21

Table 1. Theory states in our running example.

We have now described the process of the generation of a framework, given a p-theory R, as in the ASPIC-Lite system. However, we have not done anything with the probabilities associated with the rules. We turn to this in the following section, where we describe the process of generating a probabilistic framework, given a p-theory R.

4 Calculating probabilities of arguments

As we said, we interpret a p-rule $\Phi \rightarrow_p^x \psi$ as saying that the probability that the rule $\Phi \rightarrow^x \psi$ is active, is p. That is, the rule is active and can be applied, with a probability of p and it is inactive and cannot be applied, with a probability of $1 - p$. The uncertainty about the rules used in an argument carry over to uncertainty about whether or not the argument is active. That is, the probability that an argument (B, ϕ) is active, is the probability that all p-rules in B are active. To model this, we define the notion of a *theory state*. A theory state is corresponds to the event that a set of rules is active. If a theory state holds, then all and only the rules in the theory state are active. Next, we define a probability distribution over the set of theory states. We can, in principle, use any type of distribution but, for simplicity, we assume here that the events of different p-rules being active, are independent. Thus, the probability of the event that a theory state is active, is easily calculated:

Definition 20. *Let R be a p-theory. A* theory state *is a set $S \subseteq R$. If $(\Phi \to \psi) \in Tr(S)$, we say that $(\Phi \to \psi)$ is* active *in S, otherwise* inactive*. The probability of S, denoted p_S, is defined by*

$$p_S = \prod_{(\Phi \to^x_p \psi) \in S} p \prod_{(\Phi \to^x_p \psi) \in R \setminus S} (1-p)$$

Note that, from the way the probability of a theory state is calculated, it is easy to see that the sum probability of all theory states is 1. Also note that a strict rule $\Phi \to^s \psi$ is active in S if and only if all transpositions of $\Phi \to^s \psi$ are active in S. Apart from transposition, we may consider in definitions 16 and 20 closure under additional properties. For example, we may require that rules are closed under transitivity so that, in a state, two rules $\{a\} \to^s b, \{b\} \to^s c$ are active if and only if $\{a\} \to^s c$ is active. However, for the current purpose, we consider only transposition.

Example 7. (Continued from 6) There are three uncertain rules in the p-theory of our running example, namely r_2, r_3 and r_4. Let $R_{certain} = R \setminus \{r_2, r_3, r_4\}$. Every theory state not containing all rules in $R_{certain}$ receives zero probability. Table 1 lists the states that receive non-zero probability:

Given the probability distribution over theory states, the probability of a framework state (see definition 7) is defined by the sum probability of all theory states that give rise to the framework state.

Definition 21. *Let R be a p-theory and $(Args_R, Att_R)$ the framework generated by R. We say that a theory state S* gives rise to *a framework state $C \subseteq 2^{Args}$ if and only if $C = G(S)$. The* probability distribution over framework states generated by *R, denoted P_R, is defined by:*

$$P_R(C) = \sum_{\{S \in 2^R | C = G(S)\}} p_S$$

Proposition 1. *Given a p-theory R and the framework $(Args_R, Att_R)$ generated by R, the sum probability of all framework states $C \subseteq Args_R$ is 1, i.e. $\sum_{C \subseteq Args_R} P_R(C) = 1$.*

Note that, while the events of different p-rules being active are be independent, the events of different arguments being active, are in general not independent. This is not surprising, because the same p-rule may be used in several arguments.

Furthermore, some framework states may receive zero probability. This happens when no theory state gives rise to the particular framework state. These framework states are exactly those that are not well-formed, in the sense that they are not *closed under argument construction*:

Definition 22. *We say that a set of arguments C is* closed under argument construction *if and only if $C = A_{R(C)}$*

Proposition 2. *Let R be a p-theory, $(Args_R, Att_R)$ the framework generated by R and $C \subseteq Args_R$. If $P_R(C) > 0$ then C is closed under argument construction.*

Note that the right-to-left direction of proposition 2 fails to hold only because of rules $r \in R$ with probability 1. Then there may be a framework state C that is closed under argument construction but receives zero probability because $r \notin R(C)$.

Example 8. Consider a p-theory $R = \{\emptyset \rightarrow^s_{0.5} a, a \rightarrow^s_{0.5} b\}$. Two arguments that are generated by this theory are: $a_1 = (\{\emptyset \rightarrow^s_{0.5} a\}, a)$ and $a_2 = (\{\emptyset \rightarrow^s_{0.5} a, a \rightarrow^s_{0.5} b\}, b)$. We have that the framework states $\{a_1\}$ and $\{a_1, a_2\}$ are closed under argument construction, but $\{a_2\}$ is not. There are four theory states: $S_1 = \emptyset, S_2 = \{r_1\}, s_3 = \{r_2\}$ and $S_4 = \{r_1, r_2\}$. S_1 gives rise to the framework state $\emptyset$, S_2 to $\{a_1\}$, S_3 to $\emptyset$ and S_4 to $\{a_1, a_2\}$. No theory state gives rise to $\{a_2\}$. Hence it receives zero probability.

We can now generate a probabilistic framework:

Definition 23. *Given a p-theory R, the* probabilistic framework generated by R *is the probabilistic framework $PF_R = (F_R, P_R)$.*

Example 9. (Continued from 7) Given the p-theory of our running example, the probabilistic framework is (F_R, P_R), where F_R is the framework shown in example 6 and where P_R is displayed in table 2. (The column C shows the framework states that receive non-zero probability, the column $P_R(C)$ shows the corresponding probabilities and in every row, a checkmark means that the argument is active in the corresponding framework state.)

C	$P_R(C)$	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}
$C_1 = G(S_1)$	.06	✓				✓								
$C_2 = G(S_2)$	.14	✓	✓			✓	✓							
$C_3 = G(S_3)$	.09	✓		✓		✓		✓						
$C_4 = G(S_4)$	.21	✓	✓	✓		✓	✓	✓				✓		
$C_5 = G(S_5)$	.06	✓			✓	✓			✓					
$C_6 = G(S_6)$	.15	✓	✓		✓	✓	✓		✓				✓	
$C_7 = G(S_7)$	.09	✓		✓	✓	✓		✓	✓		✓			✓
$C_8 = G(S_8)$	.21	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 2. The framework states in our running example.

4.1 Evaluating arguments of a p-ASPIC framework

In the previous sections we demonstrated how, given a p-theory, we can generate a probabilistic framework. In this section we conclude our running example and show the result of calculating the probability of skeptical and credulous acceptance of a conclusion from our p-theory.

Example 10. If we take the complete semantics, then the different labelings correspond to different ways in which the constraints, expressed by the rules of our theory, can be satisfied. Focusing on just Anne, we can answer the following questions: *What is the lower bound probability that Anne joins?* and *What is the upper bound probability that Anne joins?*. These questions correspond, respectively, to the probabilities of skeptical and credulous acceptance. Regarding skeptical acceptance, we have that a_5 is labeled I in all complete labelings, given all framework states except the framework states C_7 and C_8. Therefore, this probability is the sum probability of the states $C_1, \ldots, C_6$, which is 0.7. (Indeed, Anne does not necessarily join because whenever both Chris and David want to join ($p = 0.3$), it is not sure that Anne will join.) Regarding credulous acceptance, we have that a_5 is labeled I in at least one complete labeling, given all possible framework states. Therefore, this probability is 1. (Indeed Anne can always possibly join; even if Bob, Chris and David want to join, it is possible to refuse either Chris or David, so that Anne can join.)

5 Discussion and future work

In this paper we presented an initial attempt at combining probability with instantiated and abstract argumentation. Many issues are still open and are left for future work. Here we give an overview.

One limitation of our formalization is that we apply probabilities on a meta-level rather than the object-level. That is, we associate probabilities with rules and not with elements of $\mathcal{L}$. We cannot, for example, express in our formalism that the conclusion of a rule holds with some probability, given that the premises hold. We can say that our formalism is strictly about reasoning *under* uncertainty, rather than reasoning *about* uncertainty. Furthermore, it is not very clear, from a conceptual point of view, what it means to associate a probability with a rule. However, as our example shows, it is quite natural to associate probabilities with facts (i.e. strict rules with empty sets of premises). This suggests two directions for future work: one the one hand, we can investigate argumentation *about* uncertainty and on the other hand, we can focus on the 'uncertainty about facts' part, and from there, add expressivity.

It is worth mentioning one important issue, when modeling argumentation *about* uncertainty. Namely, how do we interpret rules? An obvious candidate for such an interpretation is conditional probability, but then we lose context independence, transitivity (i.e. chaining) and transposition.

Furthermore, we have assumed in this paper that the events of individual rules being active are probabilistically independent. While this assumption reduces the complexity, it leads to serious inaccuracy when modeling many real-world problems. In many domains, one has to deal with phenomena such as common causes and effects, where different events are not probabilistically independent. We have already mentioned that, in principle, we can work with arbitrary probability distributions over theory states. An obvious choice to represent such a distribution is by using Bayesian belief networks. Since both Bayesian belief networks and abstract argumentation frameworks are intuitive and easy to understand graphical representations of knowledge, an interesting question is whether the two can be fused in a meaningful way.

6 Related work

Our concept of a probabilistic framework is similar to that of a probabilistic framework, as introduced in [14], where arguments and attacks are associated with probabilities, assuming independence of the events of individual arguments/attacks being active. They do not consider rule-based instantiation of arguments. They apply their formalism to the problem of coalition forming and present an efficient evaluation algorithm, based on Monte Carlo simulation. Our work shows that, in the instantiated setting, it is in general not possible to assume independence between different arguments.

Other related work includes [11], which presents an 'equational approach' to argumentation semantics, where arguments get numerical values $[0, 1]$, satisfying certain equations, such that the labelings of the framework are the solutions to the equations. A similar, but probabilistically oriented approach, is presented in [16], where acceptability of arguments is expressed by probabilities instead of labels. Both approaches subsume the complete semantics (interpreting I and O as a probability or numerical value of 1 and 0, respectively). These two approaches do not focus on uncertainty of the arguments themselves. Rather, they generalize the traditional 'discrete' evaluation of $I/U/O$ labelings.

An interesting approach to probabilistic argumentation, which is not related to Dung-style argumentation theory, is presented in [2]. It models judgment of a hypothesis based on from arguments against the hypothesis, each having a likelihood or reliability based on probabilities.

Also mentioned should be work that approaches the problem from another side, namely by associating numerical weights [10] or fuzzy values [12] with attacks.

7 Conclusion

In this paper we extended the model of Dung-style argumentation with uncertainty, both on the abstract level and the instantiated level. On the instantiated level, we presented an extension of the ASPIC argumentation system, which we call p-ASPIC. This system extends ASPIC in a simple manner, by allowing probabilities associated with rules. This probability is interpreted as being associated with the event that the rule is active. To capture this uncertainty on the abstract level, we extended the notion of an argumentation framework to a probabilistic framework. We have showed how to instantiate a probabilistic framework using the p-ASPIC system, and we demonstrated the system with a number of examples. This work should be understood as an initial step in the direction of probabilistic, instantiated, Dung-style argumentation, and we have discussed a number of directions for future work.

References

1. Leila Amgoud and Claudette Cayrol. Inferring from inconsistency in preference-based argumentation frameworks. *J. Autom. Reasoning*, 29(2):125–169, 2002.
2. B. Anrig, R. Bissig, R. Haenni, J. Kohlas, and N. Lehmann. *Probabilistic argumentation systems: Introduction to assumption-based modeling with ABEL*. Institute of Informatics University of Fribourg, 1999.
3. Pietro Baroni, Martin Caminada, and Massimiliano Giacomin. An introduction to argumentation semantics. *Knowledge Eng. Review*, 26(4):365–410, 2011.
4. R. Booth, S. Kaci, T. Rienstra, and L. van der Torre. Conditional acceptance functions. 2012. To appear.
5. M. Caminada and L. Amgoud. On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171(5):286–310, 2007.
6. Martin Caminada and Leila Amgoud. On the evaluation of argumentation formalisms. *Artif. Intell.*, 171(5-6):286–310, 2007.
7. Martin Caminada and Yining Wu. On the limitations of abstract argumentation. In *BNAIC 2011*. 2011.
8. M.W.A. Caminada and D.M. Gabbay. A logical account of formal argumentation. *Studia Logica*, 93(2-3):109–145, 2009. Special issue: new ideas in argumentation theory.
9. P.M. Dung. On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357, 1995.
10. Paul E. Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artif. Intell.*, 175(2):457–486, 2011.
11. Dov M. Gabbay. Introducing equational semantics for argumentation networks. In Weiru Liu, editor, *ECSQARU*, volume 6717 of *Lecture Notes in Computer Science*, pages 19–35. Springer, 2011.
12. J. Janssen, M. De Cock, and D. Vermeir. Fuzzy argumentation frameworks. In *Information Processing and Management of Uncertainty in Knowledge-based Systems*, pages 513–520, 2008.
13. Souhila Kaci and Leendert van der Torre. Preference-based argumentation: Arguments supporting multiple values. *Int. J. Approx. Reasoning*, 48(3):730–751, 2008.
14. Hengfei Li, Nir Oren, and Timothy J. Norman. Probabilistic argumentation frameworks. In Sanjay Modgil, Nir Oren, and Francesca Toni, editors, *TAFA*, volume 7132 of *Lecture Notes in Computer Science*, pages 1–16. Springer, 2011.
15. Henry Prakken. An abstract framework for argumentation with structure arguments. *Argument and Computation*, 1(2):93–124, 2010.
16. M. Thimm. A probabilistic semantics for abstract argumentation. 2012. To appear.