View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Arrow@dit

Technological University Dublin

ARROW@TU Dublin

Other

Applied Intelligence Research Centre

2019-2

Weakly-admissible Semantics and the Propagation of Ambiguity in Abstract Argumentation Semantics

Pierpaolo Dondio
Technological University Dublin, pierpaolo.dondio@tudublin.ie

Follow this and additional works at: https://arrow.tudublin.ie/aircoth

Part of the Artificial Intelligence and Robotics Commons

Recommended Citation

Dondio, P. (2019) Weakly-admissible Semantics and the Propagation of Ambiguity in Abstract Argumentation Semantics, Technological Report.

This Technical Report is brought to you for free and open access by the Applied Intelligence Research Centre at ARROW@TU Dublin. It has been accepted for inclusion in Other by an authorized administrator of ARROW@TU Dublin. For more information, please contact yvonne.desmond@tudublin.ie, arrow.admin@tudublin.ie, brian.widdis@tudublin.ie, aisling.coyne@tudublin.ie, fiona.x.farrell@tudublin.ie.

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License

Technical report firstly submitted to IJCAI 2019 for revision
Author: Pierpaolo Dondio, pierpaolo.dondio@tudublin.ie

26th February 2019, 10:59am

Weakly-admissible Semantics and the Propagation of Ambiguity in Abstract Argumentation Semantics

Abstract

The concept of *ambiguous* literals of defeasible logics is mapped to the set of undecided arguments identified by an argumentation semantics. It follows that Dung's complete semantics are all *ambiguity propagating*, since the undecided status of an attacking argument is always propagated to the attacked argument, unless the latter is defeated by another accepted argument. In this paper we investigate a novel family of abstract argumentation semantics, called *weakly-admissible* semantics, where we do not require an acceptable argument to be necessarily defended from the attacks of undecided arguments. *Weakly-admissible* semantics are conflict-free, *ambiguity blocking*, non-admissible (in Dung's sense), but employing a more relaxed defence-based notion of admissibility; they allow reinstatement and generate extensions that are super-sets of grounded semantics, and they at least accept credulously what Dung's complete semantics accept at least credulously.

1 Introduction

Abstract argumentation is a framework for non-monotonic reasoning centered on the notion of argumentation framework [Dung, 1995], a directed graph where nodes represent arguments and links represent an attack relation defined over arguments. Given an argumentation framework, various argumentation semantics have been defined to compute the acceptability status of arguments. In the labelling approach [Caminada and Gabbay, 2009], the effect of an argumentation semantics is to assign to each argument of an argumentation framework a label in, out or $undec$, meaning that the argument is accepted, rejected or deemed undecided. The $undec$ label represents a situation in which the semantics can neither accept or reject an argument. It is not an uncertain status, since there is perfect knowledge about arguments and their attacks relations. It is rather an argument that is not clearly defetated, but still not fully defended.

In this paper we explore the definition of a new family of abstract semantics, called *weakly-admissible* and *ambiguity-blocking* semantics (referred to as *ab*-semantics). These semantics are conflict-free and non-admissible semantics, but still employing a defence-based relaxed notion of admissibility; they allow reinstatement and generate extensions that are super-sets of the grounded semantics. Moreover, what is accepted at least credulously by any complete semantics is also accepted credulously by the new semantics. The genesis of such new semantics starts from a reconsideration of the way the undecided label is propagated onto an argumentation framework by abstract semantics. Under any complete semantics, an argument a attacked by an undecided arguments is never accepted. It is either undecided or, if another accepted arguments is also attacking a, it is rejected. Therefore the undecided label, if possible, is always propagated from attacker to attacked argument. In *weakly-admissible* semantics only attacks from in-labelled arguments always prevent the attacked arguments to be accepted, while attacks from undecided arguments could have no effect, blocking the undecided label to spread as it would happen in a complete labelling. The notion of admissibility is relaxed and the notion of effective attack is stronger, being harder for an argument to be excluded from the set of accepted arguments.

We frame our work in the context of *ambiguity propagation* or *ambiguity blocking* in non-monotonic semantics, well studied in defeasible logics (DL). In DL a literals is ambiguous if there are two chain or reasoning concluding a and $\neg a$ and the superiority relation cannot resolve such conflict. The DL definition of ambiguous literal has indeed analogy with the undecided argument in abstract argumentation, since the undecided label is assigned to conflicting arguments whose conflicts cannot be resolved by external attacks. As justified in section 3, our key idea is to define ambiguous arguments as undecided arguments.

Therefore in the context of this paper the problem of *ambiguity blocking* in abstract argumentation semantics become the problem of *undecidedness blocking*. Dung's complete semantics are all *ambiguity propagating*, and very few examples of *ambiguity blocking* in Dung's framework exists, as discussed in section 5.

Regarding the motivation behind *yet a new semantics*, we claim how *ambiguity blocking* is a core reasoning mechanism in defeasible reasoning that should be captured by abstract argumentation semantics. Abstract argumentation frameworks are indeed simplifying various aspects of argumentation, but it has proven to be a useful formalism to study various non-monotonic systems and their core mechanisms such as attack, defense, restatement, acceptability of arguments, all captured in the definition of different semantics. Despite ambiguity blocking semantics being present in many non-monotonic systems, such as DL, they are almost absent in Dung-like abstract semantics, and we believe this study is a contribution in that direction.

Moreover, there are plenty of reasoning patterns, both in formal contexts and informal routine situations, where humans adopt a mechanism where the ambiguity cast by conflicting arguments is confined and not propagated to other parts of the decision making process. In general, a reason for ambiguity blocking is if the penalty for being wrong outweighs the benefit of being right. In the majority of legal systems, evidence in a criminal case has to satisfy the standard of proof *beyond reasonable doubt*. If evidence versus an accused are not definitive or open to multiple interpretation, the judge rules in favor of the accused. A typical example is the situation of two testimonies,

both of them accusing x, but providing conflicting accounts. In an Dung-style argumentation graph (see figure 1), the situation can be modelled as two rebuttal arguments (the testimonies) attacking the presumption of innocence of x. None of the complete semantics accepts the innocence of x and preferred and stable semantics sceptically reject x's innocence. In this situation an *ambiguity blocking* semantics is more appropriate, since in a legal context the judge will consider x innocent *indisputably*.
We do not need to be confined to legal courts to find *ambiguity blocking*. An example is the *benefits of the doubt*, a common pattern used routinely by humans. When we reason using the *benefit of the doubt*, we tend to believe something even if we are not certain about it. Here is an example. Susan was late at school this morning and she should get extra homeworks for that. Susan blamed the traffic, but no other students were late. Despite her explanation is weak, the teacher believed her, since she has been always a very good student. The *benefit of doubt* is deeply embedded in human relationships, often granted to a person based on trust and instrumental to build and maintain mutual trust (see [Zaheer *et al.*, 1998]).

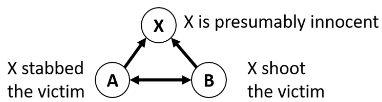

Figure 1: The floating assignment in a legal context

The paper is organized as follows. The next section introduces the required background of abstract argumentation. Section 3 introduces our semantics, while section 4 discusses their properties with proofs and examples. Section 5 contains related works before our conclusions.

2 Abstract Argumentation.

In this section we describe the required concepts of abstract grounded and preferred semantics introduced by Dung [**?**].

Definition 1. An argumentation framework AF is a pair $\langle Ar, \mathcal{R} \rangle$, where Ar is a non-empty finite set whose elements are called arguments and $\mathcal{R} \subseteq Ar \times Ar$ is a binary relation, called the attack relation. If $(a, b) \in \mathcal{R}$ we say that a attacks b. Two arguments a, b are rebuttals iff $(a, b) \in \mathcal{R} \wedge (b, a) \in \mathcal{R}$, i.e. they define a symmetric attack. An argument a is initial if it is not attacked by any arguments, including itself.

An abstract argumentation semantics identifies a set of arguments that can survive the conflicts encoded by the attack relation $\mathcal{R}$. Dung's semantics require a group of acceptable arguments to be conflict-free (we cannot accept at the same time an argument and its attacker) and admissible (the set of arguments defends itself from external attacks).

Definition 2. (conflict-free). A set $Arg \subseteq Ar$ is conflict-free iff $\forall a, b \in Arg, (a, b) \notin \mathcal{R}$.

Definition 3. (admissible set, complete set). A set $Arg \subseteq Ar$ defends an argument $a \subseteq Ar$ iff $\forall b \in Ar$ such that $(b, a) \in \mathcal{R}, \exists c \in Arg$ such that $(c, b) \in \mathcal{R}$. The set of arguments defended by Arg is denoted $\mathcal{F}(Arg)$. A conflict-free set Arg is *admissible* if $Arg \subseteq \mathcal{F}(Arg)$ and it is *complete* if $Arg = \mathcal{F}(Arg)$

We follow the labelling approach of [Caminada and Gabbay, 2009], where a semantics assigns to each argument a label in, out or $undec$.

Definition 4. (labelling). Let $AF = (Ar, \mathcal{R})$. A labelling is a total function $\mathcal{L} : Ar \rightarrow \{in, out, undec\}$. We write $in(\mathcal{L})$ for $\{a \in Ar | \mathcal{L}(a) = in\}$, $out(\mathcal{L})$ for $\{a \in Ar | \mathcal{L}(a) = out\}$, and $undec(\mathcal{L})$ for $\{a \in Ar | \mathcal{L}(a) = undec\}$.

Definition 5. (from [Caminada and Gabbay, 2009]). Let $AF = (Ar, \mathcal{R})$. A **complete labelling** is a labelling such that for every $a \in Ar$ holds that:

1. if a is labelled in then all its attackers are labelled out;
2. if a is labelled out then it has at least one attacker that is labelled in;
3. if a is labelled $undec$ then it has at least one attacker labelled $undec$ and it does not have an attacker that is labelled in.

Definition 6. (grounded and preferred labelling [**?**]) Given $AF = (Ar, \mathcal{R})$, $\mathcal{L}$ is the grounded labelling iff $\mathcal{L}$ is a complete labelling where $undec(\mathcal{L})$ is maximal (w.r.t. set inclusion) among all complete labellings of AF. $\mathcal{L}$ is the preferred labelling iff $\mathcal{L}$ is a complete labelling where $in(\mathcal{L})$ is maximal (w.r.t. set inclusion) among all complete labellings of AF.

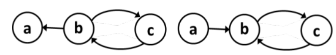

Figure 2: Two Argumentation Graphs G_1 (left) and G_2 (right)

Referring to Figure 2, the grounded labelling assigns the $undec$ label to all the arguments of G_1. Regarding the preferred semantics, there are two complete labellings that maximise the $in(\mathcal{L})$ set: one with $in(\mathcal{L}_1) = \{b\}$, $out(\mathcal{L}_1) = \{a, c\}$, $undec(\mathcal{L}_1) = \emptyset$ and the other with $in(\mathcal{L}_2) = \{a, c\}$, $out(\mathcal{L}_2) = \{b\}$, $undec(\mathcal{L}_2) = \emptyset$. Regarding G_2, there is only one complete labelling (thus representing both the grounded and preferred labelling), where argument a is in (no attackers), b is out and c is in. Note how a reinstates c.

For the rest of the discussion, we need to define the *topological ordering* of a graph. A topological ordering of a graph G is an ordering such that, $\forall a, b \in G$, if $R(a, b)$ then $a \succ b$. If a belongs to $S_1 \in G_{scc}$ and b belongs to $S_2 \in G_{scc}$ and $S_1 \succ S_2$ in the topological ordering of G_{SCC}, then $a \succ b$ in the topological order of graph G.

3 Weakly-Admissible Semantics and ambiguity blocking in Dung's framework

Our goal is to define an abstract semantics where the propagation of undecidedness is controlled by the postulates of the semantics, and it does not necessarily follow as it happens with Dung's complete semantics. Our discussion is framed around the problem of *ambiguity propagation* and *blocking* in non-monotonic reasoning. Despite little has been done about *ambiguity blocking* in abstract argumentation, the problem of *ambiguity blocking* and *propagation* has been well-studied in defeasible logics. Since the standard DL semantics is *ambiguity blocking*, researchers have investigated an *ambiguity propagation* version of such semantics (see [Stein, 1992],[Maier and Nute, 2006]). In this work, starting from Dung's framework, we go the opposite way and propose *ambiguity blocking* abstract argumentation semantics.

Let us consider the following defeasible theory D, that is a finite set of defeasible rules and a superiority relation R. The rule $x \Rightarrow y$ means that x defeasibly implies y. A conflict happens when two rules support complementary literals. The superiority relation R over the set of rules define what rule wins in case of conflicts.

$$D = \{\Rightarrow a, \Rightarrow \neg a, \Rightarrow b, a \Rightarrow \neg b\}, R = \emptyset$$

In DL, a literal a is ambiguous iff there exist two chains of reasoning and one supports the conclusion a, while the other supports the conclusion $\neg a$, and the superiority relation does not resolve the conflict. In D there are rules for concluding both a and $\neg a$, and therefore a is ambiguous. In the standard DL semantics both a and $\neg a$ are refuted (i.e. they cannot be proven) and therefore $\neg b$ cannot be proven using a $\Rightarrow \neg b$, since the antecedent of the rule a is refuted. Therefore $\neg b$ is provable and not ambiguous: the ambiguity of a is not propagated to b. In an *ambiguity propagation* semantics, both b and $\neg b$ can be proved and b result ambiguous: the ambiguity of a is transferred to b.

The standard DL semantics blocks the ambiguity by refuting the ambiguous literals. Those literals cannot be used anymore for further derivations. In a Dung-like framework, our proposal is to define as ambiguous an argument labelled undecided. Therefore, the *ambiguity blocking* semantics proposed in this study are semantics blocking undecidedness. In this sense, all complete semantics are *ambiguity propagating* (as already noticed in [Governatori *et al.*, 2004]). There is a strong analogy between undecided arguments and ambiguous literals in DL. Undecided arguments are either involved in an unresolved conflict, or attacked by those arguments. Referring to figure 2, graph G_1, a and b are responsible for generating the undecided situations, as the literals a and $\neg a$ are ambiguous in the above DL theory. In $G_1 1$ the undecided label is propagated to argument c, as the literals b and $\neg b$ are ambiguous in the *ambiguity propagating* version of the DL semantics for D.

If we want to block ambiguity in abstract argumentation, we need to block the undecided label to propagate over the argumentation graph.

We propose to consider attacks from undecided arguments not enough to remove the attacked argument from the set of accepted argument. In other words, an attack from an undecided argument is *discarded* because ambiguous, and its undecidedness does not affect anymore attacked arguments, as the literals a and $\neg a$ are both refuted in the standard ambiguity blocking DL semantics to prevent the ambiguity of a to spread.

Note how our definition of ambiguity depends on the semantics. Grounded semantics is the semantics generating the largest set of ambiguous arguments, while the same set of arguments could not be considered ambiguous by other semantics.

Our *weakly admissible* semantics are therefore based on a relaxed notion of admissibility, where arguments could be accepted also if there are $undec$ arguments attacking them. As in the complete semantics, an argument is rejected iff it has at least one in-labelled attacker. Some arguments that would be undecided in complete semantics can therefore be promoted to the label *in*, but only if the change generates a legal labelling. We formally define *weakly admissiblee* legal labellings and the corresponding semantics in the following way:

Definition 7. Given an argumentation framework $AF = \langle Ar, \mathcal{R} \rangle$, a weakly-admissible labelling is a labelling such that for every $a \in Ar$ it holds that:

1. if a is labelled in then there is no attackers of a labelled in;
2. if a is labelled out then it has at least one attacker that is labelled in;
3. if a is labelled $undec$ then it has at least one attacker labelled $undec$ and it does not have an attacker that is labelled in.

The above definition changes condition 1 (definition 5) of Dung's complete labelling by relaxing it, since now an argument attacked by undecided arguments could be accepted. The key idea is that, when it is *legally* possible, undecided arguments do not remove attacked arguments from the set of accepted arguments. Legally possible means that the above conditions are all satisfied, that imply the conflict-free and reinstatement properties. We break down condition 1 of definition 7 into two conditions:

- a is labelled in if all the attackers of a are labelled out (*complete semantics condition*);
- a is labelled in if there is at least one attacker of a labelled $undec$ and all the other attackers are labelled out (*ambiguity blocking condition*);

Weakly-admissible legal labellings generated using exclusively the *complete semantics condition* are complete, and therefore *weakly-admissible* semantics represent a non-admissible super-set of complete semantics. We call the second condition the *ambiguity blocking* condition (shortened in *ab*-condition), since it is the condition that allows an attacked argument a to be accepted if attacked by undecided arguments. Note how argument a would be otherwise labelled $undec$ (condition 3 of definition 7).

In general, there are legal *weakly-admissible* labellings where only the *complete semantics condition* is used (therefore these are also complete labellings); labellings where only the *ab*-condition is used and labellings where a combination of the two is used. These are the labellings where the *ab*-condition is applied to some part of the argumentation graph but not to all of it. These labellings represent interesting cases, where an agent might grant *ambiguity blocking* to some arguments, preventing them to be labelled undecided, but not to others. An example is shown in Figure 3. Both of the labellings are *weakly-admissible* and not complete. In the labelling on the left the *ab*-condition is applied *earlier* to the attack from b to d, while on the right the condition is applied to the attack from d to e, but not to the attack from b to d. We are interested in isolating the *weakly-admissible* semantics where the *ambiguity blocking* condition is always used, every time it is possible to do so. This subset of *weakly-admissible* semantics is called *ab*-semantics.

Definition 8. Given an argumentation framework $AF = \langle Ar, \mathcal{R} \rangle$, an *ab*-labelling is a *weakly-admissible* labelling where no undecided argument can be promoted to an in-labelled argument using the *ab*-condition without generating an illegal labelling.

In a *ab*-labelling the *ab*-condition is used *as soon and as much as possible*. In figure 3, only the one on the left is a *ab*-labelling, since we cannot apply the principle further and generate other valid labellings. The labelling on the right is not an *ab*-labelling, since the*ab*-condition could be applied to argument d, that in this case would generate again the labelling on the left.

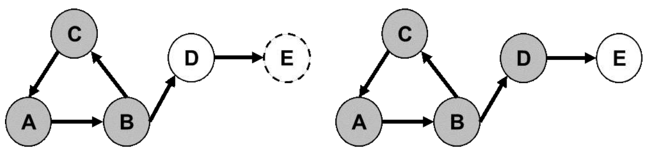

Figure 3: Two *weakly-admissible* labellings. Only the one on the left is an *ab*-labelling.

Figure 4 illustrates examples of *weakly-admissible* and *ab*-labellings. In the floating assignment example (4.3), there are four *weakly-admissible* labellings: the grounded, the two *preferred* ones (that are also *ab*-labelling) and also an *ab*-labelling where argument c is accepted. In Figure 4.2, a is labelled $undec$ and b is labelled in in the only *ab*-labelling. In Figure 4.4, there is one grounded labelling with all arguments $undec$, one *preferred* labelling (also *ab*) and an additional *ab*-labelling where the odd-length cycle a, b, c is still undecided but, by applying the *ab*-condition on the attack from b to d, d is accepted and e is rejected.

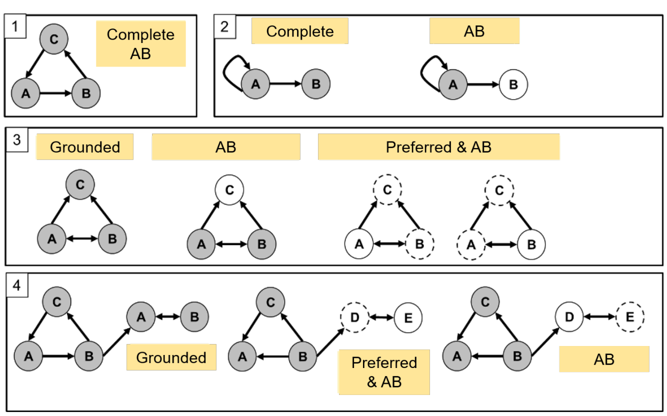

Figure 4: Complete semantics and ab-semantics labellings.

The *ab*-condition can be seen as an additional condition that a reasoner using one of Dung's complete semantics might use to make part of its labelling *ambiguity blocking*. In general, the less undecided arguments a semantics generates, the less options to apply the *ab*-condition are present. In the floating assignment example, an *ambiguity blocking grounded* reasoner would consider the rebuttal attacks a and b as a situation where the *ab*-condition can be used, while a *preferred* reasoner would not perceive the situation as doubtful, and it will find only reasons to reject c. The extreme case is represented by *stable* labellings, where the set of undecided arguments is empty. According to our definition, all stable labellings are *ab*-labellings. In these labellings the semantics does not identify any doubtful situation arising from graph, and therefore the *ab*-condition has no *raison d'être*. In other words, an agent $\mathcal{A}$ adopting a less prudent semantics might not consider certain situations doubtful and therefore, even if $\mathcal{A}$ is keen to block ambiguity, it does not find any ambiguity to be blocked.

Table 1: Properties of *weakly admissible* and *ab*-semantics

Property	*Weakly Admissible*	*ab*	*Complete*
Admissible	No	No	Yes
Conflict-free	Yes	Yes	Yes
Reinstatement	Yes	Yes	Yes
Rejection	Yes	Yes	Yes
Directionality	Yes	Yes	Yes
Abstention	Yes	No	Yes
Cardinality	≥ 1	≥ 1	≥ 1
I-maximality	No	No	No

4 Discussion and Properties

Weakly admissible semantics and *ab*-semantics satisfy the properties illustrated in Table 1.

Weakly-admissible semantics are clearly non-admissible, since they can accept arguments not defended by in-labelled arguments. They could be seen as employing a different form of admissibility since we still require an argument to be defended, but, in some cases, not from undecided arguments. Both *weakly-admissible* and *ab*-semantics satisfy the reinstatement property since, if an argument has all its attackers labelled out, it is labelled in. However, *weakly-admissible* and *ab*-semantics make reinstatement easier, since an argument a defeated by b is fully reinstated even by an argument c rebutting b, since the doubt cast by c's attack is enough to balance the attack from b to a. Our semantics are the only non-admissible semantics known to the authors satisfying reinstatement. Other non-admissible semantics, such as $Stage$ semantics, allow an initial argument to be excluded from the extension, while both $CF1$ and $CF2$ semantics allow an argument whose attackers are all labelled out to be labelled out.

Weakly-admissible and *ab* semantics are also conflict-free, they satisfy rejection (condition 2 of definition 7) and they are multiple-status semantics, with cardinality greater or equal than 1. Regarding directionality, both the semantics satisfy it, since the label assigned to an argument a does not depend on the label of arguments following a in the topological order of the graph.

Regarding abstention, the property says that, if an argument a is labelled out in at least one valid labelling, and labelled in in at least another, then there must be a valid labelling where a is labelled $undec$. Complete semantics satisfies it. Since the *ab*-condition changes the label of an argument a for which there is a valid complete labelling where a is labelled $undec$, if an argument is labelled in in one *weakly-admissible* labelling and out in another, there is also a complete labelling (and therefore *weakly admissible*) where a was $undec$, and abstension is satisfied. This is not the case for *ab*-semantics: a counter-example is in figure 4.4 for arguments e and d.

A semantics satisfies I-maximality if no extension is a strict subset of another. It is not satisfied by *weakly-admissible* semantics since it is not satisfied by *complete* semantics. *ab*-semantics does not satisfy it neither. For instance, if we consider the graph with argument a rebutting b and b attacking c, there are three valid *ab*-labellings. The set of in-labelled arguments is $\{a, c\}$ in the first, $\{b\}$ in the second (both preferred labellings as well) and $\{c\}$ in the third *ab*-labelling (that is not *complete*).

4.1 Relation between Complete and ab- acceptability

It is interesting to study the relation between the set of in-labelled arguments of a complete semantics x and the set of in-labelled arguments of the *ab*-semantics. In an *ab*-labelling some of the arguments labelled $undec$ in x could be promoted to the label in, and those arguments are now free to attack other arguments. We wonder if some of the arguments accepted by x are now discarded by the *ab*-semantics. It holds that:

Theorem 1. If an argument is at least credulously accepted by semantics x, it is at least credulously accepted by the *ab*-semantics.

Proof. We prove that, if an argument a is labelled in in a complete labelling l, there is also a labelling l_{ab} where a is labelled in. We first notice that arguments labelled in in a complete labelling l are indifferent to undecided arguments. They are either initial arguments or defended by some in-labelled arguments (potentially including themselves). The same is for arguments labelled out in l: their label is assigned by the presence of an in-labelled argument attacking them. Moreover, in a complete labelling l, in-arguments do not receive any attacks from $undec$ arguments, and $undec$ arguments only attack arguments labelled $undec$ or out. In a *ab*-labelling, attacks from a subset of undecided arguments are *de facto* neglected. There are two cases:

Case 1. The neglected attacks are directed to $undec$-labelled arguments. In this case, the attacked arguments could be promoted to the label in. However, each new in-labelled argument b does not attack any in-labelled argument in l, but only arguments labelled $undec$ and out, since b was undecided in l. In turn, arguments attacked by b only attacks argument out or $undec$, and so on. Therefore the only potential effect of the attacks from b is that some arguments labelled $undec$ in l are now labelled out or in in l_{ab}, and therefore $in(l) \subseteq in(l_{ab})$.

Case 2. The neglected attacks are directed to arguments labelled out in l. In this case, the effect is that each attacked argument c remains labelled out also in l_{ab}, since the out label of c in l was necessarily the effect of the attack from an in-labelled argument. This in-labelled argument is still labelled in in l_{ab}, since it cannot be affected by attacks included in case 1 above, and it is not affected by attacks included in case 2, since all the attacked arguments c remain labelled out, and therefore they do not affect any in-labelled arguments in l, and therefore $in(l) \subseteq in(l_{ab})$.

Theorem 1 implies that every *ab*-labelling is a super-set of *grounded* labelling. Moreover, the above proof shows how, in an *ab*-labelling, when the label of undecided arguments is modified by the *ab*-condition, only the label of other undecided arguments is modified, while in and out arguments are not affected. This implies that not only every *ab*-labelling is a super-set of grounded semantics, but that $undec_{ab} \subseteq undec_{gr}$, $in_{gr} \subseteq in_{ab}$ and $out_{gr} \subseteq out_{ab}$. It can be easily proven that that the above property is valid for to the relations between *weakly admissible* labellings and grounded labelling as well. We can also prove the following:

Lemma 1. The grounded labelling is the weakly-admissible labelling maximizing the set of undecided arguments (w.r.t set inclusion).

The lemma can be proven by observing that grounded semantics is a complete labelling maximizing the set of undecided arguments, and that the *weakly-admissible* labellings that are not complete labellings are generated by applying the *ab*-condition at least once. Therefore in these *weakly-admissible* not *complete* labellings the label of some undecided arguments under grounded semantics is changed to in or out, and consequently the set of undecided arguments of *weakly-admissible* semantics becomes a subset of the set of undecided arguments of grounded semantics.

The following lemma is interesting to show the effect of the *ab*-condition on the propagation of the undecided label over an argumentation graph.

Lemma 2. If argument b is labelled $undec$ in a *complete* labelling, and the $undec$ attackers of b are all preceding b in the topological order of the graph, then b is labelled in or out in all the *ab*-labellings.

Proof. Omitted (available on request)

As a corollary, if an argument a is not part of a cycle, then a is labelled in or out in all *ab*-labellings. This means that acyclic parts of the graph do not have $undec$ arguments, and the $undec$ label is not propagated outside the strongly connected component where it was generated. Therefore, the *ab*-condition underlying *ab*-semantics and *weakly-admissible* semantics can be seen as a mechanism to control the propagation of undecidedness over an argumentation graph, without changing the status of in- and out-labelled arguments.

5 Related Works

In defeasible logics the problem of *ambiguity blocking* and *propagation* has been extensively studied, as already discussed in section 3. In particular the work by [Governatori *et al.*, 2004] is the most relevant, since the authors propose an *ambiguity propagating* defeasible logics, and they provide a Dung-like argumentative version of both the standard (*ambiguity blocking*) and their *ambiguity propagating* semantics. Authors do not start from a modification of Dung's abstract semantics, but they propose a rule-based model of arguments using DL rules, and they study such structured argumentation systems using Dung's notion of acceptability. They show how their DL *ambiguity propagating* semantics can be obtained from Dung's grounded semantics postulates. However, the argumentative version of the standard DL *ambiguity blocking* semantics generates a new *ambiguity blocking* Dung-like argumentation semantics.

We wonder how similar is such *ambiguity blocking* semantics to ours. As their semantics is for a structured rule-based argumentation system, we need to map some of their concepts to the abstract framework. Their semantics is a two-state (accepted/rejected) semantics; there are indeed arguments that cannot be labelled neither accepted or rejected, but they do not correspond to the notion of undecided arguments, but rather to infinite undercutting arguments. Arguments built using ambiguous literals are marked as rejected. It is indeed easier for an argument to be accepted, since ambiguous literals that might prevent those argument to be accepted are rejected. The absence of a notion of undecidedness implies that, in Dung's term, the semantics would lose the reinstatement property, since a pair of otherwise unattacked rebuttal arguments are both labelled out. The authors show how grounded semantics accepts less arguments than the standard DL-based semantics, in accordance with our theorem 1, but it also rejects more than the standard DL-based semantics, in disagreement with our theorem 1. It is interesting to note that their *ambiguity blocking* semantics is obtained by relaxing the notion of acceptability. An argument is accepted not if all its attackers are attacked (as in grounded semantics) but it is enough if the attackers are attacked by arguments supported by a set of accepted arguments. Support is weaker than acceptance, since a supported argument is not necessarily accepted, but every accepted argument is supported. Even if the notion of supported arguments does not match our *ambiguity blocking* condition based on undecided arguments, we note how there are conceptual similarities: less effort is required to accept an argument, since the defensive counter-attack can come from an argument that is not fully accepted (as it happens in *weakly admissible* semantics when an argument is reinstated by an undecided argument).

In a recent study by [Dondio, 2019], the author investigates how to embed the *in dubio pro reo* principle into abstract argumentation semantics. As the *in dubio pro reo* is an example of *ambiguity blocking*, the paper is a rare example of abstract semantics where ambiguity is blocked. The author proposes a topology-based criterium to block undecidedness, obtaining that the undecided label cannot be propagated outside the strong connected component where it was generated. Their semantics, called SCC-semantics, are a subset of our *weakly admissible* semantics, and they intersects our *ab-semantics*. The main difference with our approach is that we apply ambiguity blocking inside a strong connected component as well, and we do not require any topological order constraints. Their semantics are *partially ambiguity blocking* semantics, that can be obtained from a *weakly admissible* labelling where the ab-condition is used only in the acyclic part of the argumentation graph. Other studies in the larger context of Argumentation Theory have investigated *ambiguity blocking* mechanisms. Different standard of proof have been extensively study in argumentation theory (see [Gordon and Walton, 2009]), but only few studies are relevant to abstract argumentation. In the context of structured argumentation, we mention the work by [Prakken and Sartor, 2011] on modelling *standards of proof*, and the modification of the [Brewka and Gordon, 2010]. Regarding abstract argumentation, the most explicit study about standard of proof is [Atkinson and Bench-Capon, 2007]. Here the authors consider how each Dung's semantics has a different level of *cautiousness* that is mapped to a corresponding legal *standard of proof*. Only initial arguments are beyond doubt, but they consider the skeptically preferred justification a beyond reasonable doubt position. In the floating assignment example (Figure 4.3), the authors recognize the two attackers as doubtful, but they consider the skeptically preferred rejection of c beyond reasonable doubt. It could be noticed that this position is failing to acknowledge that, if each of the attackers are considered doubtful, their effect cannot be (at last in all the situations) beyond doubt. [Brewka and Gordon, 2010] also criticises [Atkinson and Bench-Capon, 2007], since they doubt the fact that various Dung's semantics can capture the intuitive meaning of legal standard of proof (detailed discussion in here [Gordon and Walton, 2009]). In case of beyond reasonable doubt, we agree with Brewka: complete Dung's semantics are not adequate to model this principle. Prakken has analysed the floating assignment and its link to *standard of proof* in his work [Prakken, 2002], where he responds to objections advanced by [Horty, 2001]. Prakken underlines that in many problematic situations, including the floating assignment, there could be hidden assumptions about the specific problem which, if made explicit, are nothing but extra information that defeat the defeasible inference. In the case of the floating assignment, Prakken agrees that if beyond reasonable doubt is our standard of proof (like in a criminal case where there are two conflicting testimonies) we should not conclude that the accused is guilty. However, this does not mean that argumentation semantics are somehow invalid. In the case of conflicting testimonies, as already showed by Pollock [Pollock, 1995], the situation could be correctly modelled by making explicit some hidden assumptions and adding extra arguments to model such assumptions. In the conflicting testimonies, the fact that two witnesses contradict each other is a reason to add an argument undercutting the credibility of both. However, the problem of when to add arguments and how they interact with existing arguments has still to be faced, and in this work we have tackled it by embedding assumptions in an abstract argumentation semantics rather than adding arguments.

6 Conclusions

In this paper we investigated a novel family of abstract argumentation semantics, called *weakly-admissible* semantics, where we do not require an acceptable argument to be necessarily defended from the attacks of undecided arguments. We showed how these semantics retain the large majority of desirable properties: they are are conflict-free, non-admissible (in Dung's sense), but employing a defence-based relaxed notion of admissibility, they are *ambiguity blocking* semantics, they allow reinstatement and generate extensions that are super-sets of grounded semantics, and they at least accept credulously what Dung's complete semantics accept at least credulously. We have also provided numerous theoretical and practical examples. We believe to have proposed a novel and well-motivated contribution to abstract argumentation semantics and make a substantial contribution to the definition of *ambiguity blocking* Dung-like abstract semantics.

References

[Atkinson and Bench-Capon, 2007] Katie Atkinson and Trevor Bench-Capon. Argumentation and standards of proof. In *Proceedings of the 11th international conference on Artificial intelligence and law*, pages 107–116. ACM, 2007.

[Brewka and Gordon, 2010] Gerhard Brewka and Thomas F Gordon. Carneades and abstract dialectical frameworks: A reconstruction. In *Proceedings of the 2010 conference on Computational Models of Argument: Proceedings of COMMA 2010*, pages 3–12. IOS Press, 2010.

[Caminada and Gabbay, 2009] Martin WA Caminada and Dov M Gabbay. A logical account of formal argumentation. *Studia Logica*, 93(2):109–145, 2009.

[Dondio, 2019] Pierpaolo Dondio. A proposal to embed the in dubio pro reo principle into abstract argumentation semantics based on topological ordering and undecidedness propagation. In *Proceedings of the 2nd Workshop on Advances In Argumentation In Artificial Intelligence, co-located with XVII International Conference of the Italian Association for Artificial Intelligence (AI*IA 2018)*, 2019.

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

[Gordon and Walton, 2009] Thomas F Gordon and Douglas Walton. Proof burdens and standards. In *Argumentation in artificial intelligence*, pages 239–258. Springer, 2009.

[Governatori *et al.*, 2004] Guido Governatori, Michael J Maher, Grigoris Antoniou, and David Billington. Argumentation semantics for defeasible logic. *Journal of Logic and Computation*, 14(5):675–702, 2004.

[Horty, 2001] John F Horty. Argument construction and reinstatement in logics for defeasible reasoning. *Artificial intelligence and Law*, 9(1):1–28, 2001.

[Maier and Nute, 2006] Frederick Maier and Donald Nute. Ambiguity propagating defeasible logic and the well-founded semantics. In *European Workshop on Logics in Artificial Intelligence*, pages 306–318. Springer, 2006.

[Pollock, 1995] John L Pollock. *Cognitive carpentry: A blueprint for how to build a person*. Mit Press, 1995.

[Prakken and Sartor, 2011] Henry Prakken and Giovanni Sartor. On modelling burdens and standards of proof in structured argumentation. In *JURIX*, pages 83–92, 2011.

[Prakken, 2002] Henry Prakken. Intuitions and the modelling of defeasible reasoning: some case studies. *arXiv preprint cs/0207031*, 2002.

[Stein, 1992] Lynn Andrea Stein. Resolving ambiguity in nonmonotonic inheritance hierarchies. *Artificial Intelligence*, 55(2-3):259–310, 1992.

[Zaheer *et al.*, 1998] Akbar Zaheer, Bill McEvily, and Vincenzo Perrone. Does trust matter? exploring the effects of interorganizational and interpersonal trust on performance. *Organization science*, 9(2):141–159, 1998.