THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : Université Toulouse III Paul Sabatier (UPS)

Discipline ou spécialité : Informatique

Présentée et soutenue par :

Srdjan VESIC

Le vendredi 15 Juillet 2011

Titre :

Les systèmes d'argumentation basés sur les préférences : application à la décision et à la négociation

---

Directeurs de thèse :

Leila AMGOUD
Directrice de recherche CNRS
Université Paul Sabatier

Philippe BESNARD
Directeur de Recherche CNRS
Université Paul Sabatier

Rapporteurs :

Gerhard BREWKA
Professeur
Université de Leipzig

Pierre MARQUIS
Professeur
Université d'Artois

Michael WOOLDRIDGE
Professeur
Université de Liverpool

Examinateurs :

Henri PRADE
Directeur de Recherche CNRS
Université Paul Sabatier

Gilles RICHARD (Président)
Professeur
Université Paul Sabatier

École doctorale :

Mathématiques Informatique Télécommunications (MITT)

Unité de recherche :

Institut de recherche en Informatique de Toulouse (IRIT)

Preference-based argumentation frameworks: application in decision making and negotiation

Thesis presented and defended by

Srdjan Vesic

On the 15th of July 2011

To obtain the degree of

Doctorat de l'Université de Toulouse

Delivered by: Université Toulouse III Paul Sabatier (UPS)

Speciality: Computer Science

Advisers:

Leila Amgoud	Philippe Besnard
Research Director at CNRS	Research Director at CNRS
University of Toulouse III	University of Toulouse III

Reviewers:

Gerhard Brewka	Pierre Marquis	Michael Wooldridge
Professor	Professor	Professor
University of Leipzig	University of Artois	University of Liverpool

Members:

Henri Prade	Gilles Richard (President)
Research Director at CNRS	Professor
University of Toulouse III	University of Toulouse III

École doctorale: Mathematiques Informatique Télécomunications de Toulouse (MITT)

Unité de recherche: Institut de Recherche en Informatique de Toulouse (IRIT)

Acknowledgments

I would like to thank all the people who helped me during my thesis.

I thank Leila, for her energy, for all discussions, her never-ending passion for research and her countless ideas, and for always helping me in every stage of my thesis. Also, I thank Leila for listening to me and for being a true friend. I also thank Philippe, for having accepted me as a PhD student, for he has shared the knowledge with me, for his assistance and all his useful advices.

I would like to express my deepest gratitude to Gerhard Brewka, Pierre Marquis and Michael Wooldridge, for making me the honour of accepting to review my thesis.

It should be noticed that this thesis would have not been possible without my scholarship (*allocation de recherche*), which allowed me to stay completely focused on my research during those three years.

To everybody in our two research teams, ADRIA (formerly RPDMP) and LILAC from the third floor of IRIT 1 who made me feel like at home during our lunches and coffee breaks, to Bilal and Marwa with whom I became friends from the first day when I met them, François and Nadine for the good mood and all the small ideas that made our every-day life full of surprises, to Pierre, Emmanuel, Fahima, Mounira, Dídac, Sihem, Pablo, ...

I would also like to thank my family, my parents Jasmina and Nebojša for their support and for always motivating me to discover the world of science and mathematics, and my brother Jovan who is a very special person in my life.

A special thanks to Marija, my love, my wife, for her constant support, patience and understanding during my master and my PhD thesis, for listening about every single idea and article, and for being always there for me.

Abstract:

Argumentation is a promising approach for reasoning with uncertain or incoherent knowledge or more generally with common sense knowledge. It consists of constructing arguments and counter-arguments, comparing the different arguments and selecting the most acceptable among them.

This thesis contains three main parts. The first one concerns the notion of *equivalence* between two argumentation frameworks. We studied two families of equivalence: *basic* equivalence and *strong* equivalence. We proposed different equivalence criteria, investigated their links and showed under which conditions two frameworks are equivalent w.r.t. each of the proposed criteria. The notion of equivalence is then used in order to compute the *core(s)* of an argumentation framework. A core of a framework is its compact version, i.e. an equivalent sub-framework. Hence, instead of using an argumentation framework which may be infinite, it is sufficient to consider one of its cores, which is usually finite. This core determines the stability of the status of each argument.

The second part of the thesis concerns the use of *preferences* in argumentation. We investigated the roles that preferences may play in an argumentation framework. Two particular *roles* were identified: i) to privilege strong arguments over weaker attacking arguments when computing the *standard* solutions of a framework, and ii) to refine those standard solutions. We showed that the two roles are completely independent and require different procedures for modeling them. Besides, we showed that almost all the existing works have tackled only the first role. Moreover, the proposed approaches suffer from a drawback which consists of returning conflicting extensions. We proposed a general approach which solves this problem and which presents two novelties: First, it takes into account preferences at a semantic level, i.e. it defines new acceptability semantics which are grounded on attacks and preferences between arguments. Second, a semantics is defined as a dominance relation that compares any pair of subsets of arguments.

The third part illustrates our preference-based argumentation frameworks (PAF) in case of *decision making* and *negotiation*. We proposed an instantiation of our PAF which rank-orders options in a decision making problem. Then, we studied the dynamics of this model. More precisely, we showed how the ordering on options changes in light of a new argument. We also used our PAF in order to show the benefits of arguing in negotiation dialogues. For that purpose, we proposed an abstract framework for argument-based negotiation, investigated the different types of solutions that may be reached in such dialogues, and showed for the first time under which conditions arguing is beneficial during a negotiation.

Contents

Begin at the beginning and go on till you come to the end; then stop.

The King from *Alice's Adventures in Wonderland*, Lewis Carroll

1 Introduction

Argumentation is a reasoning model based on the construction and evaluation of arguments. An argument gives a reason to believe a statement, to perform an action, to choose an option, etc. The advantage of argumentation is that the reasoning process is composed of modular and intuitive steps, and thus avoids the monolithic approach of many traditional logics. An argumentation process starts with the construction of a set of arguments from a knowledge base. Then, attacks between those arguments are detected. Some argumentation frameworks also allow for specifying intrinsic strengths of arguments (e.g. on the basis of the quality of the information arguments are based on). Those elements are taken into account when determining the subsets of arguments that can be regarded as "acceptable", called *extensions*. The last step consists of analyzing whether a given statement is *justified* (i.e. follows from the knowledge base) or not. For example, this can be the case if every extension contains at least one argument having that statement as its conclusion.

Due to its explanatory power, argumentation has gained increasing interest in Artificial Intelligence. Indeed, argumentation techniques are used for revising information in a knowledge base (e.g. Rotstein et al., 2008), handling inconsistency in knowledge bases (e.g. Simari and Loui, 1992; Besnard and Hunter, 2001, 2008; Amgoud and Cayrol, 2002a; Garcia and Simari, 2004; Governatori et al., 2004), making decisions under uncertainty (e.g. Bonet and Geffner, 1996a; Fox and Parsons, 1997; Gordon and Karacapilidis, 1997; Fox and McBurney, 2002; Amgoud and Prade, 2006, 2009), merging information coming from different sources (e.g. Amgoud and Parsons, 2002; Brena et al., 2005; Amgoud and Kaci, 2007), choosing agents' intentions (e.g. Amgoud, 2003; Atkinson et al., 2004; Rahwan and Amgoud, 2006), and generating agent's goals (e.g. Hulstijn and van der Torre, 2004). Argumentation is also gaining increasing interest in multi-agent systems, namely for modeling *agents' interactions.* Since the seminal book by Walton and Krabbe (1995) on the different categories of dialogues, several argumentation-based systems were proposed for persuasion dialogues (e.g. Amgoud et al., 2000a; Prakken, 2006), negotiation (e.g. Parsons and Jennings, 1996; Kraus et al., 1998; Amgoud et al., 2000b; Amgoud and Prade, 2004; Kakas and Moraitis, 2006), and inquiry dialogues (e.g. Parsons et al., 2003; Black and Hunter, 2007).

This thesis is interested in the study of argumentation frameworks and their applications. It contains three main parts: i) the study of the notion of *equivalence* in argumentation, ii) the integration of *preferences* to argumentation frameworks, and iii) applying argumentation techniques to *decision making* and *negotiation.*

1.1 Equivalence in argumentation

The first part of the thesis studies when two argumentation frameworks are **equivalent**. Such information is useful for different purposes. First, when building an argumentation framework from a given knowledge base, several attack relations may be used. Thus, knowing under which conditions two or more attack relations lead to the same results could be useful. Second, logic-based argumentation frameworks are generally infinite, meaning that their sets of arguments are infinite. It is important to know whether such frameworks can be reduced to finite sub-frameworks. Besides, even in a finite case, building logic-based arguments from a concrete knowledge base is computationally costly. Thus, every decrease in the number of arguments is potentially useful since it reduces the burden of computation.

Despite the obvious benefits of the notion of equivalence, this issue has received little attention in the literature. To the best of our knowledge, the only work on equivalence in argumentation (Oikarinen and Woltran, 2010) is conducted for abstract argumentation frameworks, which means that the structure of arguments is supposed to be unknown. Two categories of equivalence criteria were particularly proposed. The first category (*basic equivalence*) compares directly the outputs of two frameworks (namely their extensions) while the second (*strong equivalence*) compares the outputs of their extended versions (i.e. the frameworks augmented by the same set of arguments). Oikarinen and Woltran (2010) concentrated only on strong equivalence and showed that two frameworks are strongly equivalent if and only if they coincide (i.e. they are identical) except in the rare case when self-attacking arguments are allowed.

In Chapter 3, we study both basic and strong equivalence between logic-based argumentation frameworks. We propose flexible equivalence criteria which take into account the internal structure of arguments. We study the links between those criteria and show under which conditions two frameworks are equivalent w.r.t. each of them. We then use this notion of equivalence in order to define the compact versions of an argumentation framework, called *core(s)*. A core of a framework is an equivalent sub-framework. Hence, instead of using an argumentation framework which may be infinite, it is sufficient to consider one of its cores which are usually finite. Finally, we show that a core of a framework is a threshold under which each argument of the framework has a floating status, and above it the statuses of all arguments become stable.

1.2 Preferences in argumentation

The second part of the thesis concerns the use of **preferences** in argumentation. There is a consensus in the literature that some arguments can be stronger than others and that this should be taken into account when calculating extensions. We show for the first time that there are *two distinct roles* played by preferences in argumentation: i) to protect strong arguments from attacks coming from weaker arguments, and ii) to refine the standard solutions, i.e. to choose the best extensions among those computed using any acceptability semantics. It is worth mentioning that almost all existing works (e.g. Amgoud and Cayrol, 2002b; Bench-Capon, 2003; Modgil, 2009) have modeled the first role except the work by Dimopoulos, Moraitis, and Amgoud (2009) in which the second role was considered but without identifying the nature of this role.

In Chapter 4, we start by showing that these roles are completely independent and require different procedures for modeling them. Then, we show that existing works which tackle the first role suffer from a main drawback which consists of returning conflicting extensions. Then we propose a novel approach which takes into account preferences at a semantic level, i.e. it defines new acceptability semantics which take into account attacks and preferences between arguments. Moreover, a semantics is defined as a dominance relation that compares subsets of arguments. This allows to compare any pair of sets of arguments, contrary to existing acceptability semantics which only separate those sets into two classes: extensions and non-extensions. We propose a framework in which both roles of preferences are modeled. Finally, we show that two instantiations of this framework capture the preferred sub-theories (Brewka, 1989) and democratic sub-theories (Cayrol et al., 1993), which were proposed for handling inconsistency in prioritised knowledge bases.

1.3 Argumentation for decision making and negotiation

The third contribution of the thesis consists of applying our preference-based argumentation framework for making decisions and for negotiation. In a **decision making** context, argumentation has obvious benefits as in everyday life, decisions are often based on arguments and counter-arguments.

Several argument-based decision frameworks were proposed in the literature. However, the dynamics of those frameworks has not received enough attention. In Chapter 5, we study the dynamics of a particular decision framework that is proposed by Amgoud, Dimopoulos, and Moraitis (2008). The framework rank-orders options (or decisions) on the basis of their statuses. The status of an option is based on the quality of its supporting arguments. We study how an option status changes in the light of a new argument. We provide conditions under which an accepted option becomes rejected and vice versa. Our study is undertaken under two acceptability semantics: grounded semantics and preferred one. These results may be used in negotiation dialogues, namely for defining strategies. Indeed, at a given step of a dialog, an agent may choose which argument to send to another agent in order to change the status of an option. Our results may also help to understand which arguments are useful and which ones are useless in a given situation.

Besides, even if it was claimed by many researchers that exchanging arguments may positively influence the quality of a negotiation outcome, this was never formally shown. The reason is that the quality of an outcome is not defined. In Chapter 6, we study the benefits of arguing in **negotiation dialogues**. For that purpose, we start by proposing an abstract framework for negotiation between two agents. Each agent is assumed to be equipped with a decision framework like the one discussed in Chapter 5. This framework is used for evaluating and choosing offers in a negotiation dialogue, and also for evaluating and choosing the arguments to utter in a dialogue. We define different types of solutions that may be reached in such dialogues. Finally, we study the impact of exchanging arguments on the quality of negotiation outcomes.

I dislike arguments of any kind. They are always vulgar, and often convincing.

Lady Bracknell from *The Importance of Being Earnest*, Oscar Wilde

2 Argumentation frameworks

2.1 Argumentation process

Humans engage in *argumentation* in almost all communications. They advance *arguments* and counter-arguments to justify or refute a given standpoint. Before defining what argumentation is, let us start by presenting a short dialogue between two violinists.

David "This violin is expensive since it is a Stradivarius."

Jascha "I do not think that the violin is a Stradivarius."

Here, the first violinist presents a claim which is at the time justified, thus constructing an argument. The other one challenges this justification by another argument.

Argumentation is seen as a reasoning process in which arguments are built and evaluated in order to increase or decrease the acceptability of a given standpoint. In the above dialogue, David aims at increasing the acceptability of his statement, whereas Jascha decreases it by attacking its justification. Argumentation is defined by van Eemeren, Grootendorst, and Snoeck Henkemans (1996) as follows:

> *Argumentation is a verbal and social activity of reason aimed at increasing (or decreasing) the acceptability of a controversial standpoint for the listener or reader, by putting forward a constellation of propositions intended to justify (or refute) the standpoint before a rational judge.*

In the previous definition, argumentation is defined as a verbal activity, since it is supposed to be conducted in a natural language. However, there are many approaches which aim at constructing a computational model of argumentation. In this thesis, we are interested in such models. Argumentation is also defined as a *social activity* since it is directed at other people. It is considered as an *activity of reason* since an argument is supposed to contain somehow rational justification. Argumentation always relates to a particular *opinion*, or *standpoint*, about a specific subject. A subject may be something believed or known (like in the our example) but also an action to perform, a goal to achieve, etc. Finally, argumentation is intended to justify or refute a standpoint.

Researchers in artificial intelligence are interested in building a computational model of argument. In such an approach, an argumentation process starts with construction of arguments from a knowledge base using a given logic, after which interactions between them are identified (e.g. attacks, supports). An intrinsic strength of each argument is also determined, based on quality of information it is built from. Finally, arguments are evaluated and usually extensions of arguments are calculated, where an extension represents a set of arguments that are acceptable together.

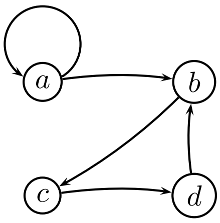
2.2 Dung's abstract argumentation framework

The most *abstract* and *general* argumentation framework in the literature is the one proposed by Dung (1995). It takes as input a set of arguments and a binary relation encoding attacks between arguments. The framework is abstract since neither the structure nor the origin of the two components are specified. Thus, it can be instantiated in different ways. The framework is general since no particular constraints are imposed on arguments or attacks.

Definition 2.2.1 (Argumentation framework)**.** *An* argumentation framework *is a pair $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, where $\mathcal{A}$ is a set of arguments and $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ is a binary relation representing attacks between arguments. For two arguments $a, b \in \mathcal{A}$, the notation $a\mathcal{R}b$ or $(a, b) \in \mathcal{R}$ means that a attacks b.*

Thus, each argumentation framework can be represented as a directed graph whose nodes represent arguments of the framework and the arcs stand for attacks between them.

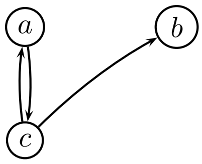
Example 2.2.1. *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework with $\mathcal{A} = \{a, b, c, d\}$ and $\mathcal{R} = \{(a, a), (a, b), (b, c), (c, d), (d, b)\}$. The graphical representation of the framework is shown below.*



Example 2.2.2. *Consider again the dialogue between the two violinists David and Jascha from the beginning of the chapter. That dialogue can be formalized by a simple framework with two arguments:*
a : *"This violin is expensive since it is a Stradivarius."*
b : *"I do not think that the violin is a Stradivarius."*
Since Jascha challenged David's argument, b attacks a.

The same dialogue can be formalized in a different manner.

Example 2.2.3. *The dialogue between Jascha and David from the beginning of the chapter can be represented by the following argumentation framework:*
a: "This violin is a Stradivarius."
b: "Since the violin is a Stradivarius, it is expensive."
c: "The violin is not a Stradivarius."
The graph associated with this framework is depicted in the figure below.



Until now, we did not make any assumptions on the cardinality of the set of arguments. We define a finite argumentation framework as follows.

Definition 2.2.2. *Argumentation framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ is* finite *iff $\mathcal{A}$ is finite.*

2.2.1 Acceptability semantics

One of the key steps in an argumentation process is the one in which arguments are evaluated using an acceptability semantics. A semantics is a set of criteria that should be satisfied by a set of arguments in order to be acceptable.

In argumentation literature, two main families of approaches for defining a semantics exist: *declarative* approaches and *labeling*-based ones. A declarative approach specifies which sets of arguments are acceptable. Examples of such semantics are those proposed by Dung (1995) (i.e. admissible, complete, preferred, stable, grounded) as well as their refinements: semi-stable (Caminada, 2006b), ideal (Dung, Mancarella, and Toni, 2007), recursive (Baroni, Giacomin, and Guida, 2005) and prudent semantics (Coste-Marquis, Devred, and Marquis, 2005). A labeling-based approach follows two steps: i) to assign a label to each argument using a particular labeling function, and ii) to compute the extensions. Generally three labels are assumed: `In`, stating that the argument is acceptable; `Out`, meaning that the argument is rejected; `Und`, describing the case where the status of the argument is floating (i.e. unknown). Examples of labeling-based semantics include robust semantics (Jakobovits and Vermeir, 1999) and stage semantics (Verheij, 1996). It was also shown by Caminada (2006a) that Dung's semantics can be redefined using labeling functions. Whatever the approach is, a semantics defines extensions which are acceptable sets of arguments. The idea behind an extension is that it represents a coherent point of view or a coherent position. Thus, each extension should be conflict-free, that is it must not contain arguments which attack each other.

Definition 2.2.3 (Conflict-freeness). *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and $\mathcal{S} \subseteq \mathcal{A}$. $\mathcal{S}$ is* conflict-free *iff $\nexists a, b \in \mathcal{S}$ s.t. $a\mathcal{R}b$.*

For the purpose of this thesis, we only need to recall Dung's semantics. They are based on a notion of defence which is defined as follows.

Definition 2.2.4 (Defence). *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework, $a \in \mathcal{A}$ and $\mathcal{S} \subseteq \mathcal{A}$. $\mathcal{S}$* defends *argument a iff $\forall b \in \mathcal{A}$ if $b\mathcal{R}a$ then $\exists c \in \mathcal{S}$ s.t. $c\mathcal{R}b$.*

Dung's semantics are based on a notion of *admissibility.* The intuition behind the notion of admissibility is that a set of arguments is acceptable if for any argument which is somehow challenged from outside, the counter-attack is possible from the arguments present in the set.

Definition 2.2.5 (Admissible semantics)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework, and $\mathcal{S} \subseteq \mathcal{A}$. $\mathcal{S}$ is an* admissible *set of $\mathcal{F}$ iff $\mathcal{S}$ is conflict-free and defends all its elements.*

Example 2.2.4. *In the framework of Example 2.2.3, there are exactly four admissible sets: $\emptyset, \{a\}, \{c\}, \{a, b\}$.*

Note that every argumentation framework has at least one admissible set; the empty set is admissible in every argumentation framework.

In the previous example, the set $\{a\}$ is admissible but it does not contain the argument b which is defended by a. However, if we accept a, why not accept b? In other words, if we accept that the violin is a Stradivari, then it should be natural to accept that it is expensive. This type of reasoning gives rise to *complete* semantics.

Definition 2.2.6 (Complete semantics)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and $\mathcal{S} \subseteq \mathcal{A}$. $\mathcal{S}$ is a* complete extension *of $\mathcal{F}$ iff $\mathcal{S}$ is conflict-free and $\mathcal{S} = \{a \in \mathcal{A} \mid \mathcal{S} \text{ defends } a\}$.*

Example 2.2.5. *In the framework of Example 2.2.3, there are exactly three complete extensions: $\emptyset, \{c\}, \{a, b\}$.*

It is easy to see that any complete extension is an admissible set. The converse is not true. For example, the set $\{a\}$ in Example 2.2.3 is an admissible set, but is not a complete extension.

It may seem surprising that the empty set is a complete extension in Example 2.2.3, since both a and c are omitted in that case. In the particular meaning we gave to those arguments, one could say that the violin is either a Stradivarius or not. This leads to the definition of *preferred* semantics, which includes a notion of maximality.

Definition 2.2.7 (Preferred semantics)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework, and $\mathcal{S} \subseteq \mathcal{A}$. $\mathcal{S}$ is a* preferred extension *iff $\mathcal{S}$ is a maximal (for set inclusion) admissible set.*

In other words, a set is a preferred extension of $\mathcal{F}$ iff it is admissible and no strict superset of that set is an admissible set.

Example 2.2.6. *In the framework of Example 2.2.3, there are exactly two preferred extensions: $\{c\}$ and $\{a, b\}$.*

Theorem 2.2.1 (Dung, 1995)**.** Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework.

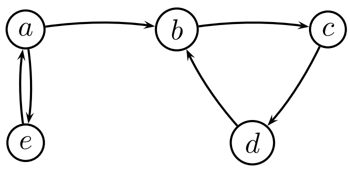
- $\mathcal{F}$ has at least one preferred extension.
- Any preferred extension of $\mathcal{F}$ is a complete extension of $\mathcal{F}$.

The contrary of the previous result is not true; for example the empty set is a complete extension but not a preferred one in the framework of Example 2.2.3.

Another semantics widely used in argumentation is *stable* semantics. According to this semantics, a set of arguments is acceptable if it is conflict-free and attacks any argument outside that set.

Definition 2.2.8 (Stable semantics)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and $\mathcal{S} \subseteq \mathcal{A}$. $\mathcal{S}$ is a* stable extension *of $\mathcal{F}$ iff $\mathcal{S}$ is conflict-free and $\forall a \in \mathcal{A} \setminus \mathcal{S}$, $\exists b \in \mathcal{S}$ s.t. $b\mathcal{R}a$.*

Example 2.2.7. *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be as depicted below. This framework has two preferred extensions: $\{a, c\}$ and $\{e\}$. However, only $\{a, c\}$ is a stable extension.*



Theorem 2.2.2 (Dung, 1995)**.** Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework. Any stable extension of $\mathcal{F}$ is also a preferred extension of $\mathcal{F}$.

A serious drawback of stable semantics is that the existence of stable extensions is not guaranteed. For instance, the framework of Example 2.2.1 has no stable extensions.

All the semantics presented so far may return more than one extension. This means that arguments may have multiple statuses: they may be accepted in some extensions and rejected in others. A semantics which assigns only one status to each argument was also proposed by Dung (1995). It is the well-known *grounded* semantics which is the minimal (for set inclusion) complete extension.

Definition 2.2.9 (Grounded semantics)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and $\mathcal{S} \subseteq \mathcal{A}$. $\mathcal{S}$ is a* grounded extension *of $\mathcal{F}$ iff $\mathcal{S}$ is a minimal for set inclusion complete extension of $\mathcal{F}$.*

Theorem 2.2.3 (Dung, 1995)**.** Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework.

- $\mathcal{F}$ has exactly one grounded extension (which may be empty).
- The grounded extension of $\mathcal{F}$ is exactly the set-theoretic intersection of all complete extensions of $\mathcal{F}$.
- The grounded extension of $\mathcal{F}$ is a subset of any preferred extension of $\mathcal{F}$.

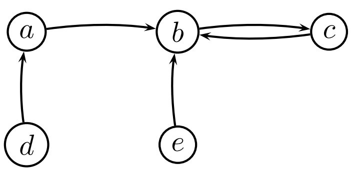
When the argumentation framework is finite, its grounded extension can be computed by iterative application of characteristic function $\mathfrak{F}$ on the empty set. For a given set $\mathcal{S}$ of arguments, the function $\mathfrak{F}$ returns the set of arguments defended by $\mathcal{S}$. In other words, $\mathfrak{F}(\mathcal{S})$ is the set containing all (and only) arguments that $\mathcal{S}$ defends.

Definition 2.2.10 (Characteristic function)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework. The* characteristic function *of $\mathcal{F}$ is defined as follows:*

- $\mathfrak{F} : 2^{\mathcal{A}} \rightarrow 2^{\mathcal{A}}$
- $\mathfrak{F}(\mathcal{S}) = \{a \in \mathcal{A} \mid \mathcal{S} \text{ defends } a\}$*, for all* $\mathcal{S} \subseteq \mathcal{A}$*.*

If $\mathcal{A}$ is finite, the grounded extension can be calculated by iterative applications of function $\mathfrak{F}$ to the empty set, i.e. it is equal to $\bigcup_{i=0}^{\infty} \mathfrak{F}^i(\emptyset)$, where $\mathfrak{F}^i(\mathcal{S}) = \underbrace{\mathfrak{F}(\mathfrak{F}(\ldots\mathfrak{F}}_{i \text{ times}}(\mathcal{S}))\ldots)$.

Example 2.2.8. *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be the argumentation framework depicted below. The grounded extension is calculated as follows: $\mathfrak{F}(\emptyset) = \{d, e\}$, $\mathfrak{F}(\{d, e\}) = \{c, d, e\}$, $\mathfrak{F}(\{c, d, e\}) = \{c, d, e\}$. Thus, the grounded extension is the set $\{c, d, e\}$.*



2.2.2 Status of arguments

Abstract argumentation frameworks return two outputs: a set of extensions of arguments under a given semantics, and a status for each argument. This latter is computed on the basis of argument's membership to extensions.

Definition 2.2.11 (Status of arguments)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework,* $\texttt{Ext}(\mathcal{F})$ *be the set of its extensions under a given semantics and $a \in \mathcal{A}$.*

- *a is sceptically accepted (or sceptical) iff $a \in \bigcap \mathcal{E}_i$, where $\mathcal{E}_i \in \texttt{Ext}(\mathcal{F})$*
- *a is credulously accepted (or credulous) iff $a \in \bigcup \mathcal{E}_i$, where $\mathcal{E}_i \in \texttt{Ext}(\mathcal{F})$*
- *a is rejected iff $a \notin \bigcup \mathcal{E}_i$, where $\mathcal{E}_i \in \texttt{Ext}(\mathcal{F})$.*

Example 2.2.9. *The framework of Example 2.2.3 has two stable extensions: $\{a, b\}$ and $\{c\}$. Thus, all the arguments are credulously accepted under stable semantics. The grounded extension of this framework is the empty set, thus, all the arguments are rejected under this semantics.*

Note that any sceptical argument is also credulous. However, there are exactly three disjunct cases, since an argument can be: i) sceptical (and credulous), ii) credulous and not sceptical iii) rejected. Let $\texttt{Status}(a, \mathcal{F})$ be a function which returns the status of an argument a in an argumentation framework $\mathcal{F}$. This function simply returns three different values in those three disjunct cases.

2.2.3 Complexity

It is well-known that argumentation reasoning is computationally costly. The concepts of credulous and sceptical acceptance motivate a number of decision problems, summarised below, that have been considered by Dimopoulos and Torres (1996) and by Dunne and Bench-Capon (2002).

Theorem 2.2.4. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework, $\mathcal{S} \subseteq \mathcal{A}$ and $a \in \mathcal{A}$.

Question	Complexity
Is $\mathcal{S} \subseteq \mathcal{A}$ a stable extension of $\mathcal{F}$?	polynomial
Does $\mathcal{F}$ have any stable extension?	NP-complete
Does $\mathcal{F}$ have a non-empty preferred extension?	NP-complete
Is a credulous under stable semantics?	NP-complete
Is a credulous under preferred semantics?	NP-complete
Is a sceptical under stable semantics?	CO-NP-complete
Is a sceptical under preferred semantics?	Π_2^p-complete
Is every preferred extension of $\mathcal{F}$ a stable one?	Π_2^p-complete

In addition to those results, Dunne (2007) has studied computational properties of argumentation frameworks which satisfy graph-theoretic constraints, e.g. when the numbers of attacks originating from and made upon any argument are bounded.

2.3 Logic-based argumentation

Until now, we have studied abstract argumentation, which means that we supposed that the origin and the structure of arguments are not known. We will now consider building arguments from a knowledge base using a given logic.

2.3.1 Logical language

A logic has two main components: a logical language $\mathcal{L}$ which is a set of well-formed formulae, and a consequence operator CN which is used to draw conclusions. Given a set of formulae $X \subseteq \mathcal{L}$, the set $\mathsf{CN}(X)$ denotes the set of conclusions that are drawn from the set X. Note that CN is a function, $\mathsf{CN} : 2^{\mathcal{L}} \to 2^{\mathcal{L}}$.

Example 2.3.1. *Let* $(\mathcal{L}, \mathsf{CN})$ *be propositional logic and* $X = \{x, y\}$. *Then,* $\mathsf{CN}(X) = \{x, y, x \wedge y, x \wedge x, x \to (x \wedge y), \ldots\}$.

Example 2.3.2. *Let us consider a simple logic for representing the colour and the size of objects. Let* $\mathcal{L} = \mathcal{L}_{col} \cup \mathcal{L}_{size} \cup \mathcal{L}_{err}$ *with* $\mathcal{L}_{col} = \{white,\ yellow,\ red,\ orange,\ blue,\ black\}$, $\mathcal{L}_{size} = \{tiny, small, big, huge\}$, $\mathcal{L}_{err} = \{\bot\}$. *In this simple example, the consequence operator captures the fact that if two different colours or two different sizes are present in the description of an object, then information concerning that object is inconsistent. We define*

CN *as follows: for all* $X \subseteq \mathcal{L}$,

$$\mathsf{CN}(X) = \begin{cases} \mathcal{L}, & \text{if } (\exists x, y \in X \text{ s.t. } x \neq y \\ & \quad \text{and } (\{x,y\} \subseteq \mathcal{L}_{col} \text{ or } \{x,y\} \subseteq \mathcal{L}_{size})) \\ & \text{or if } (\perp \in X) \\ X, & \text{else} \end{cases}$$

For example, $\mathsf{CN}(\emptyset) = \emptyset$, $\mathsf{CN}(\{red, big\}) = \{red, big\}$, $\mathsf{CN}(\{red, blue, big\}) = \mathsf{CN}\{\perp\} = \mathcal{L}$.

Two main families of logics are used in argumentation literature. The first family contains approaches where arguments are built from a Tarskian logic, while the second group of works uses rule-based systems for constructing arguments.

Tarski (1956) defined a notion of an abstract logic as follows.

Definition 2.3.1 (Tarski, 1956)**.** *A Tarskian logic is a pair* $(\mathcal{L}, \mathsf{CN})$, *where* $\mathcal{L}$ *is a set of formulae and* $\mathsf{CN} : 2^{\mathcal{L}} \to 2^{\mathcal{L}}$ *its consequence operator which verifies the following axioms:*

1. $X \subseteq \mathsf{CN}(X)$ ***(Expansion)***
2. $\mathsf{CN}(\mathsf{CN}(X)) = \mathsf{CN}(X)$ ***(Idempotence)***
3. $\mathsf{CN}(X) = \bigcup_{Y \subseteq_f X} \mathsf{CN}(Y)$ ***(Finitude)***
4. $\mathsf{CN}(\{x\}) = \mathcal{L}$ *for some* $x \in \mathcal{L}$ ***(Absurdity)***
5. $\mathsf{CN}(\emptyset) \neq \mathcal{L}$ ***(Coherence)***

Notation $Y \subseteq_f X$ *means that* Y *is a finite subset of* X.

The coherence requirement is absent from Tarski's original axioms, but added here to rule out trivial systems. Many well-known logics (e.g. propositional logic, first-order logic, modal logics, intuitionistic logic...) verify those axioms and are thus Tarskian logics.

Example 2.3.3. *Let* $(\mathcal{L}, \mathsf{CN})$ *be the logic from Example 2.3.2. It is easy to see that this simple logic verifies all the five axioms of the previous definition. Expansion and idempotence are verified directly from the definition of* CN. *Finiteness is satisfied since* $\mathcal{L}$ *is finite. Absurdity and coherence are verified since* $\mathsf{CN}(\{\perp\}) = \mathcal{L}$ *and* $\mathsf{CN}(\{\emptyset\}) = \emptyset$.

In a Tarskian logic, consistency is defined as follows.

Definition 2.3.2. *Let* $(\mathcal{L}, \mathsf{CN})$ *be a Tarskian logic and* $X \subseteq \mathcal{L}$. X *is* consistent *in* $(\mathcal{L}, \mathsf{CN})$ *iff* $\mathsf{CN}(X) \neq \mathcal{L}$. *It is* inconsistent *otherwise.*

Example 2.3.4. *Let* $(\mathcal{L}, \mathsf{CN})$ *be the logic from Example 2.3.2. The set* $\{red,$ $big\}$ *is consistent, while* $\{red, blue, big\}$ *is inconsistent.*

We can distinguish two classes of works that use Tarskian logics. The first class contains works which study argumentation frameworks built under *any* Tarskian logic, like done by Amgoud and Besnard (2009, 2010). Works in the second class concern argumentation frameworks built from a particular Tarskian logic: propositional logic (Amgoud and Cayrol, 1998; Besnard and Hunter, 2001; Simari and Loui, 1992), first-order logic (e.g. Besnard and Hunter, 2001), etc.

The second family of works relies on rule-based systems. Examples of such works include the work of Prakken and Sartor (1997) or the system of ASPIC project (Amgoud, Caminada, Cayrol, Lagasquie, and Prakken, 2004). The underlying logic language usually consists of a set $\mathcal{P}$ of literals (i.e. atomics formulae and their negations), a set $\mathcal{S}$ of strict rules and a set $\mathcal{D}$ of defeasible rules. A strict rule has the form $l_1, \ldots, l_{n-1} \rightarrow l_n$ where every l_i is a literal of $\mathcal{P}$. The meaning of this rule is that if $l_1, \ldots, l_{n-1}$ are true, then l_n is true. A defeasible rule has the form $l_1, \ldots, l_{n-1} \Rightarrow l_n$, where every l_i is a literal of $\mathcal{P}$, and expresses the fact that if $l_1, \ldots, l_{n-1}$ are true, then generally l_n is also true.

In this thesis, we are interested in instantiating Dung's argumentation framework by a Tarskian logic.

2.3.2 Arguments

An argument consists of two parts: a support and a conclusion. It is defined from formulae of a knowledge base $\Sigma \subseteq \mathcal{L}$ using a consequence operator CN.

Definition 2.3.3 (Argument)**.** *Let* $(\mathcal{L}, \mathsf{CN})$ *be a Tarskian logic and* $\Sigma \subseteq \mathcal{L}$. (H, h) *is an* argument *built from* Σ *iff:*

1. $H \subseteq \Sigma$
2. H *is consistent*
3. $h \in \mathsf{CN}(H)$

4. $\nexists H' \subset H$ *s.t.* $h \in \mathsf{CN}(H')$.

H is called support *and h* conclusion *of the argument.*

The consistency condition forbids using inconsistent sets as supports, since an argument should be based on coherent hypotheses. The third condition specifies that the conclusion can be deduced from a support, while the last condition guarantees that only relevant information is included among the hypotheses of an argument.

Notations: For an argument $a = (H, h)$, $\mathtt{Conc}(a) = h$ and $\mathtt{Supp}(a) = H$. For a set $\mathcal{S} \subseteq \mathcal{L}$, $\mathtt{Arg}(\mathcal{S}) = \{a \mid a$ is an argument (in the sense of Definition 2.3.3) and $\mathtt{Supp}(a) \subseteq \mathcal{S}\}$. For any $\mathcal{E} \subseteq \mathtt{Arg}(\mathcal{L})$, $\mathtt{Base}(\mathcal{E}) = \bigcup_{a \in \mathcal{E}} \mathtt{Supp}(a)$.

Example 2.3.5. *Let* $(\mathcal{L}, \mathsf{CN})$ *be propositional logic and* $\Sigma = \{strad, strad \rightarrow exp, \neg strad\}$. *Arguments constructed from* Σ *include the following ones:* $(\{strad\}, strad)$, $(\{strad, strad \rightarrow exp\}, exp)$, $(\{\neg strad, strad \rightarrow exp\}, \neg strad \wedge (\neg exp \rightarrow \neg strad))$, … *Note that the set of all arguments that can be built from* Σ *is infinite.*

Example 2.3.6. *Let* $(\mathcal{L}, \mathsf{CN})$ *be the logic defined in Example 2.3.2, and let* $\Sigma = \{red, blue, big\}$. $\mathtt{Arg}(\Sigma) = \{(\{red\}, red), (\{blue\}, blue), (\{big\}, big)\}$.

Example 2.3.7. *Let* $(\mathcal{L}, \mathsf{CN})$ *be the* $S5$ *logic, and* $\Sigma = \{\neg strad, \Diamond strad\}$. *Arguments constructed from* Σ *include the following ones:* $(\{\Diamond strad\}, \Box\Diamond strad)$, $(\{\neg strad\}, \Box\Diamond\neg strad)$, …

2.3.3 Interactions between arguments

Arguments can interact in different manners: they can *attack* or *support* other arguments. An attack expresses a conflict between two arguments. It is almost always represented by a binary relation on the set of arguments. A common practice in logic-based argumentation is to define an attack relation by specifying the type of logical inconsistency between two arguments whose presence implies existence of an attack. For example, if the conclusion of an argument somehow contradicts one of the formulae in the support of another argument, then the former attacks the latter. We recall below the most commonly used attack relations in the literature. For illustration purposes, we use propositional logic, but note that similar ideas may be used in definitions of attack relations for many different logics.

Definition 2.3.4 (Attack relations in propositional logic)**.** *Let us suppose two arguments* $a = (\{h_1, \ldots, h_n\}, h)$ *and* $a' = (\{h'_1, \ldots, h'_m\}, h')$ *built from the formulae of propositional logic. with* $a = (\{h_1, \ldots, h_n\}, h)$ *and* $a' = (\{h'_1, \ldots, h'_m\}, h')$*. Below are several criteria that can be used for defining an attack from* a *to* a'*:*

(1) $h \vdash \neg(h'_1 \wedge \ldots \wedge h'_n)$ *(called* defeat*)*

(2) $\exists h'_i \in \texttt{Supp}(a')$ *s.t.* $h \vdash \neg h'_i$ *(called* direct defeat*)*

(3) $\exists H'' = \{h''_1, \ldots, h''_p\} \subseteq \texttt{Supp}(a')$ *s.t.* $h \equiv \neg(h''_1 \wedge \ldots \wedge h''_p)$ *(called* undercut*)*

(4) $h \equiv \neg(h'_1 \wedge \ldots \wedge h'_n)$ *(called* canonical undercut*)*

(5) $\exists h'_i \in \texttt{Supp}(a')$ *s.t.* $h \equiv \neg h'_i$ *(called* undercut *or* direct undercut*)*

(6) $h \equiv \neg h'$ *(called* rebut*)*

(7) $h \vdash \neg h'$ *(called* defeating rebut*)*

Example 2.3.8. *Let* $(\mathcal{L}, \mathsf{CN})$ *be propositional logic,* $a = (\{strad, strad \rightarrow exp\}, exp)$ *and* $a' = (\{\neg strad\}, \neg strad)$*.* a' *attacks* a *w.r.t. (1), (2), (3), and (5). Argument* $a'' = (\{\neg strad\}, \neg(strad \wedge strad \rightarrow exp))$ *attacks* a *w.r.t. (1), (3) and (4). Argument* a *attacks* $a''' = (\{costs1000, costs1000 \rightarrow \neg exp\}, \neg exp)$ *w.r.t. (1), (6) and (7).*

Those definitions of attack relation may be adapted to other logics having negation and conjunction. If we want to provide a general definition of an attack relation for any Tarskian logic, things are more complicated since Tarski's definition is very abstract and there is no guarantee that the logic in question has any negation. The simplest solution is to define attack relation w.r.t. inconsistency.

Definition 2.3.5 (Attack relations in a Tarskian logic)**.** *Let us suppose that* $a = (H, h)$ *and* $a' = (H', h')$ *are two arguments built from a Tarskian logic* $(\mathcal{L}, \mathsf{CN})$*. Below are several criteria that can be used for defining an attack from* a *to* a'*:*

(1) $\{h\} \cup H'$ *is inconsistent*

(2) $\exists h'_i \in H'$ *s.t.* $\{h\} \cup \{h'_i\}$ *is inconsistent*

(3) $\{h\} \cup \{h'\}$ *is inconsistent.*

An argument can also support another one. This is also captured by a binary relation on the set of arguments. An argument supported by several arguments is supposed to be stronger than a non-supported argument. However, a supported argument is not necessarily accepted in an argumentation framework.

The following are several possibilities for defining such a relation.

Definition 2.3.6 (Support relations in a Tarskian logic)**.** *Let $a = (H, h)$ and $a' = (H', h')$ be two arguments built from a Tarskian logic $(\mathcal{L}, \mathsf{CN})$. Below are several criteria that can be used for defining a support from a to a':*

(1) $h = h'$

(2) $\exists h'_i \in H'$ *s.t.* $h = h'_i$

(3) the set $H \cup H'$ *is consistent and* $\exists h'_i \in H'$ *s.t.* $h = h'_i$.

Amgoud and Besnard (2009) have conducted a general study on how to choose an appropriate attack relation. They studied the link between the inconsistency in arguments' supports and conclusions and attacks between them. In order to do so, they used the notion of a minimal conflict.

Definition 2.3.7 (Minimal conflict)**.** *Let $(\mathcal{L}, \mathsf{CN})$ be a Tarskian logic and $C \subseteq \mathcal{L}$. C is a* minimal conflict *iff:*

- *C is inconsistent, and*
- *$\forall x \in C$, $C \setminus \{x\}$ is consistent.*

Let $\mathscr{C}_{\mathcal{L}}$ denote the set of all minimal conflicts of $\mathcal{L}$.

An example of property of an attack relation (Amgoud and Besnard, 2009) is conflict-dependency.

Definition 2.3.8 (Conflict-dependent)**.** *Let $(\mathcal{L}, \mathsf{CN})$ be a Tarskian logic. An attack relation $\mathcal{R} \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ is* conflict-dependent *iff for $a, b \in \mathtt{Arg}(\mathcal{L})$, $a\mathcal{R}b$ implies that there exists a minimal conflict $C \in \mathscr{C}_{\mathcal{L}}$ such that $C \subseteq \mathtt{Supp}(a) \cup \mathtt{Supp}(b)$.*

Being conflict-dependent means that $\mathcal{R}$ shows no attack from a to b unless there is a conflict between the supports of a and b.

Let us introduce several other properties an attack relation can verify.

C1 $\forall a, b, c \in \mathcal{A}$, if $\mathtt{Conc}(a) = \mathtt{Conc}(b)$ then $(a\mathcal{R}c \Leftrightarrow b\mathcal{R}c)$

C1' $\forall a, b, c \in \mathcal{A}$, if $\texttt{Conc}(a) \equiv \texttt{Conc}(b)$ then $(a\mathcal{R}c \Leftrightarrow b\mathcal{R}c)$

C2 $\forall a, b, c \in \mathcal{A}$, if $\texttt{Supp}(a) = \texttt{Supp}(b)$ then $(c\mathcal{R}a \Leftrightarrow c\mathcal{R}b)$

C2' $\forall a, b, c \in \mathcal{A}$, if $\texttt{Supp}(a) \equiv \texttt{Supp}(b)$ then $(c\mathcal{R}a \Leftrightarrow c\mathcal{R}b)$

The two first properties say that two arguments having the same (resp. equivalent) conclusions attack exactly the same set of arguments. The two remaining properties say that arguments having the same (resp. equivalent) supports are attacked by the same set of arguments.

Proposition 2.3.1. Let $\mathcal{R}$ be an attack relation.

- If $\mathcal{R}$ satisfies $C1'$ then it satisfies $C1$.
- If $\mathcal{R}$ satisfies $C2'$ then it satisfies $C2$.

Amgoud and Besnard (2009) have defined rationality postulates that any logic-based argumentation framework should satisfy. One of them concerns the consistency of the results that are returned by its extensions. Indeed, an argumentation framework satisfies extension consistency iff for every extension, the set of formulae used in its arguments is consistent.

Definition 2.3.9. *Let $(\mathcal{L}, \mathsf{CN})$ be a Tarskian logic and $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be argumentation framework whose arguments are built using that logic. $\mathcal{F}$ satisfies extension consistency iff for every $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $\texttt{Base}(\mathcal{E})$ is consistent.*

Amgoud and Besnard (2009) have shown that if a Tarskian logic is used for constructing arguments then a conflict-dependent and symmetric attack relation may violate extension consistency. Indeed, if a knowledge base Σ contains at least one minimal conflict of cardinality three or more, then $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$ violates extension consistency if $\mathcal{R}$ is symmetric and conflict-dependent.

2.3.4 Outputs of an argumentation framework

In addition to extensions and statuses of arguments, we now define other outputs of an argumentation framework.

Definition 2.3.10 (Outputs of an argumentation framework)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework.*

- $\texttt{Sc}(\mathcal{F}) = \{a \in \mathcal{A} \mid a \textit{ is sceptically accepted }\}$
- $\texttt{Cr}(\mathcal{F}) = \{a \in \mathcal{A} \mid a \textit{ is credulously accepted }\}$

- $\texttt{Output}_{sc}(\mathcal{F}) = \{\texttt{Conc}(a) \mid a \textit{ is sceptically accepted } \}$
- $\texttt{Output}_{cr}(\mathcal{F}) = \{\texttt{Conc}(a) \mid a \textit{ is credulously accepted } \}$
- $\texttt{Bases}(\mathcal{F}) = \{\texttt{Base}(\mathcal{E}) \mid \mathcal{E} \in \texttt{Ext}(\mathcal{F})\}$

The first four sets contain the sceptically and credulously accepted arguments (resp. conclusions). $\texttt{Bases}(\mathcal{F})$ contains the subsets of Σ which are returned by the extensions of $\mathcal{F}$. It is worth noticing that $\texttt{Sc}(\mathcal{F}) \subseteq \texttt{Arg}(\mathcal{L})$, $\texttt{Cr}(\mathcal{F}) \subseteq \texttt{Arg}(\mathcal{L})$, $\texttt{Output}_{sc}(\mathcal{F}) \subseteq \mathcal{L}$, $\texttt{Output}_{cr}(\mathcal{F}) \subseteq \mathcal{L}$ and $\texttt{Base}(\mathcal{E}) \subseteq \mathcal{L}$ for $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$.

2.4 Conclusion

In this chapter we have introduced abstract argumentation and logic-based argumentation. In the first part of the chapter, we studied the most abstract argumentation framework in the literature, which was proposed by Dung (1995). We showed how to define an abstract argumentation framework, how to use a semantics to calculate extensions and assign a status to each argument.

In the second part of the chapter, we introduced the basics of logic-based argumentation, where arguments are built under a given monotonic logic. We showed how to define a logic-based argument and discussed several attack relations and support relations. We also defined outputs of a logic-based argumentation framework.

Everything should be as simple as it is,
but not simpler.
Albert Einstein

3 Equivalence in argumentation

3.1 Introduction

This chapter tackles the question of equivalence between logic-based argumentation frameworks (Amgoud and Vesic, 2011c). Our study is motivated by several reasons.

First, when building an argumentation framework from a given knowledge base, it is very common that several attack relations may be used. Thus, knowing under which conditions different attack relations induce same or similar results is very likely to be useful.

Second, under many logics (e.g. propositional logic) an infinite number of arguments is built from a finite knowledge base. It would be convenient to know whether such a framework can be exchanged with an equivalent finite framework.

Besides, even in a finite case, building logic-based arguments from a concrete knowledge base is computationally complex. There are at least two tests to be done: a *consistency* test for checking whether argument's support is consistent and an *inference* test to check whether the argument's conclusion is a logical consequence of the support. In the case of propositional logic, those two tests are NP-complete and co-NP-complete, respectively. Thus, any reduction in the number of arguments of an argumentation framework would be a step forward.

A study on when two Dung's abstract frameworks are *equivalent* has been carried out by Oikarinen and Woltran (2010). Authors defined three equivalence criteria: according to them, two argumentation frameworks are equivalent if they return i) the same extensions, ii) the same sets of sceptical arguments, or iii) the same sets of credulous arguments. The main focus of the article is not on equivalence, but rather on *strong equivalence*, which is defined as follows: two frameworks are strongly equivalent iff after an arbitrary set of arguments and attacks have been added to both of them, the two enriched frameworks still return the same set of extensions (respectively sceptical / credulous arguments).

While these criteria are meaningful, they are too rigid; it has been shown by Oikarinen and Woltran (2010) that if there are no self-attacking arguments (i.e. if $\mathcal{R}$ is anti-reflexive) then any two frameworks are strongly equivalent (w.r.t. any of the above criteria) if and only if they are equal. This makes strong equivalence a nice theoretical property, but without practical applications.

In this chapter, we argue that when the structure of arguments is taken into account, similarities arise which are undetectable on the abstract level. The following example serves to illustrate this issue. Consider two argumentation frameworks built under propositional logic: the first framework has the set $\{(\{x \to y\}, x \to y)\}$ as its unique extension while the only extension of the second one is $\{(\{x \to y\}, \neg x \vee y)\}$. These two frameworks are not equivalent with respect to the above criteria since the two arguments $\{(\{x \to y\}, x \to y)\}$ and $\{(\{x \to y\}, \neg x \vee y)\}$ are different. However, under some reasonable assumptions those two arguments should be considered equivalent or exchangeable.

Thus, in order to define more accurately the notion of equivalence between two frameworks, the structure of arguments should be taken into account. First, we exploit this fact to define equivalence criteria between argumentation frameworks (Subsection 3.2.1) and study their interdependencies (Subsection 3.2.2). We also provide conditions under which two frameworks are equivalent w.r.t. a given criterion (Subsection 3.2.3). We study strong equivalence in Section 3.3. The rest of the chapter presents diverse applications. In Section 3.4, we show how to identify a core of a given argumentation framework, a core being its sub-framework containing only the essential arguments of the original framework. We also provide a condition under which a framework has a finite core. Finally, in Section 3.5, we apply our results in the case when new arguments are added or removed from a framework and we identify the cases when such a change does not influence the status of existing arguments.

3.2 Basic equivalence

In the whole chapter, we suppose Dung's argumentation framework instantiated with a Tarskian logic. More precisely, let $(\mathcal{L}, \mathsf{CN})$ be a Tarskian logic such that $\mathcal{L}$ is a countable set, and let $\Sigma \subseteq \mathcal{L}$ be a given (finite or infinite) knowledge base. We suppose that arguments are constructed as in Definition 2.3.3 and that their status is determined as in Definition 2.2.11. We say that argumentation framework $(\mathcal{A}, \mathcal{R})$ is built *from a knowledge base* Σ iff $\mathcal{A} \subseteq \texttt{Arg}(\Sigma)$. We did not restrict our attention to the single case of the whole set $\texttt{Arg}(\Sigma)$ of arguments that may be built from Σ. The reason is that we want to be more general, i.e. our results will be valid both for the case $\mathcal{A} = \texttt{Arg}(\Sigma)$ and the case $\mathcal{A} \subset \texttt{Arg}(\Sigma)$. This also allows us to study a sub-framework which is equivalent to the framework having $\texttt{Arg}(\Sigma)$ as a set of arguments. We may also need to compare two of its sub-frameworks.

We assume that arguments are evaluated using stable semantics. Note that this is not a substantial limitation since the main purpose of this chapter is to explore general ways to define equivalence in logical argumentation and not to study the subtleties of different semantics. A similar study can be conducted for any other semantics. Recall that the set of all extensions of an argumentation framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ is denoted by $\texttt{Ext}(\mathcal{F})$.

In general, an argumentation framework may have an infinite number of extensions even if the knowledge base Σ is finite. Let us consider the following example.

Example 3.2.1. *Let* $(\mathcal{L}, \mathsf{CN})$ *be propositional logic,* $\Sigma = \{x\}$ *and* $\mathcal{A} = \texttt{Arg}(\Sigma)$. *Assume that* $a\mathcal{R}b$ *iff* $a \neq b$. *It is clear that this framework has infinitely many stable extensions. Some of them are:* $(\{x\}, x)$, $(\{x\}, x \wedge x)$, $(\{x\}, x \vee y)$, $(\{x\}, x \wedge (z \vee \neg z))$.

The following result shows that if the attack relation verifies $C2$ then the argumentation framework built over a finite knowledge base has a finite number of extensions.

Proposition 3.2.1. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework built over Σ. If Σ is finite and $\mathcal{R}$ satisfies $C2$, then $(\mathcal{A}, \mathcal{R})$ has a finite number of extensions.

3.2.1 Equivalence criteria

Throughout this section, we assume a fixed Tarskian logic $(\mathcal{L}, \mathsf{CN})$ and two arbitrary argumentation frameworks $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ that are defined using this logic. Note that the two frameworks may be built over different knowledge bases. The goal of this subsection is to define equivalence criteria, i.e. to give a formal answer to the question: *"When are two argumentation frameworks equivalent?"* We propose two families of *equivalence criteria*. The first family compares directly the outputs of the two frameworks while the second family takes advantage of similarities between arguments and logical equivalence between formulae. The following definition introduces the criteria of the first family.

Definition 3.2.1 (Equivalence criteria)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built using the same Tarskian logic $(\mathcal{L}, \mathsf{CN})$. The two frameworks $\mathcal{F}$ and $\mathcal{F}'$ are* EQi-equivalent *iff criterion EQi below holds:*

EQ1: $\mathtt{Ext}(\mathcal{F}) = \mathtt{Ext}(\mathcal{F}')$

EQ2: $\mathtt{Sc}(\mathcal{F}) = \mathtt{Sc}(\mathcal{F}')$

EQ3: $\mathtt{Cr}(\mathcal{F}) = \mathtt{Cr}(\mathcal{F}')$

EQ4: $\mathtt{Output}_{sc}(\mathcal{F}) = \mathtt{Output}_{sc}(\mathcal{F}')$

EQ5: $\mathtt{Output}_{cr}(\mathcal{F}) = \mathtt{Output}_{cr}(\mathcal{F}')$

EQ6: $\mathtt{Bases}(\mathcal{F}) = \mathtt{Bases}(\mathcal{F}')$.

Note that the first three criteria were mentioned but not studied by Oikarinen and Woltran (2010). Let us consider again the example from the introduction.

Example 3.2.2. *Assume propositional logic and two argumentation frameworks $\mathcal{F}$ and $\mathcal{F}'$ having respectively $\{(\{x \to y\}, x \to y)\}$ and $\{(\{x \to y\}, \neg x \vee y)\}$ as their extensions. These two frameworks are equivalent w.r.t. criterion EQ6 since $\mathtt{Bases}(\mathcal{F}) = \mathtt{Bases}(\mathcal{F}') = \{\{x \to y\}\}$. However, they are not equivalent w.r.t. the remaining criteria, since those two arguments are considered as different even if they have the same supports and logically equivalent conclusions.*

The following example shows two frameworks which return different but somehow equivalent sub-bases of Σ.

Example 3.2.3. *Assume propositional logic and two argumentation frameworks $\mathcal{F}$ and $\mathcal{F}'$ having respectively $\{(\{x, \neg\neg y\}, x \wedge y)\}$ and $\{(\{x, y\}, x \wedge y)\}$ as extensions. The two frameworks are equivalent w.r.t. EQ4 and EQ5 but are not equivalent w.r.t. the remaining criteria, including EQ6. However, for each formula in $\mathtt{Bases}(\mathcal{F}) = \{\{x, \neg\neg y\}\}$, there is an equivalent one in $\mathtt{Bases}(\mathcal{F}) = \{\{x, y\}\}$ and vice versa.*

In order to have more refined notions of equivalence between argumentation frameworks, we take into account the logical equivalence between formulae and between sets of formulae.

Definition 3.2.2 (Equivalence between sets and formulae)**.** *Let $x, y \in \mathcal{L}$ and $X, Y \subseteq \mathcal{L}$.*

- *x and y are equivalent, denoted by $x \equiv y$, iff $\mathsf{CN}(\{x\}) = \mathsf{CN}(\{y\})$. We write $x \not\equiv y$ iff x and y are not equivalent.*
- *X and Y are equivalent, denoted by $X \cong Y$, iff $\forall x \in X$, $\exists y \in Y$ s.t. $x \equiv y$ and $\forall y \in Y, \exists x \in X$ s.t. $x \equiv y$. We write $X \ncong Y$ iff X and Y are not equivalent.*

In case of propositional logic, this allows to say that the two sets $\{x, \neg\neg y\}$ and $\{x, y\}$ are equivalent. Note that if $X \cong Y$, then $\mathsf{CN}(X) = \mathsf{CN}(Y)$. However, the converse is not true. For instance, $\mathsf{CN}(\{x \wedge y\}) = \mathsf{CN}(\{x, y\})$ while $\{x \wedge y\} \ncong \{x, y\}$. One may ask why not to use the equality of $\mathsf{CN}(X)$ and $\mathsf{CN}(Y)$ in order to say that X and Y are equivalent? The previous example have already given some of our motivation for such a definition: wanting to make a distinction between $\{x, y\}$ and $\{x \wedge y\}$. The following counter-example of two argumentation frameworks whose credulous conclusions are respectively $\{x, \neg x\}$ and $\{y, \neg y\}$ is more drastic: it is clear that $\mathsf{CN}(\{x, \neg x\})$ $= \mathsf{CN}(\{y, \neg y\})$ while the two sets are in no way similar.

In order to define an accurate notion of equivalence between two argumentation frameworks, we also take advantage of equivalence of arguments. Two arguments are equivalent if they have same or equivalent supports and conclusions.

Definition 3.2.3 (Equivalence between arguments)**.** *For two arguments $a, a' \in \mathtt{Arg}(\mathcal{L})$.*

- *$a \approx_1 a'$ iff $\mathtt{Supp}(a) = \mathtt{Supp}(a')$ and $\mathtt{Conc}(a) \equiv \mathtt{Conc}(a')$*
- *$a \approx_2 a'$ iff $\mathtt{Supp}(a) \equiv \mathtt{Supp}(a')$ and $\mathtt{Conc}(a) = \mathtt{Conc}(a')$*
- *$a \approx_3 a'$ iff $\mathtt{Supp}(a) \equiv \mathtt{Supp}(a')$ and $\mathtt{Conc}(a) \equiv \mathtt{Conc}(a')$*

Note that each relation $\approx_i$ is an *equivalence relation* (i.e. reflexive, symmetric and transitive). The equivalence between two arguments is extended to equivalence between sets of arguments as follows.

Definition 3.2.4 (Equivalence between sets of arguments). *Let* $\mathcal{E}, \mathcal{E}' \subseteq \texttt{Arg}(\mathcal{L})$ *and* $\approx_i$ *be an equivalence relation between arguments with* $i \in \{1, 2, 3\}$. *Two sets* $\mathcal{E}$ *and* $\mathcal{E}'$ *are equivalent, denoted* $\mathcal{E} \sim_i \mathcal{E}'$ *iff* $\forall a \in \mathcal{E}, \exists a' \in \mathcal{E}'$ *s.t.* $a \approx_i a'$ *and* $\forall a' \in \mathcal{E}', \exists a \in \mathcal{E}$ *s.t.* $a \approx_i a'$.

We are now ready to introduce the second family of equivalence criteria.

Definition 3.2.5 (Equivalence criteria continued). *Let* $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ *and* $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ *be two argumentation frameworks built using the same Tarskian logic* $(\mathcal{L}, \mathsf{CN})$. *Let* $\sim_i$ *be an equivalence relation between sets of arguments, with* $i \in \{1, 2, 3\}$. *The two frameworks* $\mathcal{F}$ *and* $\mathcal{F}'$ *are* EQi-equivalent *iff criterion EQi below holds:*

EQ1i: *there exists a bijection* $f : \texttt{Ext}(\mathcal{F}) \to \texttt{Ext}(\mathcal{F}')$ *such that* $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $\mathcal{E} \sim_i f(\mathcal{E})$

EQ2i: $\texttt{Sc}(\mathcal{F}) \sim_i \texttt{Sc}(\mathcal{F}')$

EQ3i: $\texttt{Cr}(\mathcal{F}) \sim_i \texttt{Cr}(\mathcal{F}')$

EQ4b $\texttt{Output}_{sc}(\mathcal{F}) \cong \texttt{Output}_{sc}(\mathcal{F}')$

EQ5b $\texttt{Output}_{cr}(\mathcal{F}) \cong \texttt{Output}_{cr}(\mathcal{F}')$

EQ6b $\forall S \in \texttt{Bases}(\mathcal{F})$, $\exists S' \in \texttt{Bases}(\mathcal{F}')$ *s.t.* $S \cong S'$ *and* $\forall S' \in \texttt{Bases}(\mathcal{F}')$, $\exists S \in \texttt{Bases}(\mathcal{F})$ *s.t.* $S \cong S'$.

Each of the above criteria refines a criterion in Definition 3.2.1 by considering the equivalences either between sets of arguments or sets of formulae. The three first criteria use an index i since they are built upon an equivalence relation $\sim_i$ between sets of arguments (with $i \in \{1, 2, 3\}$). Thus, for instance, EQ11 stands for a criterion which use relation $\sim_1$.

Example 3.2.4. *The two argumentation frameworks* $\mathcal{F}$ *and* $\mathcal{F}'$ *of Example 3.2.2 are equivalent w.r.t. criteria EQ11, EQ13, EQ21, EQ23, EQ31, EQ33, EQ4b, EQ5b and EQ6b; they are not equivalent w.r.t. EQ12, EQ22 and EQ32.*

Example 3.2.5. *The two argumentation frameworks* $\mathcal{F}$ *and* $\mathcal{F}'$ *of Example 3.2.3 are equivalent w.r.t. criteria EQ12, EQ13, EQ22, EQ23, EQ32, EQ33, EQ4b, EQ5b and EQ6b; they are not equivalent w.r.t. EQ11, EQ21 and EQ31.*

Notation: If two argumentation frameworks $\mathcal{F}$ and $\mathcal{F}'$ are equivalent w.r.t. criterion x, then we write $\mathcal{F} \equiv_x \mathcal{F}'$.

It is easy to check that each criterion is an *equivalence relation*, i.e. reflexive, symmetric and transitive.

Proposition 3.2.2. Each criterion is an equivalence relation.

Note that rejected arguments are not considered when comparing two argumentation frameworks. The reason is that rejected arguments are not an important output of a framework compared to sceptical and credulous arguments. Indeed, the set of rejected arguments is exactly the complement of the set of credulous arguments (which are themselves considered useful). Let us consider the following example.

Example 3.2.6. *Let* $(\mathcal{L}, \mathsf{CN})$ *be propositional logic, let* $a_1 = (\{t \wedge \neg x\}, \neg x)$, $a_2 = (\{x, y\}, x \wedge y)$, $a_3 = (\{w \wedge \neg y\}, \neg y)$, $\mathcal{A} = \{a_1, a_2\}$, $\mathcal{A}' = \{a_2, a_3\}$, $\mathcal{R} = \{(a_1, a_2)\}$, $\mathcal{R}' = \{(a_3, a_2)\}$. *It is easy to see that* $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ *and* $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ *would be equivalent if we compare rejected arguments, since their sets of rejected arguments coincide, i.e. for both frameworks that is the set* $\{a_2\}$*. However, those two frameworks have almost nothing in common since neither their conclusions nor their arguments coincide. Note also that arguments of those frameworks are not equivalent w.r.t. any reasonable equivalence relation.*

3.2.2 Links between criteria

It is clear that not all criteria are equally demanding and that they are not completely independent. For example, it is easy to see that when two argumentation frameworks are equivalent w.r.t. EQ1, then they are also equivalent w.r.t. EQ11, EQ12 and EQ13. In this section, we investigate all dependencies between the criteria proposed so far.

Theorem 3.2.1. Let $\mathcal{F}$ and $\mathcal{F}'$ be two argumentation frameworks built on the same logic $(\mathcal{L}, \mathsf{CN})$. Table 3.1 summarises the dependencies in the following form: $(\mathcal{F} \equiv_x \mathcal{F}') \Rightarrow (\mathcal{F} \equiv_{x'} \mathcal{F}')$.

Note that if two argumentation frameworks are equivalent w.r.t. $EQ1$, then they are equivalent w.r.t. any of the other criteria. This is not the case for its refined versions, i.e. for $EQ11, EQ12$ and $EQ13$. For instance, if two arguments are equivalent w.r.t. to $EQ11$, they are not necessarily equivalent w.r.t. $EQ21$, $EQ23$ and $EQ4b$. Later in this subsection, we will show that

EQi/EQj	1	11	12	13	2	21	22	23	3	31	32	33	4	4b	5	5b	6	6b
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
11		+		+						+		+				+	+	+
12			+	+							+	+			+	+		+
13				+								+				+		+
2					+	+	+	+					+	+				
21						+		+						+				
22							+	+					+	+				
23								+						+				
3									+	+	+	+			+	+		
31										+		+				+		
32											+	+			+	+		
33												+				+		
4													+	+				
4b														+				
5															+	+		
5b																+		
6																	+	+
6b																		+

Table 3.1: *Links between criteria. For two criteria, c in row i, and c' in column j, sign $+$ means that c implies c', more precisely, if two argumentation frameworks are equivalent w.r.t. c then they are equivalent w.r.t. c'.*

under some reasonable constraints, these implications exist. Indeed, if two argumentation frameworks are equivalent w.r.t. to $EQ11$, then they are also equivalent w.r.t. the three criteria $EQ21$, $EQ23$ and $EQ4b$ provided that the two frameworks use attack relations which verify properties $C1'$ and $C2$. Before presenting formally this result, let us study how the two properties $C1'$ and $C2$ of an attack relation are related to the equivalence relation $\approx_1$ between arguments which is used in criterion $EQ11$.

The following proposition shows that equivalent arguments w.r.t. relation $\approx_1$ behave in the same way w.r.t. attacks in case the attack relation enjoys the two properties $C1'$ and $C2$.

Proposition 3.2.3. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ verifies $C1'$ and $C2$. For all $a, a', b, b' \in \mathcal{A}$, ($a \approx_1 a'$ and $b \approx_1 b'$) $\Rightarrow$ ($a\mathcal{R}b$ iff $a'\mathcal{R}b'$).

The next result shows that equivalent arguments w.r.t. relation $\approx_1$ belong to the same extensions.

Proposition 3.2.4. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1'$ and $C2$. For all $a, a' \in \mathcal{A}$, if $a \approx_1 a'$, then $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $a \in \mathcal{E}$ iff $a' \in \mathcal{E}$.

It can also be checked that when two argumentation frameworks are equivalent w.r.t. $EQ11$, then if we consider two equivalent arguments (one from each framework), then the two arguments have the same status.

Proposition 3.2.5. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, and let $\mathcal{R}$ and $\mathcal{R}'$ verify

$C1'$ and $C2$, and $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$. For all $a \in \mathcal{A}$ and for all $a' \in \mathcal{A}'$, if $a \approx_1 a'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a', \mathcal{F}')$.

In general, when two argumentation frameworks are equivalent w.r.t. $EQ11$, they are not necessarily equivalent w.r.t. $EQ21, EQ23$ and $EQ4b$. The following result shows that when the attack relations of both frameworks verify $C1'$ and $C2$, the previous implications hold.

Theorem 3.2.2. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1'$ and $C2$. If $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ21, EQ23, EQ4b\}$.

When two argumentation frameworks are equivalent w.r.t. $EQ12$, they are also equivalent w.r.t. $EQ22, EQ23, EQ4$ and $EQ4b$ in case the attack relations of the two frameworks enjoy properties $C1$ and $C2'$. The reason is that there is a correlation between an attack relation which satisfies these two properties and the equivalence relation $\approx_2$ between arguments. Indeed, equivalent arguments w.r.t. $\approx_2$ behave in the same way w.r.t. an attack relation satisfying $C1$ and $C2'$.

Proposition 3.2.6. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1$ and $C2'$. For all $a, a', b, b' \in \mathcal{A}$, $(a \approx_2 a'$ and $b \approx_2 b') \Rightarrow (a\mathcal{R}b$ iff $a'\mathcal{R}b')$.

Equivalent arguments w.r.t. $\approx_2$ belong to the same extensions of an argumentation framework.

Proposition 3.2.7. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1$ and $C2'$. For all $a, a' \in \mathcal{A}$, if $a \approx_2 a'$ then $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $a \in \mathcal{E}$ iff $a' \in \mathcal{E}$.

Finally, two equivalent arguments pertaining to two frameworks whose attack relations satisfy $C1$ and $C2'$ have the same status.

Proposition 3.2.8. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1$ and $C2'$, and $\mathcal{F} \equiv_{EQ12} \mathcal{F}'$. For all $a \in \mathcal{A}$ and for all $a' \in \mathcal{A}'$, if $a \approx_2 a'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a', \mathcal{F}')$.

From the above properties, it follows that two argumentation frameworks which are equivalent w.r.t. $EQ12$ are also equivalent w.r.t. $EQ22, EQ23, EQ4$ and $EQ4b$.

Theorem 3.2.3. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1$ and $C2'$. If $\mathcal{F} \equiv_{EQ12} \mathcal{F}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ22, EQ23, EQ4, EQ4b\}$.

Finally, similar results can be shown when considering an attack relation satisfying the two properties $C1'$ and $C2'$ and the equivalence relation $\approx_3$ between arguments.

Proposition 3.2.9. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1'$ and $C2'$. For all $a, a', b, b' \in \mathcal{A}$, $(a \approx_3 a'$ and $b \approx_3 b') \Rightarrow (a\mathcal{R}b$ iff $a'\mathcal{R}b')$.

The following proposition shows that equivalent arguments w.r.t. $\approx_3$ belong to the same extensions in an argumentation framework whose attack relation satisfies $C1'$ and $C2'$.

Proposition 3.2.10. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1'$ and $C2'$. For all $a, a' \in \mathcal{A}$, if $a \approx_3 a'$ then $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $a \in \mathcal{E}$ iff $a' \in \mathcal{E}$.

A similar result as Proposition 3.2.8 is found in case of argumentation frameworks with attack relations satisfying $C1'$ and $C2'$ and using the equivalence relation $\approx_3$.

Proposition 3.2.11. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1'$ and $C2'$, and $\mathcal{F} \equiv_{EQ13} \mathcal{F}'$. For all $a \in \mathcal{A}$ and for all $a' \in \mathcal{A}'$, if $a \approx_3 a'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a', \mathcal{F}')$.

Finally, we show that if two argumentation frameworks whose attack relations enjoy $C1'$ and $C2'$ are equivalent w.r.t. $EQ13$, then they are also equivalent w.r.t. $EQ23$ and $EQ4b$.

Theorem 3.2.4. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1'$ and $C2'$. If $\mathcal{F} \equiv_{EQ13} \mathcal{F}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ23, EQ4b\}$.

In sum, the comparative study revealed that the most useful equivalence criteria are EQ11, EQ12 and EQ13, since they are at the same time flexible (contrary to, for example EQ1) and general (i.e. they are based on extensions and imply many other criteria). This is why, in the next subsection, we provide conditions under which two frameworks are equivalent w.r.t. those three criteria.

3.2.3 Conditions for equivalence

In subsection 3.2.1, we have proposed different criteria for the equivalence of two argumentation frameworks built from the same logic. An important question now is: *"Are there conditions under which two distinct argumentation frameworks are equivalent with respect to those criteria?"* Recall that in case of the criteria proposed by Oikarinen and Woltran (2010) the answer is negative. In this section, we show that our refined criteria make it possible to compare different frameworks.

In the rest of the subsection, we will study the case of two argumentation frameworks that may be built from two distinct knowledge bases but use the same attack relation (e.g. both frameworks use 'undercut'). Recall that $\mathtt{Arg}(\mathcal{L})$ is the set of all arguments that can be built from a fixed Tarskian logic $(\mathcal{L}, \mathsf{CN})$. We denote by $\mathcal{R}(\mathcal{L})$ the attack relation which is used in the two frameworks with $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$. The following result shows when two argumentation frameworks $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$, are equivalent w.r.t. $EQ11$.

Theorem 3.2.5. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\mathtt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \mathtt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1'$ and $C2$ and $\mathcal{A} \sim_1 \mathcal{A}'$, then $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$.

The following result follows from the previous result and Theorem 3.2.2.

Corollary 3.2.1. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\mathtt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \mathtt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1'$ and $C2$ and $\mathcal{A} \sim_1 \mathcal{A}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ13, EQ21, EQ23, EQ31, EQ33, EQ4b, EQ5b, EQ6, EQ6b\}$.

A similar result is shown for argumentation frameworks which use the same attack relation provided that the latter satisfies properties $C1$ and $C2'$.

Theorem 3.2.6. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\mathtt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \mathtt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1$ and $C2'$ and $\mathcal{A} \sim_2 \mathcal{A}'$, then $\mathcal{F} \equiv_{EQ12} \mathcal{F}'$.

As a consequence of the previous result and dependencies between criterion $EQ12$ and other equivalence criteria, which are proved in Theorem 3.2.3, the next result holds.

Corollary 3.2.2. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\mathtt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \mathtt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1$ and $C2'$ and $\mathcal{A} \sim_2 \mathcal{A}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ13, EQ22, EQ23, EQ32, EQ33, EQ4, EQ4b, EQ5, EQ5b, EQ6b\}$.

The following result shows under which conditions two frameworks are equivalent w.r.t. $EQ13$.

Theorem 3.2.7. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\mathtt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \mathtt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1'$ and $C2'$ and $\mathcal{A} \sim_3 \mathcal{A}'$, then $\mathcal{F} \equiv_{EQ13} \mathcal{F}'$.

The following follows from the previous result and Theorem 3.2.4.

Corollary 3.2.3. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\mathtt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \mathtt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1'$ and $C2'$ and $\mathcal{A} \sim_3 \mathcal{A}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ23, EQ33, EQ4b, EQ5b, EQ6b\}$.

Note that results proved for attack relations verifying C1' and C2, equivalence relation $\sim_1$ and corresponding equivalence criteria (e.g. EQ11) have their counterparts for attack relations verifying C1 and C2' (resp. C1' and C2'), equivalence relation $\sim_2$ (resp. $\sim_3$) and corresponding equivalence criteria, e.g. EQ12 (resp. EQ13). From now on, we will concentrate on attack relations verifying C1' and C2 (and corresponding equivalence relations).

This choice is motivated by the fact that this class of attack relations corresponds to equivalence relation $\sim_1$ between arguments. This relation is a good compromise between two extreme standpoints, one being asking for C1 and C2, which is too rigid, and the second one being asking for C1' and C2'. At this point, we will argue why we prefer to continue our study with relations verifying C1' and C2 (and relation $\sim_1$) instead of those verifying C1' and C2' (and the corresponding relation $\sim_3$). In logical based argumentation, arguments are supposed to be constructed from a knowledge base, which contains some information. Thus, two arguments having different (but equivalent) formulae in their support use formulae coming possibly from different sources. Moreover, in preference-based argumentation, one of those formulae may be stronger than another, and consequently the first argument may be stronger than the second one. For all these reasons, we prefer not to call them equivalent. As for C1 and C2', this is the least appealing choice since the subtleties in the supports of arguments are not taken into account, while too much attention is drawn to differences in conclusions. On the contrary, we want to keep information about the argument's support (i.e. hypotheses used in its reasoning) while we want to eliminate (usually) infinitely many alternative (but equivalent) conclusions.

Related to that point is our second remark. Namely, a knowledge base from which arguments are built is often supposed to be finite. Thus, a finite number of arguments' supports is available. On the contrary, almost all well-known logics allow for an infinite number of conclusions which can be drawn from a (finite or infinite) set of formulae. Thus, we can keep the information which will allow us to distinguish between equivalent but different formulae in the support, but if we want to be able to reduce a framework to an equivalent, finite one, we will have to reduce the number of conclusions, as will be shown later in this chapter.

The third comment we want to make here is the link between relation $\sim_1$, which corresponds to the class of attack relations verifying C1' and C2, and the notion of conservatism defined by Besnard and Hunter (2008). According to that definition, an argument (H, h) is *more conservative* than (H', h') iff $H \subseteq H'$ and $h' \vdash h$. Thus, two arguments are equivalent in the sense of $\sim_1$ iff they are more conservative than the other (i.e. the first one is more conservative than the second and vice versa).

Note also that, regardless of all the previous comments, for the majority of results which will be presented in the rest of the chapter, similar ones can be proved for relations verifying C1 and C2' (or C1' and C2').

3.3 Strong equivalence

In this subsection, we study the strong equivalence between logic-based argumentation frameworks. As mentioned before, two argumentation frameworks are strongly equivalent iff after adding the same set of arguments to both frameworks, the new frameworks are equivalent w.r.t. a given equivalence criterion.

We will be working with a framework enriched with new arguments. One of the essential questions when a new piece of information arrives is to know which arguments attack/are attacked by it. Since arguments are built from a logical language $\mathcal{L}$, and for all well-known logics $\texttt{Arg}(\mathcal{L})$ contains an infinite number of arguments, it is reasonable to suppose that the attack relation is defined by using some rule/principle allowing us to know when two arguments attack each other instead of manually specifying the attack relation $\mathcal{R}$ on $\mathcal{A}$. Thus, in the rest of the chapter, we will suppose that a general attack relation is defined on the set of all arguments $\texttt{Arg}(\mathcal{L})$ that can be built from the logic $(\mathcal{L}, \mathsf{CN})$. For any pair of arguments, this relation specifies whether they attack each other. As before, this relation will be denoted by $\mathcal{R}(\mathcal{L})$; thus $\mathcal{R}(\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$. As already mentioned, in the rest of the chapter we will suppose that $\mathcal{R}(\mathcal{L})$ verifies C1' and C2. Once this attack relation is defined, we suppose that for any argumentation framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, relation $\mathcal{R}$ is a restriction of $\mathcal{R}(\mathcal{L})$ on $\mathcal{A}$, i.e. $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$.

Augmenting a given argumentation framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ by an arbitrary set $\mathcal{B}$ of arguments results in a new framework, denoted by $\mathcal{F} \oplus \mathcal{B}$, where $\mathcal{F} \oplus \mathcal{B} = (\mathcal{A}_b, \mathcal{R}_b)$ with $\mathcal{A}_b = \mathcal{A} \cup \mathcal{B}$ and $\mathcal{R}_b = \mathcal{R}(\mathcal{L})|_{\mathcal{A}_b}$.

Definition 3.3.1 (Strong equivalence between two argumentation frameworks)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built using the same Tarski's logic $(\mathcal{L}, \mathsf{CN})$ and let EQx be an equivalence criterion. The two frameworks $\mathcal{F}$ and $\mathcal{F}'$ are* strongly equivalent *w.r.t. EQx iff*

$$\forall \mathcal{B} \subseteq \mathtt{Arg}(\mathcal{L}), \quad \mathcal{F} \oplus \mathcal{B} \equiv_{EQx} \mathcal{F}' \oplus \mathcal{B}.$$

We will also use notation $\mathcal{F} \equiv_{EQxS} \mathcal{F}'$ ('S' stands for strong*) as a synonym for the phrase "$\mathcal{F}$ and $\mathcal{F}'$ are strongly equivalent w.r.t. EQx".*

The following result is an obvious consequence of the previous definition.

Proposition 3.3.1. Let EQx be an arbitrary equivalence criterion (from Definition 3.2.1 or Definition 3.2.5), and let $\mathcal{F}$ and $\mathcal{F}'$ be two argumentation frameworks. If $\mathcal{F} \equiv_{EQxS} \mathcal{F}'$ then $\mathcal{F} \equiv_{EQx} \mathcal{F}'$.

For example, from the previous result and Theorems 3.2.1 and 3.2.2, we see that if $\mathcal{F} \equiv_{EQ11S} \mathcal{F}'$, then $\mathcal{F} \equiv_{EQ21S} \mathcal{F}'$ and $\mathcal{F} \equiv_{EQ31S} \mathcal{F}'$.

Corollary 3.3.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from a Tarskian logic $(\mathcal{L}, \mathsf{CN})$, s.t. $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ verifies C1' and C2, $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$ and $\mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{F} \equiv_{EQ11S}$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ13S, EQ21S, EQ23S, EQ31S, EQ33S, EQ4bS, EQ5bS, EQ6S, EQ6bS\}$.

The previous result showed under which conditions one form of strong equivalence implies other forms of strong equivalence between two argumentation frameworks. We will now show that the condition we introduced in the previous section is general enough to guarantee the strong equivalence.

Theorem 3.3.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from a Tarskian logic $(\mathcal{L}, \mathsf{CN})$, s.t. $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ verifies C1' and C2, $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$ and $\mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{A} \sim_1 \mathcal{A}'$, then $\mathcal{F} \equiv_{EQ11S} \mathcal{F}'$.

A consequence of the previous theorem is the following.

Corollary 3.3.2. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from a Tarskian logic $(\mathcal{L}, \mathsf{CN})$, s.t. $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ verifies C1' and C2, $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$ and $\mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{A} \sim_1 \mathcal{A}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$, with $x \in \{EQ13S, EQ21S, EQ23S, EQ31S, EQ33S, EQ4bS, EQ5bS, EQ6S, EQ6bS\}$.

Note, however, that this does *not* mean that notions of equivalence and strong equivalence coincide, as illustrated by the following example.

Example 3.3.1. *Let* $(\mathcal{L}, \mathsf{CN})$ *be a Tarskian logic defined as* $\mathcal{L} = \{rock, paper, book, scissors, \bot\}$*, and let for all* $X \subseteq \mathcal{L}$*,*

$$\mathsf{CN}(X) = \begin{cases} X, & \text{if } \bot \notin X \text{ and } |X| \leq 1 \\ \mathcal{L}, & \text{else} \end{cases}$$

As expected, rock *attacks* scissors*,* scissors *attack* paper*, and of course,* paper *attacks* rock*. We also suppose that* scissors *attack* book*. Formally,* $\forall a, b \in \mathtt{Arg}(\mathcal{L})$, $a\mathcal{R}(\mathcal{L})b$ *iff*

- $(\mathtt{Conc}(a) = scissors$ *and* $(paper \in \mathtt{Supp}(b)$ *or* $book \in \mathtt{Supp}(b)))$*, or*
- $(\mathtt{Conc}(a) = rock$ *and* $scissors \in \mathtt{Supp}(b))$*, or*
- $(\mathtt{Conc}(a) = paper$ *and* $rock \in \mathtt{Supp}(b))$*.*

It is easy to see that $\mathcal{R}(\mathcal{L})$ *verifies C1' and C2. Let* $r = (\{rock\}, rock)$*,* $p = (\{paper\}, paper)$*,* $b = (\{book\}, book)$*,* $s = (\{scissors\}, scissors)$*. We define* $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ *with* $\mathcal{A} = \{p, s\}$ *and* $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$ *(i.e.* $\mathcal{R} = \{(s, p)\}$*) and* $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ *with* $\mathcal{A}' = \{b, s\}$ *and* $\mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$ *(i.e.* $\mathcal{R}' = \{(s, b)\}$*). The only extension of* $\mathcal{F}$ *is set* $\{s\}$*, which is also the only extension of* $\mathcal{F}'$*. Thus, the two frameworks are equivalent w.r.t. all the criteria from Definition 3.2.1 and Definition 3.2.5. However, they are not strongly equivalent w.r.t. any criteria from Definition 3.2.1 or Definition 3.2.5. Namely, if both frameworks are augmented with* $\{r\}$*, framework* $\mathcal{F} \oplus \{r\}$ *has no extensions and all its arguments are rejected, while* $\mathcal{F}' \oplus \{r\}$ *has a unique extension* $\{r, b\}$*.*

3.4 Core(s) of an argumentation framework

In this subsection, we will show how to define a core of an argumentation framework, that is, to define its "sub-framework" which is equivalent to the original one. We also provide a condition under which an argumentation framework has a *finite* core. The basic idea is to simplify a given argumentation framework by taking exactly one argument from each equivalence class of $\mathcal{A}/\sim_1$.

We use the standard notation, i.e. given a set X and an equivalence relation $\sim$ on that set, $\forall x \in X$, we write $[x] = \{x' \in X \mid x' \sim x\}$ and $X/\sim = \{[x] \mid x \in X\}$. Recall also that we suppose a general attack relation $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ which satisfies C1' and C2, and that for any framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ we have $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$.

Definition 3.4.1 (Core of an argumentation framework)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework. An argumentation framework $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ is a* core *of $\mathcal{F}$ iff:*

- $\mathcal{A}' \subseteq \mathcal{A}$
- $\forall C \in \mathcal{A}/\sim_1$, $\exists! a \in C \cap \mathcal{A}'$
- $\mathcal{R}' = \mathcal{R}|_{\mathcal{A}'}$*, i.e.* $\mathcal{R}'$ *is the restriction of* $\mathcal{R}$ *on* $\mathcal{A}'$*.*

The fact that one representative of each equivalence class is included in a core allows us to show that any core of an argumentation framework is equivalent with the original framework.

Theorem 3.4.1. Let $\mathcal{F}$ be an argumentation framework and $\mathcal{F}'$ one of its cores. Then: $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$.

As a consequence of the previous result and Theorems 3.2.1 and 3.2.2, we see that all important outputs of an argumentation framework and any of its cores coincide.

Corollary 3.4.1. Let $\mathcal{F}$ be an argumentation framework and $\mathcal{F}'$ one of its cores. Then:

- $\mathtt{Sc}(\mathcal{F}) \sim_1 \mathtt{Sc}(\mathcal{F}')$
- $\mathtt{Cr}(\mathcal{F}) \sim_1 \mathtt{Cr}(\mathcal{F}')$
- $\mathtt{Output}_{sc}(\mathcal{F}) \cong \mathtt{Output}_{sc}(\mathcal{F}')$
- $\mathtt{Output}_{cr}(\mathcal{F}) \cong \mathtt{Output}_{cr}(\mathcal{F}')$
- $\mathtt{Bases}(\mathcal{F}) = \mathtt{Bases}(\mathcal{F}')$

We can also see that if $\mathcal{F}'$ is a core of $\mathcal{F}$, then each argument of $\mathcal{F}'$ has the same status in $\mathcal{F}$ and in $\mathcal{F}'$ (this follows from Proposition 3.2.5).

We will now show that when an attack relation verifies C1' and C2 then two arguments having the same support have the same status. This means that if a given standpoint (i.e. the set of hypotheses) is accepted, then all of its consequences (i.e. conclusions) must be accepted as well.

Proposition 3.4.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and let $a, a' \in \mathcal{A}$ be two arguments such that $\texttt{Supp}(a) = \texttt{Supp}(a')$. Then:
$\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a', \mathcal{F})$.

A core of an argumentation framework provides enough information to know the status of a given argument. First, we will show that if a given argument is in the core, then its status in the core is the same as in the original framework. Furthermore, we will also show how to determine a status of an argument *not* belonging to a given core.

Proposition 3.4.2. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ its core.

- If $a \in \mathcal{A}'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a, \mathcal{F}')$,
- If $a \notin \mathcal{A}'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(b, \mathcal{F}')$, where $b \in \mathcal{A}'$ is an arbitrary argument s.t. $\texttt{Supp}(a) = \texttt{Supp}(b)$.

Note also that different cores return equivalent results. This comes from the transitivity of equivalence relations between argumentation frameworks. So, if $\mathcal{F}$ is an argumentation framework and $\mathcal{F}'$ and $\mathcal{F}''$ its cores, then from $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$ and $\mathcal{F} \equiv_{EQ11} \mathcal{F}''$, we have $\mathcal{F}' \equiv_{EQ11} \mathcal{F}''$.

We now provide a condition which guarantees that any core of any argumentation framework built from a finite knowledge base is finite. This is the case for logics in which any consistent finite set of formulae has finitely many logically non-equivalent consequences. To formalize this, we use the following notation for a set of logical consequences made from consistent subsets of a given set. For any $X \subseteq \mathcal{L}$, $\texttt{Cncs}(X) = \{x \in \mathcal{L} \mid \exists Y \subseteq X \text{ s.t. } \mathsf{CN}(Y) \neq \mathcal{L} \text{ and } x \in \mathsf{CN}(Y)\}$.

We show that if $\texttt{Cncs}(\Sigma)$ has a finite number of equivalence classes, then any core of $\mathcal{F}$ is finite (i.e. has a finite set of arguments).

Theorem 3.4.2. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework built over a knowledge base Σ (i.e. let $\mathcal{A} \subseteq \texttt{Arg}(\Sigma)$). If $\texttt{Cncs}(\Sigma)/\equiv$ is finite, then any core of $\mathcal{F}$ is finite.

3.4.1 Core(s) in propositional logic

In this subsection, we will consider a particular case of the general framework we have studied so far. More precisely, we will study an argumentation framework based on *propositional logic*. Furthermore, we will suppose that the attack relation $\mathcal{R}(\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$ is defined as follows.

Definition 3.4.2 (Undercut)**.** *Let* $a, b \in \texttt{Arg}(\mathcal{L})$*. We say that* $a\mathcal{R}b$ *iff* $\exists h \in \texttt{Supp}(b)$ *s.t.* $a \equiv \neg h$*.*

Until the end of this subsection, we suppose that $\mathcal{R}(\mathcal{L})$ is as in the previous definition.

It can be checked that the condition of the previous theorem (i.e. that $\texttt{Cncs}(\Sigma)/\equiv$ is finite) is almost never verified by propositional logic (more precisely, it is not verified iff Σ contains at least one consistent formula). We provide a simple counter-example.

Example 3.4.1. *Let* $(\mathcal{L}, \mathsf{CN})$ *be propositional logic and let* $\Sigma = \{x\}$*.* $\texttt{Cncs}(\Sigma)$ *contains following formulae:* $x, x \vee z_1, x \vee z_2, x \vee z_3 \ldots$ *It is clear that in this case* $\texttt{Cncs}(\Sigma)/\equiv$ *is infinite.*

Luckily, Theorem 3.4.2 can easily be adapted to suit propositional logic (and many other well-known logics). As expected, the basic idea is to limit the number of variables which are used for the construction of arguments.

Roughly speaking, there are two sources of infiniteness for $\texttt{Arg}(\Sigma)$. The first one is due to logically equivalent conclusions and can be illustrated by the following series of arguments: $(\{x\}, x), (\{x\}, x \wedge x), (\{x\}, x \wedge x \wedge x), \ldots$ $(\{x\}, x \wedge (y \to y)), (\{x\}, x \wedge (y \vee \neg y)), (\{x\}, x \wedge ((y \to z) \leftrightarrow (\neg z \to \neg y))), \ldots$ This includes arguments with the same support and different but logically equivalent conclusions. It is easy to see that their number is infinite. The second source infiniteness is the fact that new atoms may be introduced in the conclusion. For example, if $x \in \Sigma$, then $\texttt{Arg}(\Sigma)$ contains (but is not limited to) the following arguments: $(\{x\}, x \vee z_1), (\{x\}, x \vee z_2), (\{x\}, x \vee z_3), (\{x\}, x \vee z_4), \ldots$

We have already formalized the equivalence between arguments having equivalent conclusions. Now, we will show how to limit a number of atoms in arguments' conclusions so that a core is finite. The idea is to take only arguments which are built entirely on atoms from Σ. We will first formalize this idea and then show that eliminating those arguments will allow to construct a finite core without losing any important information present in the original framework.

Let us use the following notations. $\mathtt{Atoms}(\Sigma)$ is the set of atoms occurring in Σ. $\mathtt{Arg}(\Sigma)_\downarrow$ is the subset of $\mathtt{Arg}(\Sigma)$ that contains only arguments with conclusions in the language of Σ. For instance, if $\Sigma = \{x \to y, z \vee \neg w\}$ then $\mathtt{Atoms}(\Sigma) = \{x, y, z, w\}$. Thus, an argument such as $(\{x \to y\}, (\neg x \vee y) \vee t)$ does not belong to the set $\mathtt{Arg}(\Sigma)_\downarrow$. From now on, when Σ is fixed, and if not explicitly stated otherwise, we will also use the notation $\mathcal{F}_\downarrow = (\mathtt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow)$ with $\mathcal{R}_\downarrow = \mathcal{R}(\mathcal{L})|_{\mathtt{Arg}(\Sigma)_\downarrow}$. Note that the set $\mathtt{Arg}(\Sigma)_\downarrow$ is infinite (due to equivalent arguments).

Importantly, its arguments have the same status in the two frameworks $\mathcal{F} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ and $\mathcal{F}_\downarrow = (\mathtt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow)$.

Theorem 3.4.3. Let $\mathcal{F} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ and $\mathcal{F}_\downarrow = (\mathtt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow)$. For all $a \in \mathtt{Arg}(\Sigma)_\downarrow$, $\mathtt{Status}(a, \mathcal{F}) = \mathtt{Status}(a, \mathcal{F}_\downarrow)$.

This result is important since it shows that arguments that use external variables (i.e. variables which are not in $\mathtt{Atoms}(\Sigma)$) in their conclusions can be omitted from the reasoning process. Moreover, we show next that their status is still known. It is that of any argument in $\mathtt{Arg}(\Sigma)_\downarrow$ with the same support.

Theorem 3.4.4. Let $\mathcal{F} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an argumentation framework built over Σ. For all $a \in \mathtt{Arg}(\Sigma) \setminus \mathtt{Arg}(\Sigma)_\downarrow$, $\mathtt{Status}(a, \mathcal{F}) = \mathtt{Status}(b, \mathcal{F})$ where $b \in \mathtt{Arg}(\Sigma)_\downarrow$ and $\mathtt{Supp}(a) = \mathtt{Supp}(b)$.

In sum, Theorem 3.4.3 and Theorem 3.4.4 clearly show that one can use the sub-framework $\mathcal{F}_\downarrow = (\mathtt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow)$ instead of $\mathcal{F} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ without losing any information. However, this framework is still infinite due to redundant arguments.

The following result proves that the set $\mathtt{Arg}(\Sigma)_\downarrow$ is partitioned into a finite number of equivalence classes w.r.t. the equivalence relation $\sim_1$.

Proposition 3.4.3. It holds that $|\mathtt{Arg}(\Sigma)_\downarrow / \approx_1| \leq 2^n \cdot 2^{2^m}$, where $n = |\Sigma|$ and $m = |\mathtt{Atoms}(\Sigma)|$.

This result is of great importance since it shows how it is possible to partition an infinite set of arguments into a finite number of classes. Note that each class may contain an infinite number of arguments. An example of such infinite class is the one which contains (but is not limited to) all the arguments having $\{x\}$ as a support and $x, x \wedge x, \ldots$ as conclusions.

Until now, we have shown that arguments from $\mathtt{Arg}(\Sigma) \setminus \mathtt{Arg}(\Sigma)_\downarrow$ may be omitted. Now, we show that when only atoms from Σ are used, any core of any argumentation framework built over a finite knowledge base is finite. The result is a direct consequence of Proposition 3.4.3 and Definition 3.4.1.

Proposition 3.4.4. Given $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, such that $\mathcal{A} \subseteq \mathtt{Arg}(\Sigma)_{\downarrow}$, any core of $\mathcal{F}$ has a finite number of arguments.

The last question to be answered in this subsection is the following: *"If we do not include the arguments having atoms not belonging to Σ, and if we want use argumentation for reasoning, i.e. to calculate all sceptical conclusions, is it possible to obtain all sceptical conclusions of the original framework by using its core?"* The answer is given in the next theorem, as we show that any conclusion of the original framework can be deduced from the conclusions of the framework using only atoms from Σ. Note that now we suppose that argumentation is used for calculating all conclusions, thus, $\mathcal{A} = \mathtt{Arg}(\Sigma)$.

Theorem 3.4.5. Let $\mathcal{F} = (\mathcal{A} = \mathtt{Arg}(\Sigma), \mathcal{R})$ be an argumentation framework built over a knowledge base Σ, let $\mathcal{F}' = (\mathcal{A}' = \mathtt{Arg}(\Sigma)_{\downarrow}, \mathcal{R}')$, with $\mathcal{R}' = \mathcal{R}|_{\mathcal{A}'}$, and let $\mathcal{G}$ be a core of $\mathcal{F}'$. Then, $\mathtt{Output}_{sc}(\mathcal{F}) = \{x \in \mathcal{L} \text{ s.t.} \mathtt{Output}_{sc}(\mathcal{G}) \vdash x\}$.

An important question now is how to choose a core, i.e. how to pick exactly one formula from each set of logically equivalent formulae? Since a lexicographic order on set $\mathcal{L}$ is usually available, we can take the first formula from that set according to that order. Instead of defining a lexicographic order, one could also choose to take the disjunctive (or conjunctive) normal form of a formula.

3.5 Application on dynamic frameworks

In many situations, some arguments are built, their statuses are calculated and (sceptical/credulous) conclusions of the argumentation framework are computed. In this subsection, we suppose that an argumentation framework is given and we study the impact of a new argument on the argumentation framework, in particular on the status of existing arguments and its outputs. We will also show when it is possible to know the status of the arriving argument(s) without having to recalculate the extensions of the framework.

Recall that we suppose a Tarskian logic $(\mathcal{L}, \mathsf{CN})$ and that a general function $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ verifying C1' and C2 is given. We also suppose that for any argumentation framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, we have $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$.

Considering a knowledge base corresponding to a given argumentation framework, if not explicitly stated otherwise, in this section we suppose that $\mathcal{F}$ is an arbitrary argumentation framework and $\Sigma = \mathtt{Base}(\mathcal{A})$.

We will study two situations. In the first one, we suppose that framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ is augmented with a new set of arguments $\mathcal{E}$; thus, we obtain a new framework, which will be denoted by $\mathcal{F} \oplus \mathcal{E}$. Recall that we have already defined operator $\oplus$ for merging an argumentation framework with a set of arguments on page 36. Similarly, we will define operator $\ominus$ in an expected way: $\mathcal{F} \ominus \mathcal{E} = (\mathcal{A}', \mathcal{R}')$ with $\mathcal{A}' = \mathcal{A} \setminus \mathcal{E}$ and $\mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$.

We now identify a case in revision when it is not necessary to recalculate arguments' statuses.

We will show that if argumentation framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ contains a core of argumentation framework $(\mathcal{A}' = \mathtt{Arg}(\Sigma), \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'})$ then arguments built from Σ have no impact on revision process.

Definition 3.5.1. *If $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{G}$ are argumentation frameworks, we say that $\mathcal{F}$ contains a core of $\mathcal{G}$ iff there exists an argumentation framework $\mathcal{H} = (\mathcal{A}_h, \mathcal{R}_h)$ s.t. $\mathcal{A}_h \subseteq \mathcal{A}$ and $\mathcal{R}_h = \mathcal{R}(\mathcal{L})|_{\mathcal{A}_h}$ and $\mathcal{H}$ is a core of $\mathcal{G}$.*

Theorem 3.5.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework which contains a core of $\mathcal{G} = (\mathcal{A}_g = \mathtt{Arg}(\Sigma), \mathcal{R}_g = \mathcal{R}(\mathcal{L})|_{\mathcal{A}_g})$ and let $\mathcal{E} \subseteq \mathtt{Arg}(\Sigma)$. Then:

- $\mathcal{F} \equiv_{EQ11} \mathcal{F} \oplus \mathcal{E}$
- $\forall a \in \mathcal{A}, \mathtt{Status}(a, \mathcal{F}) = \mathtt{Status}(a, \mathcal{F} \oplus \mathcal{E})$
- $\forall e \in \mathcal{E} \setminus \mathcal{A}, \mathtt{Status}(e, \mathcal{F} \oplus \mathcal{E}) = \mathtt{Status}(a, \mathcal{F})$, where $a \in \mathcal{A}$ is any argument s.t. $\mathtt{Supp}(a) = \mathtt{Supp}(e)$.

It is clear that the previous theorem is applicable when $\mathcal{F}$ is itself a core of $\mathcal{G} = (\mathcal{A}_g = \mathtt{Arg}(\Sigma), \mathcal{R}_g = \mathcal{R}(\mathcal{L})|_{\mathcal{A}_g})$ and $\mathcal{E} \subseteq \mathtt{Arg}(\Sigma)$.

We will now show that when a framework does not contain a core of the framework built over its base, new arguments may change the status of the existing ones.

Example 3.5.1. *Let $(\mathcal{L}, \mathsf{CN})$ be the propositional logic and let the attack relation $\mathcal{R}(\mathcal{L})$ be defined as: $\forall a, b \in \mathtt{Arg}(\mathcal{L})$, $a\mathcal{R}(\mathcal{L})b$ iff $\exists h \in \mathtt{Supp}(b)$ s.t. $\mathtt{Conc}(a) \equiv \neg h$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ with $\mathcal{A} = \{a_1 = (\{strad, strad \rightarrow exp\}, exp), a_2 = (\{\neg strad\}, \neg strad)\}$. Recall that we suppose that $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$; thus, $\mathcal{R} = \{(a_2, a_1)\}$. Argument a_2 is sceptically accepted and a_1 is rejected. Let $e = (\{strad\}, strad)$. It is clear that $e \in \mathtt{Arg}(\mathtt{Base}(\mathcal{A}))$. However, statuses of a_1 and a_2 change in $\mathcal{F} \oplus \{e\}$. Namely, in the revised framework there are neither sceptically accepted nor rejected arguments; all arguments are credulously accepted.*

The previous example illustrated a situation when an argumentation framework does not contain a core of the framework constructed from its base. This means that not all available information is represented in $\mathcal{F}$; thus, it is not surprising that it is possible to revise arguments' statuses.

We have already seen that extracting a core of an argumentation framework is a compact way to represent the original framework. In that process, arguments are deleted from the original framework. In some situations, one would prefer to say the same thing in several different ways, since it can be useful in a given situation; for example in a dialogue. In other situations, we want to get rid of some superfluous arguments. We show under which conditions deleting argument(s) does not influence the status of other arguments.

As expected, if a set of arguments $\mathcal{E}$ is deleted from $\mathcal{F}$ and if a resulting framework $\mathcal{F} \ominus \mathcal{E}$ contains a core of $(\mathtt{Arg}(\Sigma), \mathcal{R}(\mathcal{L})|_{\mathtt{Arg}(\Sigma)})$, then statuses of remaining arguments do not change. The following corollary follows from Theorem 3.5.1.

Corollary 3.5.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and let $\mathcal{E} \subseteq \mathcal{A}$. If $\mathcal{F} \ominus \mathcal{E}$ contains a core of $\mathcal{G} = (\mathcal{A}_g = \mathtt{Arg}(\Sigma), \mathcal{R}_g = \mathcal{R}(\mathcal{L})_{\mathcal{A}_g})$, then:

- $\mathcal{F} \equiv_{EQ11} \mathcal{F} \ominus \mathcal{E}$
- $\forall a \in \mathcal{A} \setminus \mathcal{E}$, $\mathtt{Status}(a, \mathcal{F}) = \mathtt{Status}(a, \mathcal{F} \ominus \mathcal{E})$.

The obvious consequence of the above result is that if $\mathcal{F} \ominus \mathcal{E}$ is itself one of the cores of $\mathcal{G}$, then the statuses of its arguments are not changed after the deletion of arguments from $\mathcal{E}$.

Note that in some works in the literature, behavior of an argumentation framework after addition or removal of an attack is studied. We did not consider this option since in logic-based argumentation it does not make sense. Namely, we supposed that arguments are built from a logical language, and that whether an argument attacks another is determined using logical properties of two given arguments (e.g. union of the conclusion of the first argument and the support of the second argument are inconsistent). Thus, it is not possible to become aware of a conflict in one moment without knowing that it has existed from the instant when the second of the two arguments was constructed.

3.6 Conclusion

In this chapter, we have tackled the question: *"When are two argumentation frameworks equivalent?"* First, we showed how to define equivalence between formulae, arguments, sets of formulae and sets of arguments. We have then used those equivalence relations to define equivalence criteria between argumentation frameworks. Links between criteria have also been investigated. Particularly important results are those which show under which conditions two frameworks are equivalent. We also considered strong equivalence between argumentation frameworks.

In the second part, we showed how to apply our results: first, they allow to reduce the number of arguments in an argumentation framework by obtaining an equivalent but smaller (in the terms of number of arguments) framework; second, we identified situations when adding new arguments to a framework does not influence statuses of existing ones and showed that in this case it is not necessary to recalculate extensions.

We have already noted that Oikarinen and Woltran (2010) have dealt with the problem of equivalence between argumentation frameworks. That work treated only strong equivalence in the abstract case when the structure of arguments is unknown. In that case, two frameworks are strongly equivalent if and only if they coincide, except if there are self-attacking arguments. In this thesis, thanks to taking into account the logical structure of arguments, we have identified cases when *different* argumentation frameworks are equivalent.

Equivalence between arguments and sets of arguments was also studied from the computational complexity perspective (Wooldridge, Dunne, and Parsons, 2006), in the case of propositional logic and one attack relation: undercut. According to Wooldridge et al. (2006), two arguments are logically equivalent iff their conclusions are logically equivalent. The main difference with our definitions is illustrated by the following example: Let $a = (\{y, y \to x\}, x)$ and $a' = (\{z, z \to x\}, x)$. According to Wooldridge et al. (2006), a and a' are equivalent, whereas they are not equivalent w.r.t. any of our criteria. Note that we do not consider them equivalent, since they are based on different hypotheses. It can be the case that one of those hypotheses is attacked and not the other one. For example, if $b = (\{\neg y\}, \neg y)$ then b undercuts a but not a'. This shows why that definition of equivalence is too simplistic for our purpose and is not sufficient to guarantee that all information from a knowledge base is represented in an argumentation framework.

Two sets of arguments X and Y are said to be equivalent if there is a bijection between between them f, s.t. $\forall x \in X,\ f(x)$ is equivalent with x (Wooldridge et al., 2006). In this thesis, we opted for a more flexible definition of equivalence. For example, let $X = \{(\{x\}, x), (\{x\}, \neg\neg x)\}$ and $Y = \{(\{x\}, x)\}$. We defined criteria which allow to say that those two sets are equivalent, while they are not equivalent w.r.t. the definition given by Wooldridge et al. (2006). This allows us to reduce an infinite framework to a finite one, which is impossible if using the definition demanding for a bijection between the two sets.

It should be also noted that Wooldridge et al. (2006) allow for an argument's support to be inconsistent and / or non-minimal. Thus, items 2 and 4 of Definition 2.3.3 are not verified.

Note also that in that paper, a problem of equivalence between two argumentation frameworks is not addressed. The focus of the work is on the computational complexity of different problems, e.g. the problem of checking whether an argument set is maximal (in the sense that no argument could be added without such an argument being logically equivalent to one that is already present).

Sur quelque préférence, une estime se fonde,
Et c'est n'estimer rien, qu'estimer tout le monde.

Alceste in *Le Misanthrope*, Molière

4 Preferences in argumentation frameworks

This Chapter studies the role of preferences in argumentation. In Section 4.2, after presenting some examples of preference relations, we argue that there are two roles of preferences in argumentation: conflict-resolution role and refining role. Then we present existing preference-based argumentation frameworks. In Section 4.3, we illustrate through several critical examples the drawbacks of existing frameworks which model the conflict-resolution role of preferences in argumentation. Then, we present our framework for the modeling of this role. That section uses and develops the results from several papers (Amgoud and Vesic, 2009b, 2010a). Section 4.4 presents the first framework in the literature which integrates both roles of preferences (Amgoud and Vesic, 2010c, 2011e). In Section 4.5, we show the links between well-known non-argumentative formalisms for handling inconsistency and two instantiations of our framework (Amgoud and Vesic, 2010b).

4.1 Introduction

Informally speaking, preference refers to ordering objects, on the basis of their "quality". Quality may be related to satisfaction or utility an object provides: if one is offered a drink, and (s)he can choose between orange juice, coffee and tea, one has to rank-order those three options (or, at least, identify the most preferred option) in order to choose what to drink. We say that one has to express his/her preferences.

A preference relation is a binary relation defined over a set X of objects. It is generally *reflexive* and *transitive* even if non-transitive preference relations exist. For example, if one prefers tea to coffee, and coffee to orange juice, it is reasonable to expect that (s)he prefers tea to orange juice. A reflexive and transitive relation is called *preorder*. A relation that compares any pair of objects in X is said to be *total*.

Formally, let $\geq$ be a preference relation on X, that is for $x, y \in X$, $x \geq y$

means that x is at least as good/preferred as y. If $x \geq y$ and $y \geq x$, then x and y are said to be *indifferent*. When not $x \geq y$ and not $y \geq x$, then x and y are said to be *incomparable*. The relation $\geq$ is total iff $\forall x, y \in X$, $x \geq y$ or $y \geq x$ (or both). A strict version of $\geq$ is denoted by $>$ and is defined as follows: for $x, y \in X$, $x > y$ iff $x \geq y$ and not $y \geq x$.

Example 4.1.1. *Let $X = \{oj, c, t\}$ where oj stands for orange juice, c for coffee and t for tea. If $\geq = \{(t,t), (oj, oj), (c, c)(t, oj), (t, c)\}$, then tea is the most preferred option. Note that orange juice and coffee are incomparable. Thus, this preference relation is not total. However, it is both reflexive and transitive.*

4.2 Preferences in argumentation

There is a clear consensus in the argumentation literature that arguments do not necessarily have the same strength. It may be the case that an argument relies on certain information while another argument is built from less certain ones, or that an argument promotes an important value while another promotes a weaker one. In both cases, the former argument is clearly stronger than the latter. These differences in arguments' strengths make it possible to compare them. Consequently, several preference relations between arguments have been defined in the literature (e.g. Amgoud, Cayrol, and LeBerre, 1996; Benferhat, Dubois, and Prade, 1993; Cayrol, Royer, and Saurel, 1993; Prakken and Sartor, 1997; Simari and Loui, 1992). There is also a consensus on the fact that preferences should be taken into account in the evaluation of arguments (Amgoud and Cayrol, 2002b; Bench-Capon, 2003; Modgil, 2009; Prakken and Sartor, 1997; Simari and Loui, 1992).

This section introduces examples of preference relations, studies the role of preferences in argumentation, and surveys the existing works in the area.

4.2.1 Examples of preference relations

In argumentation literature, several preference relations over arguments were defined. Those works often (but not always) assume a logic-based argumentation framework $(\mathcal{A}, \mathcal{R})$ built from a knowledge base Σ and under a monotonic logic $(\mathcal{L}, \mathsf{CN})$. They define a binary relation $\geq$ on $\mathcal{A}$ which expresses preferences between arguments of $\mathcal{A}$.

In the framework proposed by Bench-Capon (2003), each argument promotes a value, and the importance of an argument is equal to the importance of the value it promotes.

Definition 4.2.1. *Let* $(\mathcal{A}, \mathcal{R})$ *be an argumentation framework,* $\mathcal{V}$ *a set of values,* $\texttt{Pref} \subseteq \mathcal{V} \times \mathcal{V}$ *a preference relation over values and* $\texttt{val} : \mathcal{A} \to \mathcal{V}$ *a function which assigns to each argument the value it promotes. For* $a, b \in \mathcal{A}$, $a \geq b$ *iff* $(\texttt{val}(a), \texttt{val}(b)) \in \texttt{Pref}$.

Benferhat, Dubois, and Prade (1993) have proposed a preference relation based on the certainty of the formulae used as a support of an argument. This preference relation is based on *the weakest link principle*. The idea is that an argument is stronger (or equal) than another iff the weakest formula in the support of the first argument is better (or equal) than the weakest formula in the support of the second one. In this work, a knowledge base Σ contains propositional formulae. This base is equipped with a total order, that is it is *stratified* into $\Sigma_1 \cup \ldots \cup \Sigma_n$. such that $\forall i, j \in \{1, \ldots, n\}$ if $i \neq j$ then $\Sigma_i \cap \Sigma_j = \emptyset$. In other words, Σ is partitioned into a finite number of disjunct sets. Formulae in Σ_i have the same equality level and more certain that those in Σ_j iff $i \leq j$. The stratification of Σ enables to define a certainty level of each subset $\mathcal{S}$ of Σ. It is the highest number of stratum met by this subset. Formally:

$$\texttt{Level}(\mathcal{S}) = \max\{i \mid \mathcal{S} \cap \Sigma_i \neq \emptyset\} \text{ (with } \texttt{Level}(\emptyset) = 0).$$

The above certainty level is used in order to define a total preorder on the set of arguments that can be built from a stratified knowledge base.

Definition 4.2.2 (Weakest link principle)**.** *Let* $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ *be a stratified knowledge base. An argument* $(H, h) \in \texttt{Arg}(\Sigma)$ *is preferred to another argument* $(H', h') \in \texttt{Arg}(\Sigma)$*, denoted by* $(H, h) \geq_{wlp} (H', h')$*, iff* $\texttt{Level}(H) \leq \texttt{Level}(H')$.

Cayrol, Royer, and Saurel (1993) have extended this relation to the case when the knowledge base is equipped with a partial preorder, meaning that some formulae may be incomparable.

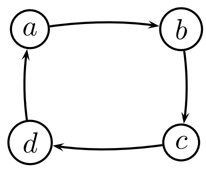
Definition 4.2.3 (Generalized weakest link principle)**.** *Let* Σ *be a knowledge base equipped with a partial preorder* $\trianglerighteq \subseteq \Sigma \times \Sigma$*. For two arguments* $(H, h), (H', h') \in \texttt{Arg}(\Sigma)$*, we write* $(H, h) \geq_{gwlp} (H', h')$ *iff* $\forall k \in H$, $\exists k' \in H'$ *such that* $k \triangleright k'$ *(i.e.* $k \trianglerighteq k'$ *and not* $(k' \trianglerighteq k)$*).*

Simari and Loui (1992) proposed another preference relation which privileges more specific information. Roughly speaking, this can be illustrated by letting an argument saying that Tweety does not fly since it is a penguin, be preferred to an argument saying that Tweety flies since it is a bird. This is since the former is grounded on the more specific rule than the latter.

4.2.2 Roles of preferences in argumentation

This subsection analyzes the role that preferences between arguments can play in an argumentation framework. We will discuss different critical examples. In this informal discussion, we will use terms *standard solutions* for the solutions calculated before taking into account the preferences and *preferred solutions* for the solutions calculated after preferences have been taken into account.[1]

Example 4.2.1. *Let us consider the argumentation framework depicted below. This framework has two standard stable extensions:* $\{a, c\}$ *and* $\{b, d\}$*.*



Assume that $a > b$ *and* $c > d$*. It can be argued that the stable extension* $\{a, c\}$ *is better than* $\{b, d\}$ *since for each element of* $\{b, d\}$ *there exists a better one in* $\{a, c\}$*. Thus, this preference-based argumentation framework would have only preferred solution* $\{a, c\}$ *as extension.*

Note that in Example 4.2.1, preferences *refine* the results obtained in the standard case. Indeed, the set of preferred solutions is a *subset of the set of the standard ones*. Preferences play here exactly the role described in non-monotonic reasoning formalisms (e.g. Brewka, Niemela, and Truszczynski, 2003). Let us now consider a different example.

Example 4.2.2. *Let* $\mathcal{A} = (\mathcal{A}, \mathcal{R})$ *with* $\mathcal{A} = \{a, b\}$ *and* $\mathcal{R} = \{(a, b)\}$*. This framework has one standard stable extension: the set* $\{a\}$*. Now, if we assume that* $b > a$*, it is clear that the standard solution cannot be refined and* $\{a\}$ *is the only preferred solution of the framework. What happened here is that the preferred argument is rejected when computing the standard solution (without taking preferences into account). Thus, there is no way to apply the preference of* b *over* a*.*
However, it is not intuitive to consider the set $\{a\}$ *as a preferred extension of the framework. Let us illustrate this by a less abstract example. Assume that the framework is built from a stratified propositional knowledge base* $\Sigma = \Sigma_1 \cup \Sigma_2$ *with* $\Sigma_1 = \{x, x \to y\}$ *and* $\Sigma_2 = \{z, z \to \neg x\}$*. Let* $a = (\{z, z \to$

[1] It should be clear that the word *preferred* does not refer to Dung's preferred semantics.

$\neg x\}, \neg x)$ $b = (\{x, x \to y\}, y)$. If the attack relation is the one which allows to undermine a premise of another argument, then a undermines b but not vice versa. If we use the preference relation which is based on the weakest link principle, then $b > a$. It is natural to expect that the conclusion y is justified and the argument b is accepted. Thus, the preferred solution of the framework should be the extension $\{b\}$.

Contrarily to Example 4.2.1, the use of preferences in Example 4.2.2 completely modifies the original set of extensions. Consequently, the set of preferred solutions of a framework is not necessarily a subset of the set of standard solutions. It can even be argued, that $\{b\}$ is the *standard* solution in the previous example, and consequently, the unique preferred solution.

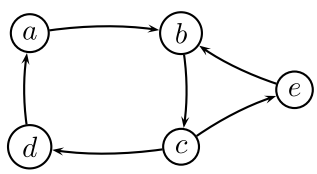
The two examples show that there are two distinct roles that preferences can play in an argumentation framework. They can be used in order to protect strong arguments from attacks coming from weaker ones. In this case, not all available preferences are exploited. Indeed, only preferences that contradict attacks are considered. In Example 4.2.2, the preference saying that a is preferred to b contradicts the attack from b to a. Such attacks are called "critical attacks".

Definition 4.2.4. *Let a, b be two arguments. There is a critical attack from a to b iff $a\mathcal{R}b$ and $b > a$.*

From now on, we will call solutions obtained after taking into account the first role of preferences *standard* solutions. The second role of preferences consists of exploiting the remaining preferences to refine the results obtained after having handled critical attacks.

The following example shows an argumentation framework in which *both* roles of preferences are needed.

Example 4.2.3. *Let us consider the argumentation framework depicted below.*



This framework has one stable extension: $\{a, c\}$. Assume now that $b > c$, $d > a$ and $b > e$. Note that only $b > e$ conflicts with the attack relation since e attacks b. Thus, only this preference is taken into account for computing *the two standard solutions* $\{a, c\}$ *and* $\{b, d\}$*. Consequently, the two remaining preferences may be used in order to refine the standard result and to prefer the extension* $\{b, d\}$*.*

To summarize, two roles of preferences are distinguished:

1. To weaken the *critical* attacks (i.e. the attacks which conflict with the preferences) in an AF, and thus to compute intuitive standard solutions.

2. To refine the standard solutions computed after considering the first role.

Example 4.2.2 shows that a refinement does not solve the problem of critical attacks whereas Example 4.2.3 shows that the first role is not sufficient and its results may need to be refined as the first role does not exploit all the available preferences.

4.2.3 Existing preference-based argumentation frameworks

As said before, there is an agreement in the literature that arguments do not necessarily have the same strength. Surprisingly, there are divergent opinions on whether the attack relation in Dung's framework already takes into account the strengths of arguments or should be augmented by a preference relation which captures these strengths. It is worth mentioning that Dung (1995) does not give an answer to this question. The only thing which is mentioned in that paper is that an argument can attack another argument meaning that it disqualifies this argument, and the two arguments cannot "survive" together.

According to some researchers, the attack relation in Dung's framework is a combination of a *symmetric* conflict relation and a preference relation between arguments (Kaci, van der Torre, and Weydert, 2006; Kaci, 2010). They argue that a conflict between two arguments should always be symmetric, and since Dung's attack relation may be asymmetric, this means that a preference relation is applied between the two arguments in order to solve the conflict.

According to other researchers (Amgoud, Caminada, Cayrol, Lagasquie, and Prakken, 2004), an argument can attack another argument by undermining one of its three basic components, that is its conclusion, a premise of its support, or a link between a premise and a conclusion. The *formal* definition of the first kind of attack induces a symmetric relation, e.g. *rebut*

(Elvang-Gøransson, Fox, and Krause, 1993), the two other kinds of attack induce asymmetric relations e.g. *assumption attack* (Elvang-Gøransson et al., 1993) or *undercut* (Pollock, 1992). Thus, the conflict relation mentioned by Kaci can be either symmetric or asymmetric. Besides, Amgoud and Besnard (2009) have shown that the choice of an attack relation is crucial for ensuring sound results, and should not be arbitrary. They have studied how to choose an attack relation when arguments are built using any logic satisfying Tarski's axioms. The results confirm that an attack relation *should not* be symmetric, in particular when the knowledge base from which arguments are built contains at least one minimal inconsistent subset with a cardinality higher than two. Indeed, symmetric relations lead to the violation of the rationality postulates identified by Caminada and Amgoud (2007). This means that the point of view defended by Kaci is not applicable, and confirms the hypothesis that attacks and preferences are two independent inputs of a preference-based argumentation framework. Thus, Dung's framework should be *extended* by preferences (at least for those applications which use a Tarskian logic for building arguments).

4.2.3.1 Handling critical attacks

We will now present the three most influential works that are done in the literature on the first role of preferences: preference-based frameworks, value-based frameworks and extended argumentation frameworks. As we will see, they all rely on the idea that preferences are used for (and only for) neutralizing attacks from weak arguments towards strong arguments.

Preference-based frameworks. (PAF) Amgoud and Cayrol (2002b) have proposed the first *abstract* preference-based argumentation framework. It takes as input a set $\mathcal{A}$ of arguments, an attack relation $\mathcal{R}$, and a preference relation $\geq$ between arguments which is abstract and can be instantiated in different ways. The basic idea behind these works is to ignore any attack coming from a weak argument towards a stronger one. This is formalised through a new relation between arguments, called *defeat*.

An argument defeats another iff the first one attacks the second one, and the second is not strictly preferred to the first one.

Definition 4.2.5. *Let* $(\mathcal{A}, \mathcal{R}, \geq)$ *be a* preference-based argumentation framework. *For* $a, b \in \mathcal{A}$*, argument* a *defeats argument* b*, denoted* $a\texttt{Def}b$ *iff* $a\mathcal{R}b$ *and not* $(b > a)$.

Extensions of a preference-based argumentation framework are then defined as extensions of $(\mathcal{A}, \texttt{Def})$.

Example 4.2.4. *Let* $\mathcal{A} = \{a, b\}$*,* $\mathcal{R} = \{(a, b)\}$ *and* $\geq = \{(a, a), (b, a), (b, b)\}$*. Then,* $\texttt{Def} = \emptyset$*. Consequently, this framework has a unique stable/preferred extension:* $\{a, b\}$*.*

Value-based argumentation frameworks. (VAF) This extension of Dung's framework was proposed by Bench-Capon (2003). It assumes that each argument promotes a value, and a preference between two arguments comes from the importance of the respective values that are promoted by the two arguments. For different audiences, different values are more or less important, which is formalized by defining an audience simply as an ordering on the set of values. An audience-specific value-based argumentation framework is defined as follows.

Definition 4.2.6. *An* audience-specific value-based argumentation framework *is a 5-tuple:* $(\mathcal{A}, \mathcal{R}, \mathcal{V}, \texttt{val}, Pref_{aud})$*, where* $\mathcal{A}$ *is a finite set of arguments,* $\mathcal{R}$ *is an irreflexive binary relation on* $\mathcal{A}$*,* $\mathcal{V}$ *is a nonempty set of values,* $\texttt{val} : \mathcal{A} \rightarrow \mathcal{V}$*, aud is an audience (i.e. an ordering on* $\mathcal{V}$*), and* $Pref_{aud}$ *is a preference relation (transitive, irreflexive and asymmetric),* $Pref_{aud} \subseteq \mathcal{V} \times \mathcal{V}$*, reflecting the value preferences of audience aud.*

This framework is a particular case of the previous PAF where the preference relation $\geq$ between arguments is defined as illustrated in Definition 4.2.1, i.e. on the basis of the importance of their corresponding values. Thus, for evaluating arguments, a VAF ignores critical attacks, exactly like in PAF.

Definition 4.2.7. *An argument a defeats b for audience aud, written* $(x, y) \in \texttt{Def}_{aud}$*, if and only if both* $a\mathcal{R}b$ *and not* $Pref_{aud}(b, a)$*.*

Like in the case of preference-based argumentation frameworks, extensions w.r.t. audience *aud* are then calculated using $(\mathcal{A}, \texttt{Def}_{aud})$.

Example 4.2.5. *Let* $\mathcal{A} = \{a, b\}$*,* $\mathcal{R} = \{(a, b)\}$*,* $\mathcal{V} = \{v_1, v_2\}$*,* $\texttt{val}(a) = v_1$*,* $\texttt{val(b)} = v_2$*, and* $Pref_{aud}(v_2, v_1)$*. Then,* $\texttt{Def}_{aud} = \emptyset$*. The only stable/preferred extension w.r.t. this audience is the set* $\{a, b\}$*.*

Extended argumentation frameworks. (EAF) Modgil (2009) has proposed to reason even about preferences. Thus, arguments may support preferences about arguments.

Definition 4.2.8. *An* extended argumentation framework *is a tuple* $(\mathcal{A}, \mathcal{R}, \mathcal{D})$ *such that* $\mathcal{A}$ *is a set of arguments, and:*

- $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$
- $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{R}$
- *if* $(x, (y, z)), (x', (z, y)) \in \mathcal{D}$ *then* $(x, x'), (x', x) \in \mathcal{R}$

The idea behind this definition is that $\mathcal{R}$ is an attack relation, like in Dung's basic framework, while $\mathcal{D}$ is a second relation which ranges from arguments to attacks. For example, if $(x, y) \in \mathcal{R}$, meaning that x attacks y, then $(z, (x, y)) \in \mathcal{D}$ means that argument z neutralizes that attack by stating that y is somehow stronger/better than x and thus there is a reason to protect it from the attack of y (in $\mathcal{R}$). This means that preferences are not defined by a given preference ordering, but are themselves claimed by arguments. The third item in the previous definition specifies that if one arguments x says that z should be preferred to y and x' says that y is preferred to z, then x and x' must attack each other w.r.t. $\mathcal{R}$.

Definition 4.2.9. *Let* $(\mathcal{A}, \mathcal{R}, \mathcal{D})$ *be an extended argumentation framework and* $S \subseteq \mathcal{A}$. S *is conflict-free iff* $\forall x, y \in S$ *if* $(x, y) \in \mathcal{R}$ *then* $(y, x) \notin \mathcal{R}$ *and* $\exists z \in S$ *s.t.* $(z, (x, y)) \in \mathcal{D}$.

Definition 4.2.10. *Let* $(\mathcal{A}, \mathcal{R}, \mathcal{D})$ *be an extended argumentation framework and* $S \subseteq \mathcal{A}$. *Then* $x\texttt{Def}_S y$ *iff* $(x, y) \in \mathcal{R}$ *and* $\nexists z \in S$ *s.t.* $(z, (x, y)) \in \mathcal{D}$.

This means that attacks from $\mathcal{D}$ "neutralize" or "delete" attacks with respect to $\mathcal{R}$. Semantics are then defined using this defeat relation. For example, a set S is a stable extension if it is conflict-free and $\forall y \notin S$, $\exists x \in S$ s.t. $x\texttt{Def}_S y$.

Example 4.2.6. *Let* $\mathcal{A} = \{x, y, z\}$, $\mathcal{R} = \{(x, y)\}$ *and* $\mathcal{D} = \{(z, (x, y))\}$. *Set* $\{x, y, z\}$ *is the only stable extension of this extended argumentation framework. Informally,* z *prevents* x *in attacking* y *and thus the framework is considered as conflict-free since the only attack w.r.t.* $\mathcal{R}$ *is "ignored".*

In sum, we have seen that even if formalizations differ, the basic idea behind preference-based, value-based and extended argumentation frameworks is to ignore attacks from weaker arguments to stronger ones.

4.2.3.2 Preferences for refining

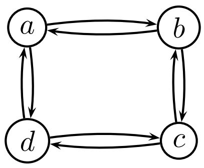
The previous section shows that most works on preferences in argumentation treat the case of critical attacks. In a recent work, Dimopoulos, Moraitis, and Amgoud (2009) have shown through a simple example that the results returned by existing approaches can be refined. The authors have focused on stable semantics, and have shown that the stable extensions returned by existing approaches can be compared, and that some of them may be better than others. They defined a new semantics which returns directly the "best" stable extensions, called super-stable extensions. For that purpose, they started by extending a preference relation $\geq$ on a set $\mathcal{A}$ of arguments as follows.

Definition 4.2.11. *Let $\mathcal{A}$ be a set of arguments and $\geq$ a preference relation on $\mathcal{A}$. A relation $\geq'$ is an extension of $\geq$ iff $\forall a, b \in \mathcal{A}$ if $a \geq b$ then $a \geq' b$ and if $a > b$ then $a >' b$. An ordering extension of $\geq$ is an extension of $\geq$ which is itself a total relation.*

A super-stable extension is then defined as follows.

Definition 4.2.12. *Let $(\mathcal{A}, \mathcal{R}, \geq)$ be a PAF such that $\mathcal{R}$ is symmetric. A set $S \subseteq \mathcal{A}$ is a* super-stable extension *of $(\mathcal{A}, \mathcal{R}, \geq)$ iff S is a stable extension of PAF $(\mathcal{A}, \mathcal{R}, \geq)$ and there exists an ordering extension $\geq'$ of $\geq$ s.t. S is a stable extension of PAF $(\mathcal{A}, \mathcal{R}, \geq')$.*

Example 4.2.7. *Let $\mathcal{A}$ and $\mathcal{R}$ be as depicted below and let $a > b, c > d$.*



The set $\{a, c\}$ is a super-stable extension, since it is a stable extension of PAF $(\mathcal{A}, \mathcal{R}, \geq')$, where $\geq'$ is an ordering extension of $\geq$ s.t. $b > c$ and $a > d$. Set $\{b, d\}$ is not a super-stable extension since there is no $\geq''$ s.t. $\geq''$ is an ordering extension of $\geq$ and $\{b, d\}$ is a stable extension of PAF $(\mathcal{A}, \mathcal{R}, \geq'')$.

Even if the idea to use preferences in order to choose between several extensions was already identified in nonmonotonic reasoning and answer set programming (Brewka et al., 2003), this paper showed for the first time that this role of preferences is also present in argumentation. In other words, preferences are used not only during the conflict-resolution phase, but also for comparing results obtained after that phase, e.g. for comparing stable extensions of $(\mathcal{A}, \mathcal{R}, \geq)$.

However, the major drawback of this framework is that it supposes a symmetric attack relation, which is shown by Amgoud and Besnard (2009) to often be undesirable. Another limitation is that the work is done only for stable semantics.

4.3 A new approach for handling critical attacks

In this section, we show the limits of existing approaches for handling critical attacks, and propose a novel solution that palliates those limits.

4.3.1 Critical examples

The three approaches (Amgoud and Cayrol, 2002b; Bench-Capon, 2003; Modgil, 2009) look for attacks from weak to stronger arguments, remove them from the attack relation, and then evaluate arguments on the basis of the remaining attacks. While this seems meaningful, we show that removing attacks may lead to *conflicting extensions* in case of non-symmetric attack relations.

Example 4.3.1. *Assume that $\mathcal{A} = \{a, b\}$ and $\mathcal{R} = \{(a, b)\}$ ($\mathcal{R}$ being not symmetric, like undercut). Assume also that b is strictly better than a. For Amgoud and Cayrol (2002b), $b > a$. In the framework of Bench-Capon (2003), the value promoted by b is more important than the value promoted by a. In the model proposed by Modgil (2009), an additional argument c is added in $\mathcal{A}$ and $\mathcal{D} = \{(c, (a, b))\}$ is used instead of $\geq$. The three approaches return only one extension, which is the set $\{a, b\}$, in case the framework by Amgoud and Cayrol (2002b) or the one by Bench-Capon (2003) is used, or the set $\{a, b, c\}$, if the framework of Modgil (2009). In all three cases, the only extension of the frameworks is not conflict-free in the sense of $\mathcal{R}$.*

The previous example illustrates a negative feature of existing preference-based argumentation frameworks, which is that an extension may contain attacks w.r.t. $\mathcal{R}$. This is in contradiction with the fact that an extension represents one of the possible points of view, each of them being coherent. The following example shows that violating conflict-freeness may lead to the violation of the rationality postulates proposed by Amgoud and Besnard (2009), see Definition 2.3.9.

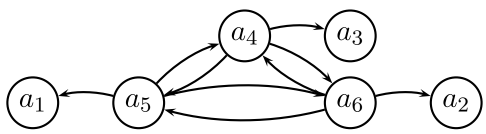
Let $\Sigma = \Sigma_1 \cup \Sigma_2$ with $\Sigma_1 = \{x\}$ and $\Sigma_2 = \{\neg y, x \to y\}$ be a stratified propositional knowledge base. (Recall that this means that the formula x is *preferred to* the two other formulae). The following framework may be constructed using Σ.

$a_1 : (\{x\}, x)$	$a_2 : (\{\neg y\}, \neg y)$
$a_3 : (\{x \to y\}, x \to y)$	$a_4 : (\{x, \neg y\}, x \wedge \neg y)$
$a_5 : (\{\neg y, x \to y\}, \neg x)$	$a_6 : (\{x, x \to y\}, y)$

Note that propositional logic verifies Tarski's axioms; thus propositional logic is a Tarskian logic. Furthermore, Amgoud and Besnard (2009) have shown that if arguments are built using a Tarskian logic and a knowledge base contains a ternary minimal conflict[2] (which is the case with Σ) then symmetric attack relations violate consistency. Thus, we should choose a non-symmetric relation like undercut (Definition 3.4.2).

Note that it has been shown by Cayrol (1995) that the corresponding argumentation framework ensures sound results. Indeed, the base of each stable extension of the framework is a maximal consistent subset of Σ.

The figure below depicts the attacks between the six above arguments.



As a preference relation, we will use *the weakest link principle* (Definition 4.2.2) as a preference relation. In our example, a_1 is strictly preferred to all the other arguments, since it is constructed only from the formulae from Σ_1. Thus, $a_1 > a_2, a_3, \ldots, a_6$.

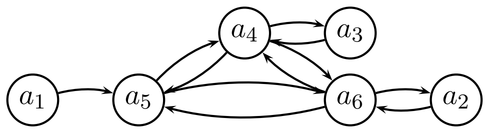
All the three existing approaches for preference-based argumentation (Amgoud and Cayrol, 2002b; Bench-Capon, 2003; Modgil, 2009) remove the attack from a_5 to a_1 and obtain the set $\mathcal{B} = \{a_1, a_2, a_3, a_5\}$ as a stable extension. Note that if we use the framework proposed by Modgil (2009), then supplementary arguments, which specify that the argument a_1 is stronger than the others, should be added. However, the framework will also return

[2]A set S is a ternary minimal conflict iff S contains exactly three formulae, S is inconsistent, and every proper subset of S is consistent.

only the extension $\mathcal{B}$. It is easy to check that this extension is not conflict-free with respect to the attack relation (undercut). Worse yet, it contains two arguments with contradictory conclusions (x and $\neg x$). It is clear that this problem is due to the addition of preferences since as said before, when preferences are ignored the framework returns sound results as shown by Cayrol (1995). What happens is that when an argument is stronger than its attacker, the attack is completely removed from the graph. By so doing, an important information is lost. This information is the conflict that exists between the two arguments, and consequently the two arguments may belong to the same extension. Note that this observation holds for any asymmetric relation and not only the one we are using in this example. Thus an approach which removes attacks is not acceptable since it does not guarantee conflict-free extensions.

One may argue that the undesirable behavior in our example is due to incompleteness of the framework, since other arguments can be constructed from Σ, e.g. $(\{x\}, x \vee y)$, $(\{x\}, x \wedge x), \ldots$ However, we will now show that even if an arbitrary set of arguments from $\texttt{Arg}(\Sigma)$ is added, the resulting framework always has a stable extension which contains conflicting arguments w.r.t. $\mathcal{R}$ and an inconsistent base. Let $\mathcal{A} \subseteq \texttt{Arg}(\Sigma)$ be an arbitrary set which contains the initial framework, i.e. s.t. $\{a_1, a_2, \ldots, a_6\} \subseteq \mathcal{A}$, let $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ be undercut, and $\geq \subseteq \mathcal{A} \times \mathcal{A}$ the weakest link principle. Let $\mathcal{B}_1 = (\texttt{Arg}(\{x\}) \cup \texttt{Arg}(\{x \to y, \neg y\})) \cap \mathcal{A}$. Let $\texttt{Def} \subseteq \mathcal{A} \times \mathcal{A}$ be the defeat relation obtained after deleting all attacks from arguments of level 2 to those having level 1. Let $a, b \in \mathcal{B}_1$, we will show that $\neg(a\texttt{Def}b)$. Let $\texttt{Supp}(b) = \{x\}$. Then, we have $a\mathcal{R}b$ only if $\texttt{Conc}(a) \equiv \neg x$. This is possible only if $\texttt{Supp}(a) = \{x \to y, \neg y\}$, and consequently $b > a$. Thus, $\neg(a\texttt{Def}b)$. Let $\texttt{Supp}(b) \neq \{x\}$. Then, for $a\mathcal{R}b$ to hold, we need to have $\texttt{Conc}(a) \equiv x \wedge \neg y$ or $\texttt{Conc}(a) \equiv y$. In the first case, we have $\texttt{Supp}(a) = \{x, \neg y\}$ and in the second case, $\texttt{Supp}(a) = \{x, x \to y\}$. In both cases, $a \notin \mathcal{B}_1$. Thus, $\mathcal{B}_1$ is conflict-free w.r.t. $\texttt{Def}$. Let us now show that $\mathcal{B}_1$ attacks any argument in $\mathcal{A} \setminus \mathcal{B}_1$ w.r.t. $\texttt{Def}$. Let $b \in \mathcal{A} \setminus \mathcal{B}_1$. Then $\texttt{Supp}(b) = \{x, x \to y\}$ or $\texttt{Supp}(b) = \{x, \neg y\}$. In both cases, $a_5\texttt{Def}b$. Thus, $\mathcal{B}_1$ is a stable extension of $(\mathcal{A}, \texttt{Def})$. This means that even if an arbitrary number of arguments is added, the framework always has an extension containing conflicting arguments and inconsistent base.

Another remark may be that there are other attack relations that could be used. For example, it may seem that the problem could be "solved" by using a relation for which a_1 attacks a_5 and a_5 attacks a_1. Can we define $\mathcal{R}$ as a union of rebut[3] and undercut? Let $\mathcal{R}$ be the union of undercut and rebut, i.e. $a\mathcal{R}b$ iff a undercuts b or a rebuts b. In this case, the attack graph w.r.t. `Def` would be as follows:



However, it is easy to see that this attack relation does not solve the problem: this framework returns $\{a_1, a_2, a_3\}$ as a stable extension, a set having an inconsistent base.

Furthermore, a good preference-based argumentation framework should return sound results for *any* input; for any set of arguments and any attack relation, extensions should be conflict-free.

4.3.2 A new approach

The previous subsection highlighted the limits of existing preference-based argumentation frameworks. Even if the idea pursued by these frameworks is meaningful, their results may violate the key property of conflict-freeness with respect to the attack relation $\mathcal{R}$. This problem is mainly due to the removing attacks from weaker to stronger arguments from the framework (critical attacks).

We propose a new approach for modeling the conflict-resolution role of preferences in argumentation which prevents the above problem. Instead of changing the original attack relation, we take into account preferences when evaluating the arguments, i.e. at the semantics level. Our aim is not to define new acceptability semantics but to generalize the existing ones with preferences. Hence, when there are no critical attacks, the extended semantics should return the same results as the basic ones (without preferences).

Our approach presents another novelty which consists of defining a semantics as a *dominance relation* on the *power set* of the set $\mathcal{A}$ of arguments. The best elements w.r.t. this relation are the acceptable sets of arguments, i.e. the extensions. Recall that existing semantics divide the power set of $\mathcal{A}$ into two subsets: extensions and non-extensions. The former are better than

[3]Recall that rebut is defined as $a\mathcal{R}b$ iff $\texttt{Conc}(a) \equiv \neg\texttt{Conc}(b)$.

the latter, but they do not say anything about non-extensions. However, in some applications, one may want to compare some sets of arguments. For instance, after a dialogue between two argents, an observer may want to compare the two sets of arguments exchanged by the two agents. Defining a semantics as a relation allows the comparison of any pair of subsets of arguments (on the basis of attacks and preferences).

Before defining formally the new semantics, let us first introduce some notations and concepts. We define a preference-based argumentation framework (PAF) as follows.

Definition 4.3.1 (PAF)**.** *A* PAF *is a tuple* $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ *where* $\mathcal{A}$ *is a set of arguments,* $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ *an attack relation, and* $\geq \subseteq \mathcal{A} \times \mathcal{A}$ *a (partial or total) preorder.*

We suppose that $\mathcal{R}$ does not contain self-attacking arguments (i.e. $\mathcal{R}$ is irreflexive). Note that whenever arguments are built from a logical knowledge base, this assumption is verified. Furthermore, all definitions and results can be presented with slight modifications even for the case when $\mathcal{R}$ is an arbitrary relation. However, we do not study this case in order to simplify notations and proofs.

Notation: Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. $\mathcal{CF}(\mathcal{T})$ denotes the conflict-free (w.r.t. $\mathcal{R}$) sets of arguments. At some places, we abuse notation and use $\mathcal{CF}(\mathcal{F})$ to denote the conflict-free sets of arguments of a basic framework $\mathcal{F} = (\mathcal{A}, \mathcal{R})$.

As already explained, a semantics for evaluating arguments of a PAF is defined as a binary relation on the power set $\mathcal{P}(\mathcal{A})$ of $\mathcal{A}$. Such a relation will be denoted by $\succeq$. For $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$, writing $(\mathcal{E}, \mathcal{E}') \in \succeq$ (or equivalently $\mathcal{E} \succeq \mathcal{E}'$) means that the set $\mathcal{E}$ is at least as good as the set $\mathcal{E}'$. The relation $\succ$ is the strict version of $\succeq$, that is for $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$, $\mathcal{E} \succ \mathcal{E}'$ iff $\mathcal{E} \succeq \mathcal{E}'$ and not $(\mathcal{E}' \succeq \mathcal{E})$. The maximal elements of such a relation are defined as follows.

Definition 4.3.2 (Maximal elements)**.** *Let* $\mathcal{A}$ *be a set of arguments,* $\mathcal{E} \in \mathcal{P}(\mathcal{A})$ *and* $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$*.* $\mathcal{E}$ *is* maximal *w.r.t.* $\succeq$ *iff:*

1. $\forall \mathcal{E}' \in \mathcal{P}(\mathcal{A})$, $\mathcal{E} \succeq \mathcal{E}'$,

2. *No strict superset of* $\mathcal{E}$ *verifies (1).*

Let $\succeq_{max}$ *denote the set of maximal sets w.r.t.* $\succeq$*.*

Like existing acceptability semantics, preference-based semantics should satisfy some basic requirements. Thus, not any relation $\succeq$ can be used for evaluating arguments in a PAF. An appropriate relation should satisfy at least three postulates.

Notation: The writing $\frac{X_1 \ldots X_n}{Y}$ means that if X_1 ... and X_n hold, then Y holds as well.

The first postulate states that any conflict-free set of arguments should be strictly preferred to a conflicting one.

Postulate 1 ($P1$). *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$.*

$$\frac{\mathcal{E} \in \mathcal{CF}(\mathcal{T}) \qquad \mathcal{E}' \notin \mathcal{CF}(\mathcal{T})}{\mathcal{E} \succ \mathcal{E}'}$$

Postulate $P1$ ensures conflict-freeness for the extensions of any PAF. Indeed, the best elements of any dominance relation satisfying this postulate are conflict-free.

Proposition 4.3.1. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. If a relation $\succeq$ satisfies postulate $P1$, then each element of the set $\succeq_{max}$ is conflict-free w.r.t. $\mathcal{R}$.

The second postulate describes the role of the attack relation. It shows that an attack should win when it is not critical. This is in some sense the basic idea behind all existing semantics in the literature.

Postulate 2 ($P2$). *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $a, a' \in \mathcal{A}$.*

$$\frac{a\mathcal{R}a' \qquad \neg(a'\mathcal{R}a) \qquad \neg(a' > a)}{\{a\} \succ \{a'\}}$$

The third postulate ensures that preferences are privileged in critical attacks. This is in fact the idea defended in previous works on PAFs (e.g. Amgoud and Cayrol, 2002b; Bench-Capon, 2003). Indeed, if an argument a attacks another argument a' and $a' > a$, then the set $\{a'\}$ is privileged. Thus, $\{a'\}$ should be strictly preferred to $\{a\}$.

Postulate 3 ($P3$). *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $a, a' \in \mathcal{A}$.*

$$\frac{a\mathcal{R}a' \qquad a' > a}{\{a'\} \succ \{a\}}$$

We are now ready to define semantics for evaluating the arguments of a PAF. A semantics is a binary relation (called also dominance relation) on the power set of the set of arguments and which satisfies the above postulates. The acceptable sets of arguments are the best elements of the dominance relation.

Definition 4.3.3 (Semantics for PAFs)**.** *An* acceptability semantics *for a PAF $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ is defined by a dominance relation $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$ which satisfies postulates $P1$, $P2$ and $P3$. Extensions of $\mathcal{T}$ under semantics $\succeq$ are the elements of $\succeq_{max}$.*

4.3.3 Generalising Dung's semantics with preferences

In this subsection, we propose three new semantics which generalize respectively stable, preferred and grounded semantics. Before presenting them, let us first define formally when a semantics generalizes another one.

Definition 4.3.4 (Generalising a semantics)**.** *A dominance relation $\succeq$* generalises semantics *x iff for all $(\mathcal{A}, \mathcal{R}, \geq)$, if $\nexists a, b \in \mathcal{A}$ such that $a\mathcal{R}b$ and $b > a$, then $\succeq_{max} = \texttt{Ext}((\mathcal{A}, \mathcal{R}))$ where $\texttt{Ext}((\mathcal{A}, \mathcal{R}))$ is the set of all extensions of the argumentation framework $(\mathcal{A}, \mathcal{R})$ w.r.t. semantics x.*

Informally speaking, a dominance relation generalises a given semantics iff its best elements are exactly the extensions of the basic framework (i.e. without preferences) w.r.t. that semantics, unless there are critical attacks.

4.3.3.1 Generalising stable semantics

Before showing how to extend stable semantics with preferences, we show that it is possible to encode this semantics in the new setting, i.e. to define it as a dominance relation on the power set of the set of arguments. The following theorem characterizes the dominance relations that encode stable semantics.

Theorem 4.3.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$. Let $\texttt{Ext}(\mathcal{F})$ be the set containing all the stable extensions of $\mathcal{F}$. The equality $\texttt{Ext}(\mathcal{F}) = \succeq_{max}$ holds iff $\forall \mathcal{E} \in \mathcal{P}(\mathcal{A})$,

1. if $\mathcal{E} \notin \mathcal{CF}(\mathcal{F})$ then $\exists \mathcal{E}' \in \mathcal{P}(\mathcal{A})$ s.t. not$(\mathcal{E} \succeq \mathcal{E}')$, and
2. if $\mathcal{E} \in \mathcal{CF}(\mathcal{F})$ and $\forall a' \notin \mathcal{E}$, $\exists a \in \mathcal{E}$ s.t. $a\mathcal{R}a'$, then $\forall \mathcal{E}' \in \mathcal{P}(\mathcal{A})$, $\mathcal{E} \succeq \mathcal{E}'$, and

3. if $\mathcal{E} \in \mathcal{CF}(\mathcal{F})$ and $\exists a' \in \mathcal{A} \setminus \mathcal{E}$ s.t. $\nexists a \in \mathcal{E}$ s.t. $a\mathcal{R}a'$, then $\exists \mathcal{E}' \in \mathcal{P}(\mathcal{A})$ s.t. not$(\mathcal{E} \succeq \mathcal{E}')$.

It is worth mentioning that there are several relations $\succeq$ that encode stable semantics. All these relations return the same maximal elements (i.e. the stable extensions). However, they compare in different ways the remaining sets of arguments. An example of a relation that encodes stable semantics is the following:

Relation 1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. $\mathcal{E} \succeq_1 \mathcal{E}'$ iff

- $\mathcal{E} \in \mathcal{CF}(\mathcal{F})$ and $\mathcal{E}' \notin \mathcal{CF}(\mathcal{F})$, or
- $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{F})$ and $\forall a' \in \mathcal{E}' \setminus \mathcal{E}$, $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $a\mathcal{R}a'$.

Let us illustrate this relation on the following simple example.

Example 4.3.2. *Consider the argumentation framework depicted in the figure below.*

It can be checked that: $\{a\} \succeq_1 \emptyset$, $\{b\} \succeq_1 \emptyset$, $\emptyset \succeq_1 \{a, b\}$. The two sets $\{a\}$ and $\{b\}$ are equally preferred. The maximal elements of $\succeq_1$ (its stable extensions) are $\{a\}$ and $\{b\}$.

Note that Dung's approach returns only two classes of subsets of arguments: the extensions and the non-extensions. In Example 4.3.2, the two sets $\{a\}$ and $\{b\}$ are stable extensions while it does not say anything about the sets $\{a, b\}$ and $\{\}$. Our approach compares even the non-extensions. According to relation $\succeq_1$, the set $\{\}$ is preferred to the set $\{a, b\}$.

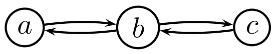
In what follows, we present a new semantics, called *pref-stable* semantics, that generalises stable semantics with preferences. This amounts to define a dominance relation which will be denoted by $\succeq_s$ and its best elements by $\succeq^s_{max}$. The idea behind this relation is the following: given two conflict-free sets of arguments, $\mathcal{E}$ and $\mathcal{E}'$, we say that $\mathcal{E}$ is better than $\mathcal{E}'$ iff any argument in $\mathcal{E}' \setminus \mathcal{E}$ is weaker than at least one argument in $\mathcal{E} \setminus \mathcal{E}'$ or is attacked by it. Moreover, a conflict-free set of arguments is strictly preferred to a conflicting one, while conflicting sets are all incomparable. In fact, the relation $\succeq_s$ extends the relation $\succeq_1$ with preferences.

Definition 4.3.5 (Pref-stable semantics)**.** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. It holds that $\mathcal{E} \succeq_s \mathcal{E}'$ iff:*

- *$\mathcal{E} \in \mathcal{CF}(\mathcal{T})$ and $\mathcal{E}' \notin \mathcal{CF}(\mathcal{T})$, or*
- *$\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{T})$ and $\forall a' \in \mathcal{E}' \setminus \mathcal{E}$, $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. ($a\mathcal{R}a'$ and not($a' > a$)) or ($a > a'$).*

Let us illustrate this definition through the following simple example.

Example 4.3.3. *Let $\mathcal{A} = \{a, b, c\}$, $a > b$ and let $\mathcal{R}$ be as depicted in the figure below:*



The conflict-free sets of arguments are: $\mathcal{E}_1 = \emptyset$, $\mathcal{E}_2 = \{a\}$, $\mathcal{E}_3 = \{b\}$, $\mathcal{E}_4 = \{c\}$, and $\mathcal{E}_5 = \{a, c\}$. It can be checked that the following relations hold: $\mathcal{E}_2 \succeq_s \mathcal{E}_1$, $\mathcal{E}_3 \succeq_s \mathcal{E}_1$, $\mathcal{E}_4 \succeq_s \mathcal{E}_1$, $\mathcal{E}_5 \succeq_s \mathcal{E}_1$, $\mathcal{E}_5 \succeq_s \mathcal{E}_4$, $\mathcal{E}_5 \succeq_s \mathcal{E}_2$, $\mathcal{E}_5 \succeq_s \mathcal{E}_3$, $\mathcal{E}_4 \succeq_s \mathcal{E}_3$, $\mathcal{E}_3 \succeq_s \mathcal{E}_4$, $\mathcal{E}_2 \succeq_s \mathcal{E}_3$. It can also be checked that $\succeq_{max} = \{\mathcal{E}_5\}$.

The relation $\succeq_s$ is in conformity with Definition 4.3.3. Indeed, it satisfies the three postulates $P1, P2$ and $P3$.

Proposition 4.3.2. The relation $\succeq_s$ satisfies postulates $P1, P2$ and $P3$.

Since the relation $\succeq_s$ satisfies postulate $P1$, its extensions are conflict-free. The following result shows that they are even maximal (for set inclusion). Indeed, the relation $\succeq_s$ privileges maximal sets.

Proposition 4.3.3. Let $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. If $\mathcal{E} \subsetneq \mathcal{E}'$ then $\mathcal{E}' \succ_s \mathcal{E}$ (i.e. $\mathcal{E}' \succeq_s \mathcal{E}$ and not ($\mathcal{E} \succeq_s \mathcal{E}'$)).

However, not any maximal conflict-free set of arguments is an extension (i.e. an element of $\succeq^s_{max}$) as shown by the following example.

Example 4.3.4. *The set $\mathcal{E}_3$ from Example 4.3.3 is a maximal conflict-free set but does not belong to $\succeq^s_{max}$.*

From Proposition 4.3.3, it follows that Definition 4.3.2 can be simplified as follows: $\mathcal{E} \in \succeq^s_{max}$ iff $\forall \mathcal{E}' \in \mathcal{P}(\mathcal{A}), \mathcal{E} \succeq_s \mathcal{E}'$. Finally, notice that the relation $\succeq_s$ is *not transitive*. Indeed, in the previous example, $\mathcal{E}_2 \succeq_s \mathcal{E}_3$ and $\mathcal{E}_3 \succeq_s \mathcal{E}_4$ however, the two sets $\mathcal{E}_2$ and $\mathcal{E}_4$ are incomparable w.r.t. $\succeq_s$. Informally speaking, this is a consequence of two facts: first, $\succeq_s$ takes into account attacks from $\mathcal{R}$; second, $\mathcal{R}$ is not transitive. For example, if $\mathcal{A} = \{a, b, c\}$, $\mathcal{R} = \{(a, b), (b, c)\}$ and all the arguments are equally preferred, then $\{a\} \succ_s \{b\}$ and $\{b\} \succ_s \{c\}$, as expected. However, there is not obvious reason to prefer $\{a\}$ to $\{c\}$. (Note that those two sets are not comparable w.r.t. $\succeq_s$.

The following theorem shows that pref-stable semantics generalises stable semantics. Recall that this means that the two semantics coincide in case any attacked argument is not stronger than its attacker.

Theorem 4.3.2. The relation $\succeq_s$ generalises stable semantics.

Finally, we can show that the proposed approach handles correctly the example discussed on pages 58–60. Namely, it can be checked that the corresponding PAF has exactly two extensions: $\{a_1, a_2, a_4\}$ (whose base is $\{x, \neg y\}$) and $\{a_1, a_3, a_6\}$ (whose base is $\{x, x \rightarrow y\}$), and that both of them are conflict-free and support consistent conclusions.

4.3.3.2 Generalising preferred semantics

We now propose a new semantics, called *pref-preferred*, that generalises preferred semantics with preferences. It is defined by a dominance relation, denoted by $\succeq_p$. The basic idea behind this relation is that a set $\mathcal{E}$ of arguments is better than another set $\mathcal{E}'$ of arguments iff for every attack from $\mathcal{E}'$ to $\mathcal{E}$ which does not fail, $\mathcal{E}$ is capable to defend the attacked argument and that for every attack from $\mathcal{E}$ to $\mathcal{E}'$ which fails, there is another attack from $\mathcal{E}$ that defends the argument which failed in its attack.

Definition 4.3.6 (Pref-preferred semantics)**.** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. $\mathcal{E} \succeq_p \mathcal{E}'$ iff:*

- *$\mathcal{E} \in \mathcal{CF}(\mathcal{T})$ and $\mathcal{E}' \notin \mathcal{CF}(\mathcal{T})$, or*
- *$\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{T})$ and $\forall a \in \mathcal{E}$, $\forall a' \in \mathcal{E}'$, if ($a'\mathcal{R}a$ and not($a > a'$)) or ($a\mathcal{R}a'$ and $a' > a$), then $\exists b \in \mathcal{E}$ such that ($b\mathcal{R}a'$ and not($a' > b$)) or ($a'\mathcal{R}b$ and $b > a'$).*

From now on, $\succeq^p_{max}$ will denote the best elements w.r.t. this relation. Let us illustrate this definition through the next example.

Example 4.3.5. *In Example 4.3.3, it holds that $\mathcal{E}_2 \succ_p \mathcal{E}_3$, $\mathcal{E}_3 \succeq_p \mathcal{E}_4$, $\mathcal{E}_4 \succeq_p \mathcal{E}_3$, $\mathcal{E}_5 \succ_p \mathcal{E}_3, \ldots$. It can also be checked that $\succeq^p_{max} = \{\mathcal{E}_5\}$ for this framework.*

Note that the relation $\succeq_p$ is not transitive. However, it can be checked that it satisfies the three postulates $P1, P2$ and $P3$. Thus, it encodes a semantics in the sense of Definition 4.3.3.

Proposition 4.3.4. The relation $\succeq_p$ satisfies postulates $P1, P2$ and $P3$.

The above proposition ensures that the extensions of a PAF under pref-preferred semantics are conflict-free. The following result shows that this semantics generalises Dung's preferred semantics.

Theorem 4.3.3. The relation $\succeq_p$ generalises preferred semantics.

In Dung's basic framework, every stable extension is a preferred one. We show that the same link holds in our setting. Namely, every pref-stable extension is a pref-preferred extension.

Theorem 4.3.4. For any $(\mathcal{A}, \mathcal{R}, \geq)$, it holds that $\succeq^s_{max} \subseteq \succeq^p_{max}$.

4.3.3.3 Generalising grounded semantics

We now focus on grounded semantics and generalise it with preferences. The new semantics is called *pref-grounded* and is defined by a dominance relation which is denoted by $\succeq_g$. The basic idea behind this relation is that a set is not worse than another if it can strongly defend all its arguments against all attacks that come from the other set.

Before giving the formal definition of $\succeq_g$, let us first generalise the notion of strong defense by preferences. The idea is that an argument has either to be preferred to its attacker or has to be defended by arguments that themselves can be strongly defended without using the argument in question. Note that, for simplicity reasons, in this sub-subsection we suppose that the set of arguments $\mathcal{A}$ is finite. While it is certainly possible to define generalisations of grounded semantics for infinite sets of arguments (which we will do later in this chapter), we conducted this first study of generalising grounded semantics by dominance relations for a finite case. Consequently, we suppose a finite set of arguments in the results concerning grounded semantics in Subsection 4.3.5 (namely Theorem 4.3.12 and Theorem 4.3.13).

Definition 4.3.7 (Strong defense)**.** *Let $\mathcal{E} \subseteq \mathcal{A}$. $\mathcal{E}$* strongly defends *an argument a from attacks of set $\mathcal{E}'$, denoted by $sd(a, \mathcal{E}, \mathcal{E}')$, iff $\forall b \in \mathcal{E}'$ if $(b\mathcal{R}a$ and not$(a > b))$ or $(a\mathcal{R}b$ and $b > a)$, then $\exists c \in \mathcal{E} \setminus \{a\}$ such that $((c\mathcal{R}b$ and not$(b > c))$ or $(b\mathcal{R}c$ and $c > b))$ and $sd(c, \mathcal{E} \setminus \{a\}, \mathcal{E}')$.*
If the third argument of sd is not specified, then we define $sd(a, \mathcal{E})$ as $sd(a, \mathcal{E}, \mathcal{A})$.

Let us illustrate this notion through the following example.

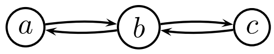
Example 4.3.6. *In the framework of Example 4.3.3, we have* $sd(a, \{a\}, \{b\})$ *since* a *is strictly preferred to* b *thus it can defend itself. However, we have* $\neg sd(b, \{b\}, \{c\})$ *since* b *cannot strongly defend itself against* c*. On the other hand,* $sd(c, \{a, c\}, \{b\})$ *holds since* a *can defend* c *against* b *and* a *is protected from* b *since it is strictly preferred to it.*

The relation $\succeq_g$ prefers subsets that strongly defend all their arguments. Namely, $\mathcal{E} \succeq_g \mathcal{E}'$ iff $\mathcal{E}$ strongly defends all its arguments against attacks from $\mathcal{E}'$.

Definition 4.3.8 (Pref-grounded semantics)**.** *Let* $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ *be a PAF and* $\mathcal{E}, \mathcal{E}'$ *be two subsets of* $\mathcal{A}$*. It holds that* $\mathcal{E} \succeq_g \mathcal{E}'$ *iff:*

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$ *and* $\mathcal{E}' \notin \mathcal{CF}(\mathcal{T})$*, or*
- $\forall a \in \mathcal{E}$*, it holds that* $sd(a, \mathcal{E}, \mathcal{E}')$*.*

Example 4.3.7. *Let* $\mathcal{A} = \{a, b, c\}$*,* $b > a$*, and* $\mathcal{R}$ *is as depicted in the figure below:*



One can check that there is exactly one subset of $\mathcal{A}$ *which is preferred to all other subsets of arguments w.r.t.* $\succeq_g$*. This set is the empty one. While we do have* $\{b\} \succeq_g \{a\}$*, we have* $\neg(\{b\} \succeq_g \{c\})$*, so* $\{b\}$ *is not an extension of this PAF. We have also* $\neg(\{a\} \succeq_g \{b\})$*,* $\neg(\{c\} \succeq_g \{b\})$ *and* $\neg(\{a, c\} \succeq_g \{b\})$*. This is expected and natural output since neither* b *nor* c *are capable to defend strongly themselves and, on the other hand, it can be said that* a *is the worst argument in this framework, thus not strong enough to be better than* b*.*

The relation $\succeq_g$ has exactly one best element, i.e. the set $\succeq^g_{max}$ contains only one set of arguments. This is not surprising since pref-grounded semantics intends to generalise the principle underlying Dung's grounded semantics.

Proposition 4.3.5. The equality $| \succeq^g_{max} | = 1$ holds.

The following result shows that the relation $\succeq_g$ satisfies the three postulates $P1, P2$ and $P3$. Thus, its unique extension is conflict-free.

Proposition 4.3.6. The relation $\succeq_g$ satisfies postulates $P1, P2$ and $P3$.

Finally, the dominance relation $\succeq_g$ generalises grounded semantics.

Theorem 4.3.5. The relation $\succeq_g$ generalises grounded semantics.

In Dung's basic framework, the grounded extension is a subset of the intersection of all preferred extensions. The same link exists between pref-grounded and pref-preferred extensions:

Theorem 4.3.6. For any $(\mathcal{A}, \mathcal{R}, \geq)$, if $\mathcal{E} \in \succeq^g_{max}$ then $\mathcal{E} \subseteq \bigcap_{\mathcal{E}_i \in \succeq^p_{max}} \mathcal{E}_i$.

4.3.4 Characterizing pref-stable semantics

In the previous subsection, we have proposed three particular semantics which generalise respectively stable, preferred and grounded semantics with preferences. What is worth mentioning is that the three corresponding dominance relations are *not unique*. There exist, for instance, other relations which may generalise stable semantics by preferences. Not surprisingly, the same is true for preferred and grounded semantics. This remark opens many new questions: How many dominance relations that generalise a given semantics do exist? Are some of them "better" than others? What are their properties? What are the differences between them? In the rest of the section we focus on stable semantics and give a formal and precise answer to these questions.

4.3.4.1 Postulates

In this subsection, we characterize all the dominance relations $\succeq$ that generalise stable semantics with preferences. For that purpose, we identify a set of postulates that such relations should satisfy. It is clear that the three postulates $P1, P2$ and $P3$ are in that set. Postulate $P1$ ensures that the extensions of a PAF are conflict-free w.r.t. the attack relation. This is important since an extension represents a coherent point of view. Postulates P2 and P3 describe when the attack relation should take precedence over the preference relation and when this latter is privileged. These two postulates are given in order to underline the basic ideas on how to combine attacks and preferences. However, they specify the correct behavior only in case of singletons (sets containing exactly one argument). To completely define a dominance relation, we will need to supply additional postulates which will describe its desirable properties. We provide $P4$ and $P5$ with a motivation to generalise basic principles behind Dung's stable semantics.

The first postulate describes when a set should not be preferred to another. The idea is that: if an argument of a set $\mathcal{E}$ cannot be compared with arguments in another set $\mathcal{E}'$ (since it is neither attacked nor less preferred to any argument of that set), then $\mathcal{E}$ cannot be less preferred to $\mathcal{E}'$.

Postulate 4 (P4)**.** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF, and $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{T})$ such that $\mathcal{E} \cap \mathcal{E}' = \emptyset$.*

$$\frac{(\exists a' \in \mathcal{E}')(\forall a \in \mathcal{E})\ \neg(a\mathcal{R}a' \wedge \neg(a' > a)) \ \wedge\ \neg(a > a')}{\neg(\mathcal{E} \succeq \mathcal{E}')}$$

The second postulate describes when a set is preferred to another. The idea is that if for any argument of a set, there is at least one argument in another set which 'wins the conflict' with it, then the latter should be preferred to the former. There are two situations in which an argument a wins a conflict against a': either a attacks a' and a' does not defend itself since it is not stronger than a w.r.t. $\geq$, or a' attacks a but a is strictly preferred to a'.

Postulate 5 (P5)**.** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{T})$ such that $\mathcal{E} \cap \mathcal{E}' = \emptyset$.*

$$\frac{(\forall a' \in \mathcal{E}')(\exists a \in \mathcal{E})\ s.t.\ (a\mathcal{R}a' \wedge \neg(a' > a))\ or\ (a'\mathcal{R}a \wedge a > a')}{\mathcal{E} \succeq \mathcal{E}'}$$

Proposition 4.3.7. Let $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$. If $\succeq$ satisfies postulates $P4$ and $P5$, then it also satisfies postulates $P2$ and $P3$.

The following requirement ensures that a dominance relation is entirely based on the distinct elements of any two subsets of arguments.

Postulate 6 (P6)**.** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF, and $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{T})$. Then:*

$$\frac{\mathcal{E} \succeq \mathcal{E}'}{\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}} \qquad\qquad \frac{\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}}{\mathcal{E} \succeq \mathcal{E}'}$$

Note that in Definition 4.3.5 we have defined a particular relation $\succeq_s$ that we called pref-stable semantics. From now on, we will **redefine** this notion, by letting any relation satisfying P1, P4, P5 and P6 be called *pref-stable semantics.*

Definition 4.3.9 (Pref-stable semantics)**.** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. A relation $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$ encodes pref-stable semantics iff it satisfies postulates $P1, P4, P5$ and $P6$.*

From now on, a relation that encodes pref-stable semantics will be called *pref-stable relation*, and its maximal elements will be called *pref-stable extensions*.

It can be checked that the relation $\succeq_s$ given in Definition 4.3.5 is a pref-stable relation and satisfies the four postulates.

Proposition 4.3.8. $\succeq_s$ is a pref-stable relation.

There are several relations that encode pref-stable semantics. However, they all return the same pref-stable extensions.

Theorem 4.3.7. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\succeq, \succeq' \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$. If $\succeq$ and $\succeq'$ are pref-stable relations, then $\succeq_{max} = \succeq'_{max}$.

Note that postulates $P1$, $P4$, $P5$ and $P6$ encode important properties of stable semantics enriched with preferences. However, it is worth noticing that no relation which generalises stable semantics and verifies $P1$ and $P5$ is transitive. As already mentioned on page 65, that this is not surprising since $P5$ describes one of the basic properties of stable semantics, which is that a set attacking another one should win. This notion is not necessarily transitive since it is based on an attack relation which does not exhibit any property. Indeed, an attack relation is generally not a preorder. We formally show that transitivity is incompatible with postulates $P1$ and $P5$.

Proposition 4.3.9. There exists no transitive relation which generalises stable semantics and satisfies postulates $P1$ and $P5$.

Finally, we can show that a pref-stable semantics generalises stable semantics.

Theorem 4.3.8. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. Any pref-stable relation $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$ generalises stable semantics.

4.3.4.2 General and specific pref-stable relations

As already said, there are several relations that encode pref-stable semantics. Our aim now is to define the upper and lower bounds of these relations. The most general pref-stable relation, denoted by $\succeq_{gn}$, returns $\mathcal{E} \succeq_{gn} \mathcal{E}'$ if and only if it can be proved from the four postulates that $\mathcal{E}$ must be preferred to $\mathcal{E}'$.

Definition 4.3.10 (General pref-stable relation)**.** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. $\mathcal{E} \succeq_{gn} \mathcal{E}'$ iff:*

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$ *and* $\mathcal{E}' \notin \mathcal{CF}(\mathcal{T})$*, or*
- $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{T})$ *and* $\forall a' \in \mathcal{E}' \setminus \mathcal{E}, \exists a \in \mathcal{E} \setminus \mathcal{E}'$ *such that* $(a\mathcal{R}a'$ *and* *not*$(a' > a))$ *or* $(a'\mathcal{R}a$ *and* $a > a')$.

Proposition 4.3.10. $\succeq_{gn}$ is a pref-stable relation.

The most specific pref-stable relation, denoted by $\succeq_{sp}$, returns $\mathcal{E} \succeq_{sp} \mathcal{E}'$ if and only if from the four postulates, it cannot be proved that $\neg(\mathcal{E} \succeq_{sp} \mathcal{E}')$.

Definition 4.3.11 (Specific pref-stable relation)**.** *Let* $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ *be a PAF and* $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$*.* $\mathcal{E} \succeq_{sp} \mathcal{E}'$ *iff:*

- $\mathcal{E}' \notin \mathcal{CF}(\mathcal{T})$*, or*
- $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{T})$ *and* $\forall a' \in \mathcal{E}' \setminus \mathcal{E}, \exists a \in \mathcal{E} \setminus \mathcal{E}'$ *such that* $(a\mathcal{R}a'$ *and* *not*$(a' > a))$ *or* $(a > a')$.

Proposition 4.3.11. $\succeq_{sp}$ is a pref-stable relation.

Let us illustrate the differences between the three particular relations $\succeq_s$, $\succeq_{sp}$ and $\succeq_{gn}$ on the following example.

Example 4.3.8. *Let* $\mathcal{A} = \{a, b, c\}, \mathcal{R} = \{(a, b)\}$ *and* $\geq = \{(a, a), (b, b), (c, c), (a, c)\}$*. For example, it holds that* $\{a\} \succeq_s \{c\}$*,* $\{a\} \succeq_{sp} \{c\}$ *and* $\neg(\{a\} \succeq_{gn} \{c\})$*. That is, for relations* $\succeq_s$ *and* $\succeq_{sp}$ *the strict preference between a and c is enough to prefer* $\{a\}$ *to* $\{c\}$*. For relation* $\succeq_{gn}$*, since c is not attacked by a, there is no preference between sets* $\{a\}$ *and* $\{c\}$*. The fact that a is stronger is not important, because there is no conflict between those arguments.*

Another difference is that for relation $\succeq_{sp}$*, all conflicting sets are equally preferred. For example,* $\{a, b, c\} \succeq_{sp} \{a, b\}$ *and* $\{a, b\} \succeq_{sp} \{a, b, c\}$*. Relations* $\succeq_s$ *and* $\succeq_{gn}$ *encode the idea that a contradictory point of view cannot be accepted as a standpoint. Thus, it is not even possible to compare two contradictory sets of arguments. For example* $\neg(\{a, b, c\} \succeq_s \{a, b\})$*.*

The next result shows that any pref-stable relation is "between" the general and the specific relations.

Theorem 4.3.9. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. Let $\succeq$ be a pref-stable relation.

- If $\mathcal{E} \succeq_{gn} \mathcal{E}'$ then $\mathcal{E} \succeq \mathcal{E}'$.
- If $\mathcal{E} \succeq \mathcal{E}'$ then $\mathcal{E} \succeq_{sp} \mathcal{E}'$.

A simple consequence of the previous result is that, if $\mathcal{E} \succeq_{gn} \mathcal{E}'$ and $\mathcal{E} \succeq_{sp} \mathcal{E}'$, then for any pref-stable relation $\succeq$, it holds that $\mathcal{E} \succeq \mathcal{E}'$.

4.3.5 Characterizing pref-stable, pref-preferred and pref-grounded extensions

As already said, the new approach for taking into account the strengths of arguments in an argumentation framework is sound and rich. It is sound since it guarantees conflict-free extensions, and it is rich since it provides more information than existing approaches. Indeed, not only it computes the acceptable sets of arguments, but it also compares the remaining ones. This comparison is of great importance in some applications like decision making and dialogues. However, it is less crucial in some other applications like handling inconsistency in knowledge bases. In this case, one looks only for the sets of arguments which support ‘good’ conclusions and does not bother about the other arguments. It is thus important to be able to characterize the extensions under a given semantics without comparing all the subsets of arguments, i.e. without referring to pref-stable relations. This subsection provides those characterizations.

Theorem 4.3.10. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\succeq$ be a pref-stable relation.
$\mathcal{E} \in \succeq_{max}$ iff:

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$, and
- $\forall a' \in \mathcal{A} \setminus \mathcal{E}$, $\exists a \in \mathcal{E}$ such that ($a\mathcal{R}a'$ and not($a' > a$)) or ($a'\mathcal{R}a$ and $a > a'$).

This shows the link between our approach (based on dominance relations) and the existing approaches (based on changing $\mathcal{R}$ into $\texttt{Def}$ and then applying Dung's semantics on $(\mathcal{A}, \texttt{Def})$). Namely, another way to compute the pref-stable extensions of a PAF is to “invert” the direction of attacks when they are not in accordance with the preferences between arguments. We apply then stable semantics on the basic framework that is obtained. More precisely, we start with a PAF $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$. We compute an AF $\mathcal{F} = (\mathcal{A}, \mathcal{R}')$ where $\mathcal{R}'$ is defined as follows:

$$\begin{array}{rcl} \mathcal{R}' & = & \{(a,b) \in \mathcal{A} \times \mathcal{A} \mid (a\mathcal{R}b \text{ and not } (b > a))\} \\ & \cup & \{(a,b) \in \mathcal{A} \times \mathcal{A} \mid (b\mathcal{R}a \text{ and } a > b)\}. \end{array}$$

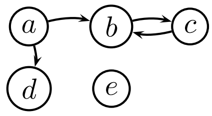
Then, we apply stable semantics on the new framework $(\mathcal{A}, \mathcal{R}')$.

The following result is a consequence of Theorem 4.3.10.

Corollary 4.3.1. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\succeq$ be a pref-stable relation. Let $\mathcal{R}' = \{(a, b) \mid a, b \in \mathcal{A},\ (a\mathcal{R}b \text{ and not}(b > a)) \text{ or } (b\mathcal{R}a \text{ and } a > b)\}$. It holds that $\succeq_{max}$ is exactly the set of stable extensions of framework $(\mathcal{A}, \mathcal{R}')$.

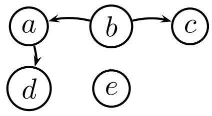
Let us illustrate this result through an example.

Example 4.3.9. *Let $\mathcal{A} = \{a, b, c, d, e\}$ and let $\mathcal{R}$ be as depicted in figure below:*



Assume that $b > a$, $b > c$ and $e > d$. Note that this framework has two critical attacks: (a, b) and (c, b).
It can be checked that any pref-stable relation will return exactly one pref-stable extension: $\succeq_{max} = \{\{b, d, e\}\}$.

Let us now consider the following argumentation framework that is obtained after inverting the arrows of the two critical attacks.



It is easy to check that the only stable extension of this framework is the set $\{b, d, e\}$.

We will show that the same result can be obtained for two relations we proposed for generalising preferred and grounded semantics.

Theorem 4.3.11. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF, let $\succeq_p$ be the relation from Definition 4.3.6 and let $\succeq^p_{max}$ be the set of maximal elements of $\mathcal{T}$ w.r.t. that relation. Then, $\mathcal{E} \in \succeq^p_{max}$ iff:

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$, and
- $(\forall a' \in \mathcal{E}')$ $(\forall a \in \mathcal{A} \setminus \mathcal{E}')$ if $(((a, a') \in \mathcal{R} \wedge (a', a) \notin >)$ or $((a', a) \in \mathcal{R} \wedge (a, a') \in >))$ then $(\exists b' \in \mathcal{E}')$ s.t. $((b', a) \in \mathcal{R}$ and $(a, b') \notin >)$ or $((a, b') \in \mathcal{R}$ and $b' > a)$, and
- $\mathcal{E}'$ is a maximal set (w.r.t. set inclusion) which satisfies previous two items.

The following result is a consequence of the previous theorem.

Corollary 4.3.2. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF, let $\mathcal{R}' = \{(a,b) \mid a,b \in \mathcal{A},\ (a\mathcal{R}b \text{ and not}(b > a)) \text{ or } (b\mathcal{R}a \text{ and } a > b)\}$, and let $\succeq^p_{max}$ be the set of all maximal elements w.r.t. $\succeq_p$. Then: $\succeq^p_{max}$ is exactly the set of all preferred extensions of $(\mathcal{A}, \mathcal{R}')$.

The corresponding results for relation $\succeq_g$ are as follows. First, we show how to characterize the pref-grounded extension.

Theorem 4.3.12. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF, let $\succeq_g$ be a relation from Definition 4.3.8 and let $\succeq^g_{max}$ be the set of maximal elements of $\mathcal{T}$ w.r.t. that relation. Then, $\mathcal{E} \in \succeq^g_{max}$ iff:

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$, and
- $(\forall a \in \mathcal{E})\ sd(a, \mathcal{E})$ and
- $\mathcal{E}$ is a maximal set (w.r.t. set inclusion) which satisfies previous two items.

Now, we can show that the pref-grounded extension can also be obtained by inverting critical attacks.

Theorem 4.3.13. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and let $\mathcal{R}' = \{(a,b) \mid a,b \in \mathcal{A},\ (a\mathcal{R}b \text{ and not}(b > a)) \text{ or } (b\mathcal{R}a \text{ and } a > b)\}$, and let $\succeq^g_{max}$ be the set of all maximal elements w.r.t. $\succeq_g$. Then: $\succeq^g_{max}$ contains only one set which is exactly the grounded extension of $(\mathcal{A}, \mathcal{R}')$.

4.4 Rich preference-based argumentation framework

In the previous section, we proposed a PAF for handling critical attacks. Now, we propose a model that integrates both roles of preferences.

The general procedure we propose for modeling both roles follows two steps. Given an input $(\mathcal{A}, \mathcal{R}, \geq)$, the first step handles critical attacks using our approach. The output of this step is a set $\{\mathcal{E}_1, \ldots, \mathcal{E}_n\}$ of extensions under a given semantics. The second step consists of using a refinement relation to compare those extensions.

A refinement relation can be any preorder on the set $\mathcal{P}(\mathcal{A})$. An example of such a relation is the so-called *democratic relation* (Cayrol, Royer, and Saurel, 1993).

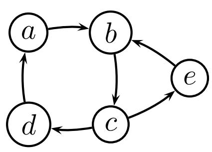
Definition 4.4.1 (Democratic relation)**.** *Let X be a set of objects and $\geq \subseteq X \times X$ be a preorder. For $\mathcal{S}, \mathcal{S}' \subseteq X$, $\mathcal{S} \succeq_d \mathcal{S}'$ iff $\forall x' \in \mathcal{S}' \setminus \mathcal{S}$, $\exists x \in \mathcal{S} \setminus \mathcal{S}'$ such that $x > x'$.*

Let us define our rich model which integrates both roles of preferences. For simplicity reasons, the first role is encoded by inverting the arrows of critical attacks.

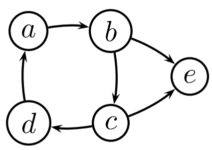
Definition 4.4.2. *A* rich PAF *is a tuple $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq, \succeq)$ where $\mathcal{A}$ is a set of arguments, $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ an attack relation, $\geq \subseteq \mathcal{A} \times \mathcal{A}$ is a preference relation and $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$ is a refinement relation s.t. both preference and refinement relations are reflexive and transitive.*
The basic part of $\mathcal{T}$ is PAF $(\mathcal{A}, \mathcal{R}, \geq)$. The extensions of $(\mathcal{A}, \mathcal{R}, \geq)$, denoted $\texttt{Ext}((\mathcal{A}, \mathcal{R}, \geq))$, are exactly the extensions of the argumentation framework $(\mathcal{A}, \mathcal{R}')$ (w.r.t. the same semantics), where $\mathcal{R}' = \{(a,b) \in \mathcal{A} \times \mathcal{A} \mid (a\mathcal{R}b$ and $not(b > a))$ or $(b\mathcal{R}a$ and $a > b)\}$.
The set of extensions of the rich PAF $\mathcal{T}$ is the set $\{\mathcal{E} \in \texttt{Ext}((\mathcal{A}, \mathcal{R}, \geq)) \mid \nexists \mathcal{E}' \in \texttt{Ext}((\mathcal{A}, \mathcal{R}, \geq))$ s.t. $\mathcal{E}' \succ \mathcal{E}\}$.

In other words, the extensions of a rich PAF are the best elements among the extensions of its basic part.

Example 4.4.1. *Let us consider the argumentation framework depicted below.*



Assume that $b > c$, $d > a$ and $b > e$ and let the refinement relation be the democratic relation $\succeq_d$. Basic PAF $(\mathcal{A}, \mathcal{R}, \geq)$ has exactly one critical attack, that from e to b. The framework $(\mathcal{A}, \mathcal{R}')$ is depicted below:



Thus, basic PAF $(\mathcal{A}, \mathcal{R}, \geq)$ has exactly two stable extensions: $\{a, c\}$ and $\{b, d\}$. According to $\succeq_d$, we have $\{b, d\} \succ_d \{a, c\}$; consequently rich PAF $(\mathcal{A}, \mathcal{R}, \geq, \succeq_d)$ has exactly one extension, $\{b, d\}$.

In the previous example, we used democratic relation as a refinement relation. Other relations may be used instead. The choice of a refinement relation is related to the particular application. There is a huge literature on ranking sets of objects based on preferences, see for example the work by Barbera, Bossert, and Pattanaik (2001).

It is easy to see that any extension of a rich PAF is conflict-free with respect to $\mathcal{R}$. This is a consequence of the fact that a set is conflict-free w.r.t. $\mathcal{R}$ iff it is conflict-free w.r.t. $\mathcal{R}'$.

Proposition 4.4.1. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq, \succeq)$ be a rich PAF. All extensions of $\mathcal{T}$ are conflict-free.

The following result shows that in the particular case where the preference relation $\geq$ is a linear order (i.e. reflexive, antisymmetric, transitive and total), then the basic part $(\mathcal{A}, \mathcal{R}, \geq)$ of any corresponding rich PAF has a unique stable/preferred extension. It is clear that in this case, there is no need to refine the result. Namely, the rich PAF has the same extension independently of the refinement relation.

Proposition 4.4.2. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a basic PAF s.t. $\mathcal{R}$ is irreflexive and $\geq$ is a linear order.

- Stable, preferred and grounded extensions of $\mathcal{T}$ coincide.
- $\mathcal{T}$ has exactly one stable extension.
- If $|\mathcal{A}| = n$, then this extension is computed in $\mathcal{O}(n^2)$ time.

Our proposition of rich PAF has several advantages. First, it is the only framework that models both roles of preferences for any semantics and any attack relation. Moreover, if it is desirable to compare any pair of sets, dominance relations may be used in the first step instead of inverting arrows. Another interesting feature of our framework is that at the second step, for comparing the basic extensions computed after the first step, one can choose any preference relation. It is guaranteed that extensions obtained after the first step are already useful (e.g. they are conflict-free). Thus, in the second step, one needs only to choose between them, i.e. to refine this result.

4.5 Links with non-argumentative approaches

This section shows that two particular instantiations of the rich PAF presented in the previous section capture respectively the preferred (Brewka,

1989) and democratic (Cayrol et al., 1993) sub-theories, which were proposed for handling inconsistency in prioritised knowledge bases. Throughout this section, we assume a propositional knowledge base Σ.

The two instantiations of our rich PAF use the set $\texttt{Arg}(\Sigma)$ as set of arguments and undercut (Definition 3.4.2) as attack relation.

Note that inconsistent formulae are not used in construction of arguments; they do not appear in preferred (or democratic) sub-theories neither. Thus, in the rest of the chapter, we assume that a knowledge base Σ contains only consistent formulae.

Proposition 4.5.1. Let Σ be a propositional knowledge base and ($\texttt{Arg}(\Sigma)$, Undercut) the argumentation framework built from Σ.

- For any consistent set $\mathcal{S} \subseteq \Sigma$, $\mathcal{S} = \texttt{Base}(\texttt{Arg}(\mathcal{S}))$.
- The function $\texttt{Base} : \texttt{Arg}(\Sigma) \to \Sigma$ is surjective.
- For any $\mathcal{E} \subseteq \texttt{Arg}(\Sigma)$, $\mathcal{E} \subseteq \texttt{Arg}(\texttt{Base}(\mathcal{E}))$.
- The function $\texttt{Arg} : \Sigma \to \texttt{Arg}(\Sigma)$ is injective.

Another property that is important for the rest of the chapter relates the notion of consistency of a set of formulae to that of conflict-freeness of a set of arguments.

Proposition 4.5.2. A set $\mathcal{S} \subseteq \Sigma$ is consistent *iff* $\texttt{Arg}(\mathcal{S})$ is conflict-free w.r.t. undercut.

The following example shows that the previous proposition does not hold for function $\texttt{Base}$.

Example 4.5.1. *Let* $\mathcal{E} = \{(\{x\}, x), (\{x \to y\}, x \to y), (\{\neg y\}, \neg y)\}$. *It is obvious that* $\mathcal{E}$ *is conflict-free w.r.t. undercut while* $\texttt{Base}(\mathcal{E})$ *is not consistent.*

We show that if a preference relation $\geq$ between arguments is a total preorder, then the stable extensions of $(\texttt{Arg}(\Sigma), \text{Undercut}, \geq)$ are all incomparable w.r.t. the democratic relation $\succeq_d$.

Proposition 4.5.3. Let $\mathcal{T} = (\texttt{Arg}(\Sigma), \text{Undercut}, \geq)$ be a basic PAF and let $\geq$ be a total preorder (i.e. any pair of arguments is comparable). Then: for all stable extensions $\mathcal{E}$ and $\mathcal{E}'$ of $\mathcal{T}$, if $\mathcal{E} \neq \mathcal{E}'$, then $\neg(\mathcal{E} \succeq_d \mathcal{E}')$.

From the previous proposition, it follows that the stable extensions of basic PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq$) coincide with those of the rich PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq$, $\succeq_d$).

Corollary 4.5.1. If $\geq$ is a total preorder, then the stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq$, $\succeq_d$) are exactly the stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq$).

As a consequence, in the first part of our study (i.e. when $\geq$ is total, and consequently, Σ is stratified), we will use a basic framework ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp}$) since refinement is not necessary in this case.

4.5.1 Recovering preferred sub-theories

A notion of *preferred sub-theory* has been defined by Brewka (1989). It supposes a stratified knowledge base $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ and uses these preferences on the set of formulae in order to choose the best sets among the maximal consistent subsets of Σ.

Definition 4.5.1 (Brewka, 1989)**.** *Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a stratified knowledge base. Let $\mathcal{S} \subseteq \Sigma$ and let $\forall i \in \{1, \ldots, n\}$, $\mathcal{S}_i = \mathcal{S} \cap \Sigma_i$. $\mathcal{S}$ is a* preferred sub-theory *of Σ iff $\forall k \in \{1, \ldots, n\}$, $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_k$ is a maximal (for set inclusion) consistent set in $\Sigma_1 \cup \ldots \cup \Sigma_k$.*

The following proposition is a consequence of Definition 4.5.1.

Proposition 4.5.4. Every preferred sub-theory of $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ is a maximal consistent set in Σ.

Example 4.5.2. *Let $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$ with $\Sigma_1 = \{strad\}$, $\Sigma_2 = \{strad \to exp\}$, $\Sigma_3 = \{\neg exp\}$, where strad stands for "the violin is a Stradivarius" and exp for "the violin is a expensive". There are three maximal consistent subsets of Σ: $\mathcal{S}_1 = \{strad, strad \to exp\}$, $\mathcal{S}_2 = \{strad, \neg exp\}$ and $\mathcal{S}_3 = \{strad \to exp, \neg exp\}$. Only $\mathcal{S}_1$ is a preferred sub-theory.*

In the rest of the subsection, we will show that there is a full correspondence between the preferred sub-theories of a stratified knowledge base Σ and the stable extensions of the basic PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp}$). Recall that the relation $\geq_{wlp}$ is based on the weakest link principle and privileges the arguments whose less important formulae are more important than the less important formulae of the other arguments. This relation is a total preorder (and is defined over a knowledge base that is itself equipped with a total preorder). Recall that, according to Corollary 4.5.1, the stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp}$) coincide with those of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp}, \succeq_d$).

The first result shows that from each preferred sub-theory is built a stable extension of PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp}$).

Theorem 4.5.1. Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a stratified knowledge base. For every preferred sub-theory $\mathcal{S}$ of Σ, it holds that:

- $\texttt{Arg}(\mathcal{S})$ is a stable extension of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp})$
- $\mathcal{S} = \texttt{Base}(\texttt{Arg}(\mathcal{S}))$

Similarly, we show that the base of each stable extension of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp})$ is a preferred sub-theory of Σ and that it contains all arguments that can be built from its base.

Theorem 4.5.2. Let Σ be a stratified knowledge base. For every stable extension $\mathcal{E}$ of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp})$, it holds that:

- $\texttt{Base}(\mathcal{E})$ is a preferred sub-theory of Σ
- $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$

The next theorem shows that there exists a one-to-one correspondence between preferred sub-theories of Σ and stable extensions of the framework $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp})$.

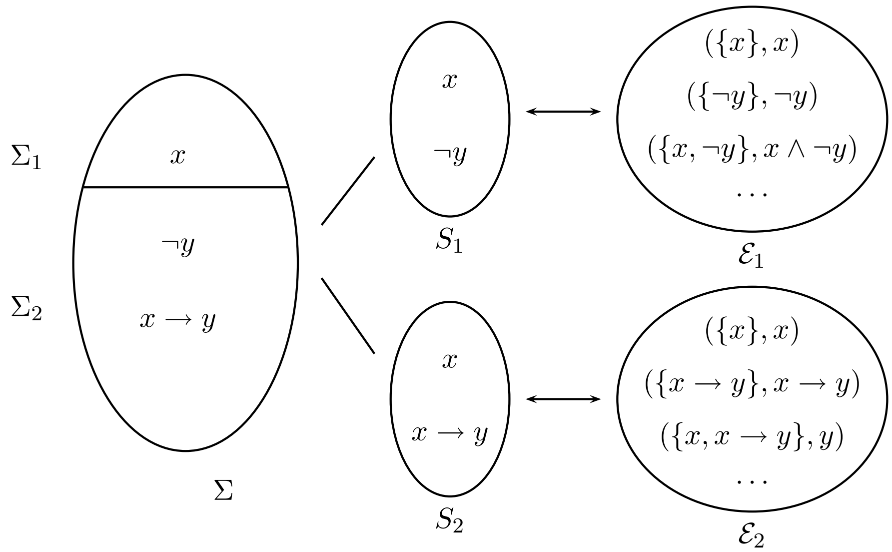
Theorem 4.5.3. Let $\mathcal{T} = (\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp})$ be a basic PAF built from a stratified knowledge base Σ. The stable extensions of $\mathcal{T}$ are exactly $\texttt{Arg}(\mathcal{S})$ where $\mathcal{S}$ ranges over the preferred sub-theories of Σ.

From the above result, since any Σ has at least one preferred sub-theory, it follows that the basic PAF $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp})$ has at least one stable extension.

Corollary 4.5.2. The PAF $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp})$ has at least one stable extension.

Example 4.5.3. *Let $\Sigma = \Sigma_1 \cup \Sigma_2$ with $\Sigma_1 = \{x\}$ and $\Sigma_2 = \{\neg y, x \rightarrow y\}$ be a stratified propositional knowledge base. There are two preferred sub-theories, $S_1 = \{x, \neg y\}$ and $S_2 = \{x, x \rightarrow y\}$. The argumentation framework $(\texttt{Arg}(\Sigma), Undercut, \geq_{wlp})$ has exactly two stable extensions: $\mathcal{E}_1 = \texttt{Arg}(S_1)$ and $\mathcal{E}_2 = \texttt{Arg}(S_2)$. Figure 4.1 shows the two preferred sub-theories of Σ as well as the two stable extensions of the corresponding PAF.*

Figure 4.1: Preferred sub-theories of Σ and stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp}$)



4.5.2 Recovering the democratic sub-theories

Cayrol, Royer, and Saurel (1993) have extended the notion of preferred sub-theory to the case where Σ is equipped with an arbitrary preorder $\trianglerighteq$, i.e. not necessarily a total one. The basic idea is to define a preference relation on the power set of Σ. The best elements according to this relation are called *democratic sub-theories*.[4] The relation that generalises preferred sub-theories is the democratic relation (Definition 4.4.1).

Definition 4.5.2 (Cayrol et al., 1993)**.** *Let Σ be a propositional knowledge base and $\trianglerighteq \subseteq \Sigma \times \Sigma$ be a partial preorder. A* democratic sub-theory *is a set $\mathcal{S} \subseteq \Sigma$ such that $\mathcal{S}$ is consistent and $(\nexists \mathcal{S}' \subseteq \Sigma)$ s.t. $\mathcal{S}'$ is consistent and $\mathcal{S}' \succeq_d \mathcal{S}$.*

From the previous definition, we see that any democratic sub-theory is a maximal consistent set of Σ. Thus, like in the case of preferred sub-theories, preferences are used to *refine* the results obtained without preferences, i.e. to keep only some and not all maximal consistent sets.

[4]They are called "demo-preferred sets" by Cayrol et al. (1993).

Example 4.5.4. *Let us suppose the knowledge base* $\Sigma = \{strad, strad \rightarrow exp, \neg exp\}$. *Let us suppose that the formula* $strad \rightarrow exp$ *is preferred to the two other formulae, which are themselves incomparable. In this case, there are exactly two democratic sub-theories,* $\{strad \rightarrow exp, strad\}$ *and* $\{strad \rightarrow exp, \neg exp\}$.

It can be shown that democratic sub-theories generalise preferred sub-theories.

Proposition 4.5.5. Let $(\Sigma, \trianglerighteq)$ be a prioritized knowledge base, $\trianglerighteq$ be a total preorder and let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a corresponding stratified knowledge base, i.e. $\forall i, j \in \{1, \ldots, n\}$ $\forall x \in \Sigma_i$, $\forall y \in \Sigma_j$ we have $x \trianglerighteq y$ iff $i \leq j$. Then: $\forall \mathcal{S} \subseteq \Sigma$, $\mathcal{S}$ is a preferred sub-theory of $\Sigma_1 \cup \ldots \cup \Sigma_n$ iff $\mathcal{S}$ is a democratic sub-theory of $(\Sigma, \trianglerighteq)$.

In order to capture democratic sub-theories, we will use the generalised version of the preference relation $\geq_{wlp}$. We use relation $\geq_{gwlp}$ as defined in Definition 4.2.3. However, two remarks have to be made at this point. The first is that the relation $\geq_{gwlp}$ is not reflexive. If needed, it can easily be redefined in order to become reflexive. Another remark is that this relation does not formally generalize the relation $\geq_{wlp}$. Namely, when $\trianglerighteq$ is a total preorder, $\geq_{wlp}$ and $\geq_{gwlp}$ do *not* coincide. However, the strict version $>_{gwlp}$ of $\geq_{gwlp}$ generalises the strict version $>_{wlp}$ of the relation based on the weakest link principle. Since we are using those relations in order to treat critical attacks, then when $\trianglerighteq$ is a total preorder, whether $\geq_{wlp}$ or $\geq_{gwlp}$ is used for calculating $\mathcal{R}'$ is irrelevant since attack are inverted only in case of strict preference. This shows why we call $\geq_{gwlp}$ a generalisation of $\geq_{wlp}$. Note also that in this case (when $\trianglerighteq$ is total), there is no refinement.

At this point, it becomes clear that the results from this subsection are generalisations of the results from the previous one, i.e. $(\texttt{Arg}(\Sigma), \text{Undercut}, \geq_{wlp})$ and $(\texttt{Arg}(\Sigma), \text{Undercut}, \geq_{gwlp}, \succeq_d)$ return identical results when $\trianglerighteq$ is a total preorder. However, we presented the particular case first, since we think that it is easier to understand ideas and proofs in this case, and then to pass to the more general case.

It can be shown that from each democratic sub-theory of a knowledge base Σ, a stable extension of $(\texttt{Arg}(\Sigma), \text{Undercut}, \geq_{gwlp})$ is built.

Theorem 4.5.4. Let Σ be a knowledge base which is equipped with a partial preorder $\trianglerighteq$. For every democratic sub-theory $\mathcal{S}$ of Σ, it holds that $\texttt{Arg}(S)$ is a stable extension of basic PAF $(\texttt{Arg}(\Sigma), \text{Undercut}, \geq_{gwlp})$.

The following result shows that each stable extension of the basic PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}$) returns a maximal consistent subset of Σ.

Theorem 4.5.5. Let Σ be a knowledge base equipped with a partial preorder $\trianglerighteq$. For every stable extension $\mathcal{E}$ of ($\texttt{Arg}(\Sigma)$, Undercut , $\geq_{gwlp}$), it holds that:

- $\texttt{Base}(\mathcal{E})$ is a maximal (for set inclusion) consistent subset of Σ.
- $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$.

The following example shows that the stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}$) do not necessarily return democratic sub-theories.

Example 4.5.5. *Recall that* $\Sigma = \{x, \neg x, y, \neg y\}$, $\neg x \trianglerighteq y$ *and* $\neg y \trianglerighteq x$. *Let* $\mathcal{S} = \{x, y\}$. *It can be checked that the set* $\texttt{Arg}(\mathcal{S})$ *is a stable extension of* ($\texttt{Arg}(\Sigma)$, *Undercut,* $\geq_{gwlp}$). *However,* $\mathcal{S}$ *is not a democratic sub-theory since* $\{\neg x, \neg y\} \succ_d \mathcal{S}$.

It can also be shown that a knowledge base may have a maximal consistent subset $\mathcal{S}$ s.t. $\texttt{Arg}(\mathcal{S})$ is not a stable extension of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}$). Let us consider the following example.

Example 4.5.6. *Let* $\Sigma = \{x, \neg x\}$ *and* $x \rhd \neg x$. *It is clear that* $\{\neg x\}$ *is a maximal consistent subset of* Σ *while* $\texttt{Arg}(\{\neg x\})$ *is not a stable extension of* ($\texttt{Arg}(\Sigma)$, *Undercut,* $\geq_{gwlp}$).

The following result establishes a link between the 'best' maximal consistent subsets of Σ w.r.t. the democratic relation $\succeq_d$ and the 'best' sets of arguments w.r.t. the same relation $\succeq_d$.

Theorem 4.5.6. Let $\mathcal{S}, \mathcal{S}' \subseteq \Sigma$ be maximal (for set inclusion) consistent subsets of Σ. It holds that $\mathcal{S} \succeq_d \mathcal{S}'$ iff $\texttt{Arg}(\mathcal{S}) \succeq_d \texttt{Arg}(\mathcal{S}')$.

Theorem 4.5.7. Let Σ be equipped with a partial preorder $\trianglerighteq$.

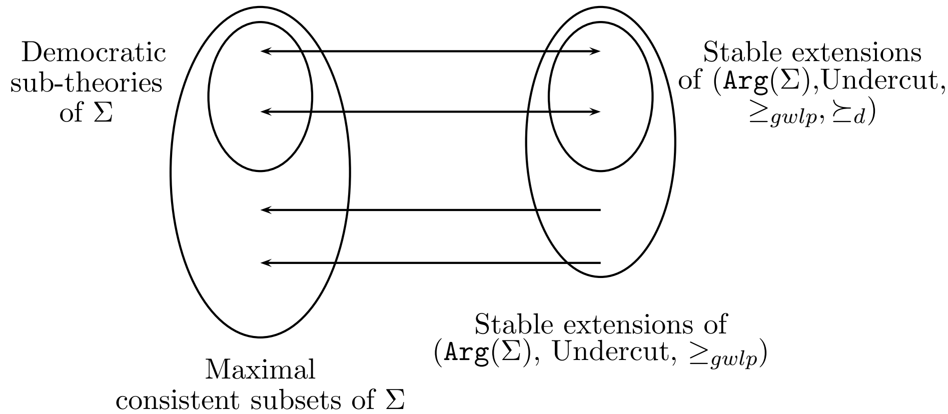
- For every democratic sub-theory $\mathcal{S}$ of Σ, $\texttt{Arg}(\mathcal{S})$ is a stable extension of the rich PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}, \succeq_d$).
- For each stable extension $\mathcal{E}$ of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}, \succeq_d$), $\texttt{Base}(\mathcal{E})$ is a democratic sub-theory of Σ.

Finally, we show that there is a one-to-one correspondence between the democratic sub-theories of a base Σ and the stable extensions of its corresponding rich PAF.

Theorem 4.5.8. The stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}, \succeq_d$) are exactly the $\texttt{Arg}(\mathcal{S})$ where $\mathcal{S}$ ranges over the democratic subtheories of Σ.

Figure 4.2 synthesizes different links between Σ and the corresponding PAF and rich PAF.

Figure 4.2: Democratic sub-theories of Σ and stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp}, \succeq_d$)



4.6 Conclusion

In this chapter, we have studied the role of preferences in argumentation frameworks. We started by presenting several examples of preference relations. Then, we showed that there are two roles of preferences, namely conflict-resolution role and refinement role. A survey of the state of the art in preference-based argumentation showed that the most of the frameworks model only the first role. We showed that there are situations when existing frameworks do not return desirable results. Then, we proposed a new approach for modeling this role of preferences. Our approach presents two novelties: First, it takes into account preferences at a semantic level, i.e. it defines new acceptability semantics which are grounded on attacks and preferences between arguments. Second, a semantics is defined as a dominance relation that compares any pair of subsets of arguments. We proposed three particular relations which generalise stable, preferred and grounded semantics. Then, we focused on stable semantics and studied all relations that generalise this semantics.

We also proposed a way to take into account both roles of preferences, namely through the definition of rich PAF. At the end, we showed that our proposition is general and sound since there are full correspondences between several instantiations of our rich PAF and preferred and democratic sub-theories.

Note that in related non-argumentative approaches preferences do not play the same role as in argumentation. For example, in nonmonotonic reasoning, preferences are used in order to narrow down the number of possible belief sets of a base theory (Brewka, Niemela, and Truszczynski, 2008). To say it differently, from a given base theory, a first set of *standard* solutions (belief sets) is computed, then a subset of those solutions (called *preferred* solutions) is chosen on the basis of available preferences. Thus, preferences *refine* the standard solutions.

Brewka, Truszczynski, and Woltran (2010) have proposed another way to integrate preferences into argumentation. They argue that the extensions should be calculated without taking into account preferences and that the set of obtained extensions should be then refined by the use of preferences. This corresponds exactly to the second role of preferences we have identified (Amgoud and Vesic, 2010c,b, 2011e). However, we believe that the first role of preferences cannot be omitted and that the refinement role does not subsume the conflict-resolution role of preferences in argumentation.

We have shown that preferences *intervene twice* in an argumentation framework. They are mandatory for: i) computing the standard solutions of an AF, and then ii) for narrowing the number of those solutions. We have seen in this chapter that those two roles of preferences are completely independent and none of them can be modeled by the other one.

Another work which handles the problem of critical attacks is the framework proposed by Prakken (2011). In that paper, a logic-based instantiation of Dung's framework is developed, in which three kinds of attacks are considered: rebuttal, assumption attack and undercut. For each relation, the author has found a way to avoid the problem of critical attack and ensured conflict-free extensions. We think that our work is more general since we solved the problem at an abstract level (for any set of arguments, any attack relation and any preference relation).

We would like also to mention the work done by Kaci (2010). In that paper, the author made a survey of the critics presented in existing works (Amgoud and Vesic, 2009b; Dimopoulos, Moraitis, and Amgoud, 2009) against existing approaches for PAFs. The author concluded that one should use a symmetric attack relation in order to avoid the problem of conflicting extensions. That suggestion is certainly not realistic, especially in light of results by Amgoud and Besnard (2009) stating that there are many cases when symmetric relations should be avoided.

An extension of our work would be to characterize the different dominance relations that generalise preferred semantics and those which generalise grounded semantics. A similar work can be done on those semantics proposed by Baroni, Giacomin, and Guida (2005), ideal semantics (Dung, Mancarella, and Toni, 2007) and semi-stable semantics (Caminada, 2006b). Another future work consists of studying how the new semantics can be used in a decision making context in order to rank order a set of alternatives.

In my house I'm the boss,
my wife is just the decision maker.
Woody Allen

5 Argumentation-based decision making

In this chapter, we study argumentation-based decision making. We start by defining a decision-making problem. Then, we survey the most important works in argumentation-based decision making. In Section 5.2, we recall the decision model we will be using, initially proposed by Amgoud et al. (2008). The contribution of this chapter is the study of dynamics of that framework. In particular, in Section 5.3, we show how the status of options changes when a new argument is received (Amgoud and Vesic, 2009a, 2011d). The last section concludes.

5.1 Introduction

Decision making is as a process leading to the selection of an option between several possible alternatives. We will illustrate the decision making problem by presenting an example borrowed from Savage (1954).

Example 5.1.1. *An agent knows that there are five good eggs in a bowl and has another egg in his hand. We suppose that he has only three actions available:*

- *join, to break the egg to join the other five eggs,*
- *inspect, to break the egg into a saucer for inspection,*
- *throw, to throw the egg away without inspection.*

The uncertainty is reflected in the fact that the agent does not know if the egg is good or bad. Consequences of the actions are given in the table bellow:

act / state	*good egg*	*bad egg*
join	*six-egg omelet*	*no omelet*
inspect	*six-egg omelet, a saucer to wash*	*five-egg omelet, a saucer to wash*
throw	*five-egg omelet*	*five-egg omelet*

States of the world (relevant to this problem) are:

- *good, the egg is god,*
- *bad, the egg is bad.*

Decision making problem aims at choosing exactly one action between possible alternatives.

An agent is supposed to have *preferences* between different outcomes. For example, an agent may prefer the outcome "six-egg omelet" to the outcome "no omelet" etc. In the general case, this preference may be difficult to determine completely. In the best case, an agent can attribute utility value to every outcome, thus yielding to a total preorder on the set of outcomes. In that approach, the agent can compare all outcomes, for example (s)he must be able to say if (s)he prefers "six-egg omelet, a saucer to wash" or "five-egg omelet". In other approaches, an agent can express his/her preferences using different criteria. For example, (s)he could, on the one hand prefer "six-egg omelet" to "five-egg omelet", and on the other, prefer "no saucer to wash" to "a saucer to wash". The task to be solved here is to aggregate those criteria; this is called multi criteria decision making (Roy, 1985). There is even another case in which an agent has a unique decision criterion but is involved in group decision making. Aggregation of different agents' preferences is studied in social choice theory (Arrow, 1951).

Even if decision making theory is largely inspired by human decision making, humans may make errors and can sometimes act irrationally in the sense that the choice they make is not always in accordance with their preferences. That is why the notion of a *rational agent* has been defined. A rational agent is an agent which always chooses to perform the action that results in the optimal outcome for itself from among all feasible actions. The action a rational agent takes depends on: the set of actions available to the agent, the preferences of the agent, the agent's information about the current state of the world, and the estimated benefits of the actions.

5.1.1 Argumentation-based decision making

In a decision making context, argumentation has obvious benefits. Indeed, in everyday life, decision is often based on arguments and counter-arguments. Argumentation can also be useful for explaining a choice already made.

Another advantage of argumentation is that it is a powerful approach for handling inconsistency in knowledge bases. Thus, not only it can rank order options in a decision problem, but it can do that under inconsistent information.

Argumentation has been used for decision making by different authors. In particular, Fox and Parsons (1997) have developed an inference-based decision support system. An implementation of this system was made for medical applications (Fox and Das, 2000). Another example of argument-based decision system that is purely based on an inference system is proposed by Chesnevar et al. (2006) for advising about language usage assessment on the basis of corpus available on the web.

Bonet and Geffner (1996b) proposed an original approach to qualitative decision making, inspired by Tan and Pearl (1994), based on *action rules* that link a situation and an action with the satisfaction of a positive or a negative goal. This framework contains: a set of actions, a set of input propositions where to each proposition is attached a degree of plausibility (e.g. likely, plausible), two sets of prioritized goals, one containing the positive goals and the other the negative ones (i.e. those that should be avoided), and a set of action rules, where the left side contains an action and input literals and the right side contains a goal (e.g. $goBeach \wedge \neg rain \rightarrow enjoyBeach$). To each action is associated a priority level which is the priority of the goal, and a plausibility level, which is defined on the base of plausibility of input literals appearing in the rule. In this approach only input propositions are weighted in terms of plausibility. Action rules inherit these weights in an empirical manner which depends on the chosen plausibility scale. The action rules themselves are not weighted since they are potentially understood as defeasible rules, although no non-monotonic reasoning system is associated with them.

Dubois and Fargier (2006) studied a framework where a candidate decision d is associated with two distinct sets of positive arguments and negative arguments. The authors provided an axiomatic characterization of different rules in this setting, with a possibility theory interpretation of their meaning. For example, a *bipolar lexicographic* preference relation is characterized.

Another trend of works relating argumentation and decision is mainly interested in the use of arguments for explaining and justifying multiple criteria decisions once they have been made using some definite aggregation function. A systematic study for different aggregation functions was done by Labreuche (2006).

Besides, a general and abstract argument-based framework for decision making was proposed by Amgoud and Prade (2009). This framework follows two main steps. At the first step, arguments for beliefs and arguments for options are built and evaluated using classical acceptability semantics. At the second step, pairs of options are compared using decision principles. Decision principles are based on the accepted arguments supporting the options. Three classes of decision principles are distinguished: unipolar, bipolar or non-polar principles depending on whether i) only arguments pro or only arguments con, or ii) both types, or iii) an aggregation of them into a meta-argument are used. The abstract model is then instantiated by expressing formally the mental states (beliefs and preferences) of a decision maker. In the proposed framework, information is given in the form of a stratified set of beliefs. The bipolar nature of preferences is emphasized by making an explicit distinction between prioritized goals to be pursued, and prioritized rejections that are stumbling blocks to be avoided. A typology that identifies four types of argument is also proposed. Indeed, each decision is supported by arguments emphasizing its positive consequences in terms of goals certainly satisfied and rejections certainly avoided. A decision can also be attacked by arguments emphasizing its negative consequences in terms of certainly missed goals, or rejections certainly led to by that decision.

While there are several works on modeling decision problems by argumentation techniques, there is no work on the dynamics of these models. To say it differently, there is no work that shows how the status of options (i.e. decisions) change when a new argument arrives. The goal of this chapter is to answer that question. In order to do so, we must study the evolution of the status of *a given argument* without having to compute the extensions of the new argumentation framework, as done by Cayrol et al. (2008). Furthermore, we will study the most general case, i.e. the new argument may attack and be attacked by an arbitrary number of arguments of the initial argumentation framework. Finally, we are interested in two acceptability semantics: grounded and preferred semantics.

5.2 An argumentation-based decision framework

In the rest of this chapter, we are interested by a decision model proposed by Amgoud et al. (2008). Our choice is mainly motivated by the fact that this model is general enough to encode different decision criteria.

In what follows, $\mathcal{L}$ will denote a logical language, from which a finite set $\mathcal{O} = \{o_1, \ldots, o_n\}$ of n distinct *options* is identified; the decision maker has to choose exactly one of them. Note that an option o may be the "conjunction" of other options in $\mathcal{O}$. Let us consider the following example borrowed from Amgoud et al. (2008).

> Assume that Carla wants a drink and has to choose between tea, milk or both. Thus, there are three options: o_1: tea, o_2: milk, and o_3: tea and milk.

Two kinds of arguments are distinguished: arguments supporting options, called *practical* arguments and gathered in a set $\mathcal{A}_o$, and arguments supporting beliefs, called *epistemic* arguments and gathered in a set $\mathcal{A}_b$, such that $\mathcal{A}_o \cap A_b = \emptyset$. The structure of these arguments is not specified. For instance, an epistemic argument may involve beliefs while a practical argument involves beliefs and benefits/goals that may be reached if the option supported by that argument is chosen. We will suppose that those arguments are collected by an agent, thus both $\mathcal{A}_b$ and $\mathcal{A}_o$ are finite.

Practical arguments are linked to the options they support by a function $\mathcal{H}$ defined as follows:

$$\mathcal{H}\colon \mathcal{O} \to 2^{\mathcal{A}_o} \text{ where } \forall i, j \text{ if } i \neq j \text{ then } \mathcal{H}(o_i) \cap \mathcal{H}(o_j) = \emptyset \text{ and } \mathcal{A}_o = \bigcup_{i=1}^{n} \mathcal{H}(o_i) \text{ with } \mathcal{O} = \{o_1, \ldots, o_n\}.$$

Each practical argument a supports exactly one option o. We say that o is the conclusion of the practical argument a, and write $\texttt{Conc}(a) = o$. Note that there may exist options that do not have arguments in their favor (i.e. such that $\mathcal{H}(o) = \emptyset$).

Example 5.2.1. *Let* $\mathcal{O} = \{o_1, o_2, o_3\}$, $\mathcal{A}_b = \{b_1, b_2, b_3\}$, $\mathcal{A}_o = \{a_1, a_2, a_3\}$ *and let the arguments supporting the three options be as in the table below.*

$\mathcal{H}(o_1)$	$= \{a_1\}$
$\mathcal{H}(o_2)$	$= \{a_2, a_3\}$
$\mathcal{H}(o_3)$	$= \emptyset$

Three preference relations between arguments are considered. They express the fact that some arguments may be stronger than others. The first preference relation, denoted by $\geq_b$, is a preorder[1] on the set $\mathcal{A}_b$. For example, an argument which is built from more certain information may be considered as stronger than an argument based on less certain information. The second relation, denoted by $\geq_o$, is a preorder on the set $\mathcal{A}_o$. It should be based both on the certainty degrees of the information involved in the arguments and on the importance of the benefits of the options. Finally, a third preorder,

[1] Recall that a relation is a preorder iff it is *reflexive* and *transitive*.

denoted by $\geq_m$ (m for *mixed* relation), captures the idea that any epistemic argument is stronger than any practical argument. The role of epistemic arguments in a decision problem is to validate or to undermine the beliefs on which practical arguments are built. Thus, $(\forall a \in \mathcal{A}_b)(\forall a' \in \mathcal{A}_o)\ (a, a') \in \geq_m \wedge (a', a) \notin \geq_m$. Note that $(a, a') \in \geq_x$ (with $x \in \{b, o, m\}$) means that a is *at least as good as* a'. In what follows, $>_x$ denotes the strict relation associated with $\geq_x$. It is defined as $(a, a') \in >_x$ iff $(a, a') \in \geq_x$ and $(a', a) \notin \geq_x$.

Three conflict relations among arguments are also distinguished. The first one, denoted by $\mathcal{R}_b$, captures the conflicts that may hold between epistemic arguments. In the framework of Amgoud et al. (2008), the structure of this relation is not specified. The second relation, denoted $\mathcal{R}_o$, captures the conflicts among practical arguments. Two practical arguments are conflicting if they support distinct options. This is mainly due to the fact that the options are mutually exclusive and competitive. Formally, for all $a, b \in \mathcal{A}_o$, $(a, b) \in \mathcal{R}_o$ iff $\texttt{Conc}(a) \neq \texttt{Conc}(b)$. Finally, practical arguments may be attacked by epistemic ones. The idea is that an epistemic argument may challenge the belief part of a practical argument. However, practical arguments are not allowed to attack epistemic ones. This avoids wishful thinking, i.e. avoids making decisions according to what might be pleasing to imagine instead of by appealing to evidence. This relation, denoted by $\mathcal{R}_m$, contains pairs (a, a') where $a \in \mathcal{A}_b$ and $a' \in \mathcal{A}_o$.

In the framework of Amgoud et al. (2008), each conflict relation $\mathcal{R}_x$ (with $x \in \{b, o, m\}$) is combined with the preference relation $\geq_x$ into a unique relation between arguments, called defeat and denoted by $\texttt{Def}_x$, as follows: For all $a, b \in \mathcal{A}_b \cup \mathcal{A}_o$, $(a, b) \in \texttt{Def}_x$ iff ($a\mathcal{R}_x b$ and $\neg(b \geq_x a)$).

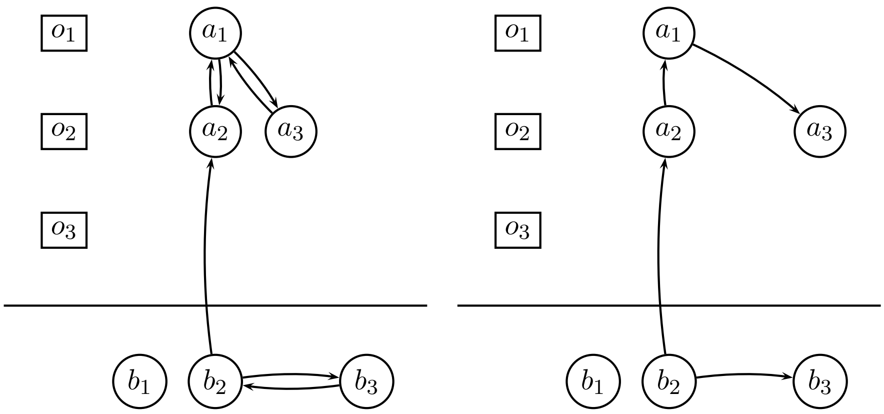
However, note that we have shown in the previous chapter that this may lead to contradictory extensions and counter-intuitive results. Recall also that we concluded that deleting attacks should be avoided; instead, preferences should be taken into account by inverting arrows corresponding to critical attacks. Consequently, we will define defeat relations as follows: $\forall a, b \in \mathcal{A}_b \cup \mathcal{A}_o$, $(a, b) \in \texttt{Def}_x$ iff ($a\mathcal{R}_x b$ and $\neg(b >_x a)$) or ($b\mathcal{R}_x a$ and $a >_x b$).

Let $\texttt{Def}_b$, $\texttt{Def}_o$ and $\texttt{Def}_m$ denote the three defeat relations corresponding to the three conflict relations. Since arguments in favor of beliefs are always preferred (in the sense of $\geq_m$) to arguments in favor of options, it trivially holds that $\mathcal{R}_m = \texttt{Def}_m$.

Throughout the paper, we use the following convention when depicting decision frameworks. Options, put in squares, are on the same line as their arguments. Epistemic arguments are separated from practical ones by a horizontal line.

Example 5.2.2. *Let us suppose options and arguments from Example 5.2.1. Let the graph on the left of Figure 5.1 depict the conflicts (w.r.t. $\mathcal{R}$) among arguments. Assume that $(b_2, b_3) \in >_b$, $(a_2, a_1) \in >_o$ and $(a_1, a_3) \in >_o$. The graph of* $\texttt{Def}$ *is depicted on the right of the same figure.*

Figure 5.1: Attack relation (left) and corresponding defeat relation (right).



Definition 5.2.1 (Decision framework)**.** *A* decision framework *is a tuple $\mathcal{AF} = (\mathcal{O}, \mathcal{A}_b \cup \mathcal{A}_o, \texttt{Def}_b \cup \texttt{Def}_o \cup \texttt{Def}_m, \mathcal{H})$.*

In the rest of the chapter, if not specified otherwise, we will use notation $\mathcal{A} = \mathcal{A}_b \cup \mathcal{A}_o$, $\geq = \geq_b \cup \geq_o \cup \geq_m$, $\mathcal{R} = \mathcal{R}_b \cup \mathcal{R}_o \cup \mathcal{R}_m$ and $\texttt{Def} = \texttt{Def}_b \cup \texttt{Def}_o \cup \texttt{Def}_m$. The arguments of $\mathcal{A}$ are evaluated in $(\mathcal{A}, \texttt{Def})$ using a given acceptability semantics.

Until now, we have used the usual definition of argument's status (i.e. Definition 2.2.11). That is, in the literature, an argument is credulously accepted if it is in at least one of the extensions. Thus, each argument that is sceptically accepted is also credulously accepted. In the framework defined by Amgoud et al. (2008), this definition was slightly modified. The reason is that in a decision making context, one looks for a preference relation on the set of options. Thus, it is important to distinguish between options that are supported by arguments in all the extensions, and those supported by arguments in only some extensions. From now on, we will call an argument credulously accepted only if it is not sceptically accepted. Formally, we will use the definition given by Amgoud et al. (2008).

Definition 5.2.2 (Status of arguments)**.** *Let* $\mathcal{AF} = (\mathcal{O}, \mathcal{A}, \mathtt{Def}, \mathcal{H})$ *be a decision framework, let* $\mathtt{Ext}(\mathcal{AF})$ *be the set of its extensions of* $\mathcal{F} = (\mathcal{A}, \mathtt{Def})$ *with respect to a given semantics and let* $a \in \mathcal{A}$*. The status of argument* a *is defined as follows:*

- a *is sceptically accepted (or sceptical) iff* $a \in \bigcap_{\mathcal{E} \in \mathtt{Ext}(\mathcal{F})} \mathcal{E}$
- a *is credulously accepted (or credulous) iff* $a \in \bigcup_{\mathcal{E} \in \mathtt{Ext}(\mathcal{F})} \mathcal{E}$ *and* $a \notin \bigcap_{\mathcal{E} \in \mathtt{Ext}(\mathcal{F})} \mathcal{E}$
- a *is rejected iff* $a \notin \bigcup_{\mathcal{E} \in \mathtt{Ext}(\mathcal{F})} \mathcal{E}$.

Let $\mathtt{Status}(a, \mathcal{F})$ *be a function which returns the status of an argument* a *in argumentation framework* $\mathcal{F}$*. By abuse of notation, we will sometimes use the same notation in the case when* $\mathcal{F}$ *is a decision framework; in that case, we suppose that the status is calculated using the set of all arguments and the set of all defeats of the corresponding decision framework.*

Example 5.2.3. *The decision framework of Figure 5.1 (graph on the right) has one preferred extension, which is also the grounded one,* $\{a_1, b_1, b_2\}$*. Thus, the three arguments* a_1*,* b_1*, and* b_2 *are sceptically accepted while* a_2*,* a_3 *and* b_3 *are rejected.*

Let the sets of sceptical, credulous and rejected arguments of a given framework be denoted by $\mathtt{Sc}(\mathcal{AF})$, $\mathtt{Cr}(\mathcal{AF})$ and $\mathtt{Rej}(\mathcal{AF})$. It is easy to see that those three sets are disjunct while their union is set $\mathcal{A}$.

From the status of arguments, a status is assigned to each option of the set $\mathcal{O}$. Four disjoint cases are distinguished. An option may be:

- *acceptable* if it is supported by at least one sceptically accepted argument,
- *negotiable* if it has no sceptically accepted arguments, but it is supported by at least one credulously accepted argument,
- *non-supported* if it is not supported at all by arguments,
- *rejected* if it has arguments but all of them are rejected.

Definition 5.2.3 (Status of options)**.** *Let* $\mathcal{AF} = (\mathcal{O}, \mathcal{A}, \mathtt{Def}, \mathcal{H})$ *be a decision framework and* $o \in \mathcal{O}$*.*

- o *is* acceptable *iff* $\exists a \in \mathcal{H}(o)$ *such that* $a \in \texttt{Sc}(\mathcal{AF})$.
- o *is* negotiable *iff* $\nexists a \in \mathcal{H}(o)$ *s.t* $a \in \texttt{Sc}(\mathcal{AF})$ *and* $\exists a' \in \mathcal{H}(o)$ *s.t.* $a' \in \texttt{Cr}(\mathcal{AF})$.
- o *is* non-supported *iff* $\mathcal{H}(o) = \emptyset$.
- o *is* rejected *iff* $\mathcal{H}(o) \neq \emptyset$ *and* $\forall a \in \mathcal{H}(o)$, $a \in \texttt{Rej}(\mathcal{AF})$.

Let us denote by $\mathcal{O}_x(\mathcal{AF})$ the set of all options of a framework $\mathcal{AF}$ having status x, where $x \in \{a, n, ns, r\}$, and a stands for acceptable, n for negotiable, ns for non-supported and r for rejected. For example, $\mathcal{O}_a(\mathcal{AF})$ is the set of acceptable options of the framework $\mathcal{AF}$.

Example 5.2.4. *In Example 5.2.1, option o_1 is acceptable, o_2 is rejected and o_3 is non-supported under stable, preferred and grounded semantics.*

It can be checked that an option has exactly one status. This status may change in light of new arguments as we will see in the next section. The following property compares the sets of acceptable options under different semantics. As expected, since the empty set is an admissible extension for any argumentation framework, then there are no acceptable options under this semantics. Consequently, this semantics is not interesting in our application. We will not study stable semantics neither, since stable extensions do not always exist. Thus, in the rest of the chapter, we will concentrate on grounded and preferred semantics.

The last output of the decision framework proposed by Amgoud et al. (2008) is a total pre-ordering on the set $\mathcal{O}$. Indeed, it has been argued that an acceptable option is preferred to any negotiable option. A negotiable option is preferred to a non-supported one, which is itself preferred to a rejected option.

5.3 Revising decision frameworks

In the previous section, we introduced an argumentation-based decision making framework. The goal of this section is to study how the ordering on options changes in light of a new argument, and what is the impact of a new argument on the ordering without having to re-compute this latter. This issue is very important, especially in negotiation dialogues in which agents use argument-based decision making models for rank-ordering the possible values of the negotiation object, and for generating and evaluating arguments. From a strategical point of view, it is important for an agent to know what will be the impact of a given argument on the ordering of the receiving agent. This avoids sending useless arguments.

We assume that the new argument concerns an option. This means that new information about an option is received. Moreover, the original set of options remains the same. Thus, the new argument is about an existing option. We investigate under which conditions this option changes its status, and under which conditions the new argument does not influence neither positively nor negatively the quality of this option. Similarly, we investigate the impact of the new argument on the status of other arguments. For that purpose, we study how the acceptability of arguments evolves when the decision framework is extended by new arguments. We particularly focus on the sceptical grounded semantics, and the credulous preferred semantics.

Let $\mathcal{AF} = (\mathcal{O}, \mathcal{A}_b \cup \mathcal{A}_o, \mathtt{Def}_b \cup \mathtt{Def}_o \cup \mathtt{Def}_m, \mathcal{H})$ be a decision framework. Recall that function $\mathcal{H}$ relates options of $\mathcal{O}$ with the arguments that support them (i.e. $\mathcal{H} : \mathcal{O} \rightarrow 2^{\mathcal{A}_o}$).

Assume that a new argument, denoted e, is received (for instance, from another agent). Thus, the decision framework $\mathcal{AF}$ is extended by this argument and by new defeats. Let $\mathcal{AF} \oplus e = (\mathcal{O}', \mathcal{A}', \mathtt{Def}', \mathcal{H}')$ denote the new framework. It is clear that when $e \in \mathcal{A}$, then $\mathcal{O}' = \mathcal{O}$, $\mathcal{A}' = \mathcal{A}$, $\mathtt{Def}' = \mathtt{Def}_b \cup \mathtt{Def}_o \cup \mathtt{Def}_m$ and $\mathcal{H}' = \mathcal{H}$. The more interesting case is when $e \notin \mathcal{A}$, thus $\mathcal{A}' = \mathcal{A} \cup \{e\}$. In this paper, we assume that the argument e is practical, meaning that it supports an option. Moreover, we assume this option is already in the set $\mathcal{O}$. Thus, $\mathcal{O}' = \mathcal{O}$ and $\exists o \in \mathcal{O}$ such that $\mathtt{Conc}(e) = o$.

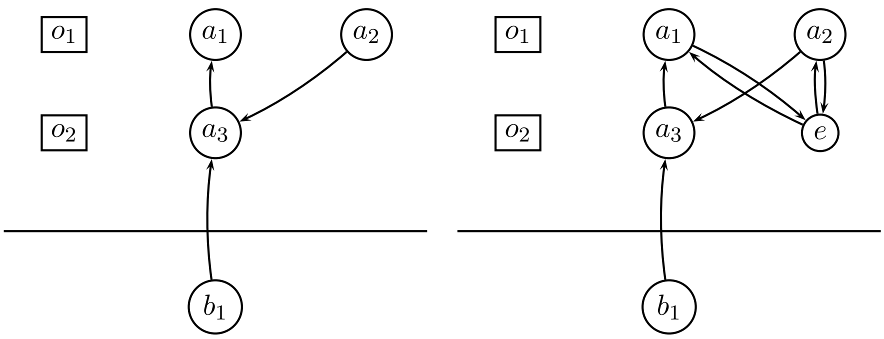
Regarding the relation $\mathtt{Def}'$, it contains all the elements of $\mathtt{Def}$, all the defeats between e and the arguments of $\mathcal{A}_o$ that support other options than $\mathtt{Conc}(e)$, and all the defeats emanating from epistemic arguments in $\mathcal{A}_b$ towards the argument e. Recall that a practical argument is not allowed to attack an epistemic one. The question now is how to recognize an attack from an epistemic argument towards e? This is done by checking the formal definition of the attack relation that is used. For instance, if $\mathcal{R}_m$ is defined as undercut, then an argument $x \in \mathcal{A}_b$ attacks e if the conclusion of x undermines a premise in e. For our purpose, we assume that $\mathcal{R}_m^{\mathcal{L}}$ contains all the conflicts that may exist between all the epistemic arguments and the practical arguments that may be built from the logical language $\mathcal{L}$. Thus, $\mathcal{R}'_m = \mathcal{R}_m^{\mathcal{L}}|_{\mathcal{A}'}$.

Defeats between practical arguments of $\mathcal{A}_o$ and the new argument e are based on i) the conflicts between arguments, and these capture the idea that two arguments support different options, and ii) a preference relation between the arguments. The new argument needs then to be compared to the other arguments of $\mathcal{A}_o$. The question is how this can be done? Here again by applying the formal definition of the preference relation that is used in the decision framework. For instance, if $\geq_o$ privileges the argument that is based on the most certain information and the most important benefit, then the new argument e is compared to any argument in $\mathcal{A}_o$ using these criteria (which should, of course, be defined precisely in each concrete application). At an abstract level, we assume that this is captured by a new preference relation, denoted by $\geq'_o$, on the set $\mathcal{A}'_o$. The definition of $\texttt{Def}'$ of the extended framework $\mathcal{AF} \oplus e$ is summarised below.

$$\texttt{Def}' = \texttt{Def} \cup \{(x, e) \mid x \in \mathcal{A}_b \text{ and } (x, e) \in \mathcal{R}_m^{\mathcal{L}}\} \cup$$
$$\{(e, y) \mid y \in \mathcal{A}_o \text{ and } \texttt{Conc}(y) \neq \texttt{Conc}(e) \text{ and } (y, e) \notin >'_o\} \cup$$
$$\{(y, e) \mid y \in \mathcal{A}_o \text{ and } \texttt{Conc}(y) \neq \texttt{Conc}(e) \text{ and } (e, y) \notin >'_o\}.$$

Extending a decision framework by a new argument may have an impact on the output of the original framework, namely on the status of the arguments, the status of options, and on the ordering on options. This is illustrated by the following example.

Figure 5.2: Decision framework before and after the new argument e arrives.



Example 5.3.1. *Let* $\mathcal{AF} = (\mathcal{O}, \mathcal{A}_b \cup \mathcal{A}_o, \texttt{Def}_b \cup \texttt{Def}_o \cup \texttt{Def}_m, \mathcal{H})$ *be a decision framework such that* $\mathcal{O} = \{o_1, o_2\}$, $\mathcal{A}_o = \{a_1, a_2, a_3\}$, $\mathcal{A}_b = \{b_1\}$, $\mathcal{H}(o_1) = \{a_1, a_2\}$, $\mathcal{H}(o_2) = \{a_3\}$, $\mathcal{R}_b = \emptyset$, *and* $\mathcal{R}_m = \{(b_1, a_3)\}$. *Assume that* $(a_3, a_1) \in >_o$ *and* $(a_2, a_3) \in >_o$. *The graph of defeat is depicted on*

the left side of Figure 5.2. It can be checked that the grounded extension of this framework is $\texttt{GE} = \{a_1, a_2, b_1\}$*. Thus,* $\texttt{Sc}(\mathcal{AF}) = \{a_1, a_2, b_1\}$ *and* $\texttt{Rej}(\mathcal{AF}) = \{a_3\}$*. Consequently, the option* o_1 *is acceptable while* o_2 *is rejected, and* o_1 *is strictly preferred to* o_2*.*

Assume now that the framework is extended by a new practical argument e *in favor of option* o_2 *(i.e.* $\texttt{Conc}(e) = o_2$*), and that this argument is incomparable with the other practical arguments. The new graph of defeat is depicted on the right side of Figure 5.2. The grounded extension of the extended framework is* $\texttt{GE} = \{b_1\}$*. Thus,* $\texttt{Sc}(\mathcal{AF} \oplus e) = \{b_1\}$ *and* $\texttt{Rej}(\mathcal{AF} \oplus e) = \{a_1, a_2, a_3, e\}$*. Consequently, the two options* o_1 *and* o_2 *are rejected, and are thus equally preferred.*

The aim of this section is to study the impact of a new practical argument e on the result of a decision framework. We first study under which conditions statuses of existing arguments change. Then, we show when an option changes its status in the new framework.

5.3.1 Revision under grounded semantics

Let $\mathcal{AF} = (\mathcal{O}, \mathcal{A}_b \cup \mathcal{A}_o, \texttt{Def}_b \cup \texttt{Def}_o \cup \texttt{Def}_m, \mathcal{H})$ be a decision framework, and $\mathcal{AF} \oplus e = (\mathcal{O}, \mathcal{A}_b \cup \mathcal{A}_o \cup \{e\}, \texttt{Def}_b \cup \texttt{Def}'_o \cup \texttt{Def}'_m, \mathcal{H}')$ its extension by a practical argument e. In this subsection, we assume that arguments in both $\mathcal{AF}$ and $\mathcal{AF} \oplus e$ are evaluated under grounded semantics. In this case, an argument is either sceptically accepted or rejected. The set of credulously accepted arguments is empty since there exists exactly one extension under this semantics. Consequently, an option may be non-supported, acceptable or rejected (i.e. there are no negotiable options).

Notation: The grounded extensions of a given framework will be denoted by $\texttt{GE}(\mathcal{AF})$, or $\texttt{GE}$ if there is no risk of confusion. Recall also that a characteristic function of an argumentation framework is denoted by $\mathfrak{F}$ (see Definition 2.2.10). For decision framework $\mathcal{AF}$, we will define $\texttt{Sc}^i(\mathcal{AF}) = \underbrace{\mathfrak{F}(\mathfrak{F}(\ldots\mathfrak{F}}_{i \text{ times}}(\emptyset))\ldots)$. (This notation is mostly used in proofs.)

The following property shows that a new practical argument will never influence the status of existing epistemic arguments. This means that the status of any epistemic argument in the framework $\mathcal{AF}$ remains the same in $\mathcal{AF} \oplus e$. This is mainly due to the fact that practical arguments are not allowed to attack epistemic ones. Recall that $\texttt{Status}(a, \mathcal{AF})$ be the function that returns the status of an argument a in the decision framework $\mathcal{AF}$.

Proposition 5.3.1. Let $\mathcal{AF}$ be a decision framework. For all $a \in \mathcal{A}_b$, $\texttt{Status}(a, \mathcal{AF}) = \texttt{Status}(a, \mathcal{AF} \oplus e)$.

Example 5.3.1 shows that this result is not always true for the practical arguments of the set $\mathcal{A}_o$. However, it holds in case the new argument is defeated by a sceptically accepted epistemic argument. In this case, the argument e has clearly no impact on the results of the original framework $\mathcal{AF}$.

Proposition 5.3.2. Let $\mathcal{AF}$ be a decision framework. If $\exists a \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF})$ such that $(a, e) \in \texttt{Def}'_m$, then

- $e \in \texttt{Rej}(\mathcal{AF} \oplus e)$,
- $\texttt{GE}(\mathcal{AF}) = \texttt{GE}(\mathcal{AF} \oplus e)$
- for all $a \in \mathcal{A}_o$, $\texttt{Status}(a, \mathcal{AF}) = \texttt{Status}(a, \mathcal{AF} \oplus e)$.

In case the new argument e is not defeated by an accepted epistemic argument, we show that the status of practical arguments in $\mathcal{A}_o$ which are in favor of $\texttt{Conc}(e)$ may either be the same as in the original framework or improved, moving thus from a rejection to an acceptance. However, things are different with the practical arguments that support other options than $\texttt{Conc}(e)$. Indeed, the status of these arguments may either remain the same or be worsened. This means that the new argument can *improve* only the status of the other arguments supporting its own option.

Proposition 5.3.3. Let $\mathcal{AF}$ be a decision framework.

- For all $a \in \mathcal{H}(\texttt{Conc}(e))$, if $a \in \texttt{Sc}(\mathcal{AF})$ then $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$.
- For all $a \in \mathcal{A}_o$, if $a \in \texttt{Rej}(\mathcal{AF})$ and $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$, then $e \in \mathcal{H}(\texttt{Conc}(a))$.

The result proved in the previous proposition can be summarised as follows. Let $a \in \mathcal{H}(o)$ and $a' \in \mathcal{H}(o')$ with $o \neq o'$. Symbol $\times$ means that the status of the argument does not change in the new framework, symbol $-$ denotes the fact the argument moves from an acceptance to a rejection, while $+$ means that the status of the argument is improved (i.e. the argument moves from a rejection to an acceptance).

There are four possible situations (corresponding to the four columns of the table). In the first situation, both the argument supporting $\texttt{Conc}(e)$ and that supporting the other option keep their original status. In the second situation, the argument in favor of $\texttt{Conc}(e)$ does not change its status while the argument supporting the other option is weakened. In the two remaining situations, the argument in favor of $\texttt{Conc}(e)$ improve its status while the argument supporting the other options either does not change its status or is weakened.

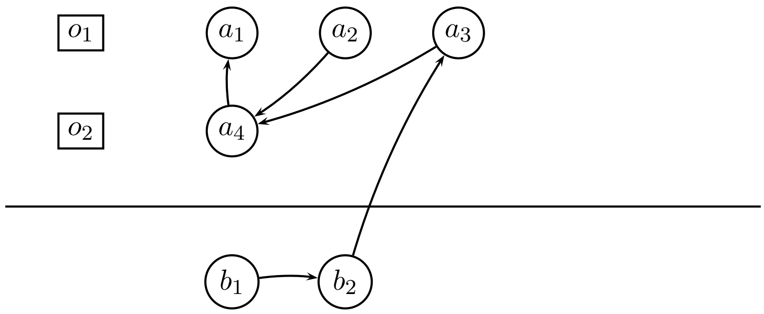
	1	2	3	4
$a \in \mathcal{A}_o$ s.t. $\texttt{Conc}(a) = \texttt{Conc}(e)$	$\times$	$\times$	$+$	$+$
$a' \in \mathcal{A}_o$ s.t. $\texttt{Conc}(a') \neq \texttt{Conc}(e)$	$\times$	$-$	$\times$	$-$

Recall that our main goal is to show under which conditions an option changes its status. To characterize that situations, we will need the following notion.

Definition 5.3.1. *Let $\mathcal{AF} = (\mathcal{O}, \mathcal{A}, \texttt{Def}, \mathcal{H})$ be a decision framework and let $a \in \mathcal{A}$. We say that a is defended by epistemic arguments in $\mathcal{AF}$, and we write $a \in \texttt{Dbe}(\mathcal{AF})$, iff $\forall x \in \mathcal{A}$, if $(x, a) \in \texttt{Def}$ then $\exists b \in \texttt{Sc}(\mathcal{AF}) \cap \mathcal{A}_b$ such that $(b, x) \in \texttt{Def}$.*

Example 5.3.2. *Let $\mathcal{AF}$ be a decision framework such that $\mathcal{O} = \{o_1, o_2\}$, $\mathcal{A}_b = \{b_1, b_2\}$, $\mathcal{A}_o = \{a_1, a_2, a_3, a_4\}$, $\mathcal{H}(o_1) = \{a_1, a_2, a_3\}$ and $\mathcal{H}(o_2) = \{a_4\}$. The defeat relation $\texttt{Def}$ are depicted in Figure 5.3.*

Figure 5.3: Defense by epistemic arguments



The grounded extension of this framework is $\texttt{GE} = \{a_1, a_2, a_3, b_1\}$. It can be checked that $\texttt{Dbe}(\mathcal{AF}) = \{b_1, a_2, a_3\}$. Note that $a_1 \notin \texttt{Dbe}(\mathcal{AF})$ even if it is indirectly defended by argument b_1. In fact, definition of $\texttt{Dbe}$ uses only direct defense.

It is worth noticing that non-defeated arguments are (trivially) defended by epistemic arguments.

Let us now come back to the status of options. Recall that, under grounded semantics, an option may be acceptable, rejected or non-supported. We are interested in i) the case where an option is rejected in the framework $\mathcal{AF}$ and becomes acceptable in $\mathcal{AF} \oplus e$, and ii) the case where an option is acceptable in $\mathcal{AF}$ and becomes rejected in $\mathcal{AF} \oplus e$. From the previous results, it is clear that the first case holds only for the option that is supported by the new argument. Indeed, the new argument may improve the status of its own conclusion. However, it never improves the status of the other options in the framework. This is formally shown by the following result. (To see that if e supports a given option and is sceptically accepted then the option becomes accepted is trivial, however, the other part of theorem is not trivial, even if it may seem so at the first sight.)

Proposition 5.3.4. Let $\mathcal{AF}$ be a decision framework and $o \in \mathcal{O}_r(\mathcal{AF})$. It holds that $o \in \mathcal{O}_a(\mathcal{AF} \oplus e)$ iff $e \in \mathcal{H}(o)$ and $e \in \mathtt{Sc}(\mathcal{AF} \oplus e)$.

Note that the above result depends on the status of the new argument in the extended framework. This is why we provide the following result which characterizes when this argument is sceptically accepted in $\mathcal{AF} \oplus e$ without computing the grounded extension of this framework. We show that the new argument is accepted iff for every attack from an argument $x \in \mathcal{A}_b \cup \mathcal{A}_o$ to e, there exists an argument which either supports $\mathtt{Conc}(e)$ or is epistemic, which defeats x, and which is in the grounded extension of the original framework.

Proposition 5.3.5. Let $\mathcal{AF} = (\mathcal{O}, \mathcal{A}, \mathtt{Def}, \mathcal{H})$ be a decision framework, and $\mathcal{AF}$ its extension with argument e. It holds that $e \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ iff for all $a \in \mathcal{A}$, if $(a, e) \in \mathtt{Def}'$, then $\exists b \in \mathtt{Sc}(\mathcal{AF}) \cap (\mathcal{A}_b \cup \mathcal{H}(\mathtt{Conc}(e)))$ s.t. $(b, a) \in \mathtt{Def}$.

Let us now analyze the case where an option is acceptable in $\mathcal{AF}$ and becomes rejected in $\mathcal{AF} \oplus e$. This case concerns only the options that are not supported by the new argument e. Indeed, since practical arguments supporting other options than $\mathtt{Conc}(e)$ may be weakened by the new argument, their conclusions may be weakened as well. The following result shows the conditions under which this is possible.

Proposition 5.3.6. Let $\mathcal{AF}$ be a decision framework and $o \in \mathcal{O}_a(\mathcal{AF})$. It holds that $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$ iff

1. $e \notin \mathcal{H}(o)$, and

2. $\nexists a \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF})$ s.t. $(a, e) \in \texttt{Def}'_m$, and

3. $\forall a \in \texttt{Dbe}(\mathcal{AF}) \cap \mathcal{H}(o)$, $(e, a) \in \texttt{Def}'_o$.

The first condition says that the new argument does not support the rejected option. The second condition states that the new argument is not defeated by an epistemic argument which is accepted in the original framework $\mathcal{AF}$. This is important because otherwise the new argument is rejected in $\mathcal{AF} \oplus e$ and has no impact on the result. The last condition says that all the practical arguments supporting the option in question which are defended by epistemic arguments are defeated by the new argument.

5.3.2 Revision under preferred semantics

In this subsection, the arguments of a decision framework $\mathcal{AF}$ and those of its extension $\mathcal{AF} \oplus e$ are evaluated under preferred semantics. Thus, an argument may be either sceptically accepted, credulously accepted or rejected. Consequently, an option may have one of the corresponding statuses: acceptable, negotiable, rejected or non-supported.

Like in the case of grounded semantics, epistemic arguments will not change their status when a new practical argument is received. This shows that the system is protected against wishful thinking.

Proposition 5.3.7. Let $\mathcal{AF}$ be a decision framework. For all $a \in \mathcal{A}_b$, $\texttt{Status}(a, \mathcal{AF}) = \texttt{Status}(a, \mathcal{AF} \oplus e)$.

We now prove that if the new practical argument is attacked by a sceptically accepted epistemic argument in $\mathcal{AF}$, then the preferred extensions of $\mathcal{AF}$ and $\mathcal{AF} \oplus e$ coincide. As a consequence, all the existing arguments keep their status. Moreover, the new argument e is rejected. This means that such an argument does not influence the output of the decision framework.

Proposition 5.3.8. Let $\mathcal{AF}$ be a decision framework. If $\exists a \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF})$ such that $(a, e) \in \texttt{Def}'_m$, then

- $e \in \texttt{Rej}(\mathcal{AF} \oplus e)$,
- $\forall \mathcal{E} \subseteq \mathcal{A}$, $\mathcal{E}$ is a preferred extension of $\mathcal{AF}$ iff $\mathcal{E}$ is a preferred extension of $\mathcal{AF} \oplus e$,
- for all $a \in \mathcal{A}_o$, $\texttt{Status}(\mathcal{AF}, a) = \texttt{Status}(\mathcal{AF} \oplus e, a)$.

Like in the case of grounded semantics, the status of the arguments supporting $\texttt{Conc}(e)$ in $\mathcal{AF}$ can be improved but never weakened in $\mathcal{AF} \oplus e$.

Proposition 5.3.9. Let $\mathcal{AF}$ be a decision framework. For all $a \in \mathcal{A}_o$ such that $\texttt{Conc}(a) = \texttt{Conc}(e)$, it holds that:

- If $a \in \texttt{Sc}(\mathcal{AF})$ then $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$
- If $a \in \texttt{Cr}(\mathcal{AF})$ then $a \in \texttt{Sc}(\mathcal{AF} \oplus e) \cup \texttt{Cr}(\mathcal{AF} \oplus e)$

On the other hand, as in the case of grounded semantics, an argument supporting an option different that $\texttt{Conc}(e)$ is never improved in $\mathcal{AF} \oplus e$.

Proposition 5.3.10. Let $\mathcal{AF}$ be a decision framework, and $a \in \mathcal{A}_o$. If $a \in \texttt{Rej}(\mathcal{AF})$ and $a \in \texttt{Sc}(\mathcal{AF} \oplus e) \cup \texttt{Cr}(\mathcal{AF} \oplus e)$ then $\texttt{Conc}(a) = \texttt{Conc}(e)$.

Using the above results on the status of arguments, we can show under which conditions a given option may change its status in the extended decision framework $\mathcal{AF} \oplus e$. We have seen that the quality of the arguments of $\mathcal{A}_o$ that support $\texttt{Conc}(e)$ may be improved. Thus, it is expected that the status of $\texttt{Conc}(e)$ may be improved as well. The following result shows, in particular, when $\texttt{Conc}(e)$ moves from a rejection to a better status (i.e. becomes either negotiable or acceptable).

Proposition 5.3.11. Let $\mathcal{AF}$ be a decision framework and $o \in \mathcal{O}_r(\mathcal{AF})$. Then $o \in \mathcal{O}_a(\mathcal{AF} \oplus e) \cup \mathcal{O}_n(\mathcal{AF} \oplus e)$ iff $e \in \mathcal{H}(o) \land e \notin \texttt{Rej}(\mathcal{AF} \oplus e)$.

Since in the previous result the status of the new argument in $\mathcal{AF} \oplus e$ is used, we provide the characterization of its status, based only on information from $\mathcal{AF}$.

Proposition 5.3.12. Let $\mathcal{AF} = (\mathcal{O}, \mathcal{A}_b \cup \mathcal{A}_o, \texttt{Def}_b \cup \texttt{Def}_o \cup \texttt{Def}_m, \mathcal{H})$ be a decision framework. It holds that $e \notin \texttt{Rej}(\mathcal{AF} \oplus e)$ iff $\exists \mathcal{E} \subseteq \mathcal{A}_b$ and $\exists \mathcal{E}' \subseteq \mathcal{H}(\texttt{Conc}(e))$ such that:

1. $\mathcal{E} \cup \mathcal{E}'$ is conflict-free, and
2. $\mathcal{E}$ is a preferred extension of $(\mathcal{A}_b, \texttt{Def}_b)$, and
3. $\forall a \in \mathcal{E}' \cup \{e\}$, if $\exists x \in \mathcal{A}$ s.t. $(x, a) \in \texttt{Def}$, then $\exists a' \in \mathcal{E} \cup \mathcal{E}' \cup \{e\}$ s.t. $(a', x) \in \texttt{Def}$.

The following result summarises under which conditions an option may become rejected in the extended decision framework. The first condition says that for an option to become rejected, it should not be supported by the new argument. The second condition says that the new argument should not be attacked by an epistemic argument which is in a preferred extension that contains arguments in favor of this option. The last condition claims that the new argument should be preferred to some arguments in favor of the option.

Proposition 5.3.13. Let $\mathcal{AF}$ be a decision framework and $o \in \mathcal{O}_a(\mathcal{AF}) \cup \mathcal{O}_n(\mathcal{AF})$. Then $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$ iff

1. $e \notin \mathcal{H}(o)$, and
2. there does not exist a preferred extension $\mathcal{E}$ of $\mathcal{AF}$ s.t. $\mathcal{E} \cap \mathcal{H}(o) \neq \emptyset$ and $\exists a \in \mathcal{E} \cap \mathcal{A}_b$ s.t. $(a, e) \in \mathtt{Def}'_m$, and
3. there does not exist a preferred extension $\mathcal{E}$ of $\mathcal{AF}$ s.t. there exists an admissible set $\mathcal{E}''$ of $\mathcal{AF}$ with $\mathcal{E}'' \cap \mathcal{A}_o \subseteq \mathcal{E} \cap \mathcal{H}(o)$ and $\mathcal{E}'' \cap \mathcal{A}_b = \mathcal{E} \cap \mathcal{A}_b$ and $\forall a \in \mathcal{E}'' \cap \mathcal{H}(o), (a, e) \in \succ'_o$ or $\exists a' \in \mathcal{E}'' \cap \mathcal{H}(o)$ s.t. $(e, a) \notin \succ'_o$.

5.4 Conclusion

In this chapter, we showed what a decision making process is, and how argumentation may be used for decision. We mainly focused on the dynamics of argumentation-based decision frameworks.

We conducted the first investigation of the impact of a new argument on an argumentation-based decision framework. We used the decision framework proposed by Amgoud, Dimopoulos, and Moraitis (2008) with slight modifications. (The modifications are related to the fact that we use our preference-based argumentation framework which inverts the arrows of critical attacks instead of deleting them.)

We provided a full characterization of acceptable options that become rejected, and of rejected options that become acceptable in the extended framework. A characterization of the evolution of the status of arguments is also provided. Our study is undertaken under two acceptability semantics: grounded semantics and preferred one.

These results may be used in negotiation dialogues, namely to determine strategies. Indeed, at a given step of a dialog, an agent may choose which argument to send to another agent in order to change the status of an option. Our results may help to understand which arguments are useful and which ones are useless in a given situation.

J'ai entendu vos points de vue. Ils ne rencontrent pas les miens. La décision est prise à l'unanimité.

Charles de Gaulle

6 Argumentation-based negotiation

This chapter studies argument-based negotiation. After presenting the most influential works in the literature, we focus our attention on the impact of argumentation on the quality of negotiation outcome and answer the question: *when does argumentation improve the quality of negotiation outcomes?* (Amgoud and Vesic, 2011a,b).

6.1 Introduction

Negotiation is one of the most common approaches used to make decisions and manage disputes. It occurs between parents and children, managers and staff, employers and employees, professionals and clients, within and between organizations and between agencies and the public.

Negotiation is a process that aims at finding some *compromise* or *consensus* on an issue between two or more agents having different *goals*. In the negotiation literature, the issue under negotiation is called the *negotiation object*. Examples of negotiation objects are: the price of a given product, the date and/or the place of a meeting and so on. In the seminal book by Walton and Krabbe Walton and Krabbe (1995), the object concerns the share of some goods or services.

6.2 Main approaches to negotiation

A huge amount of work was done for modeling negotiation. Negotiation techniques are often separated in the three classes: *game-theoretic* approaches, *heuristic-based* approaches and *argumentation-based* approaches. This classification was proposed by Jennings et al. (2001), and later adopted by other researches (e.g. Rahwan et al., 2004).

6.2.1 Game-theoretic approaches

Game theory models strategic situations, or games, in which an individual's success in making choices depends on the choices of others (Myerson, 1997). It has its roots in the work of von Neumann and Morgenstern (1944). It has been used by many researches to study the interaction between self-interested negotiating agents (e.g. Rosenschein and Zlotkin, 1994). In game-theoretic analysis of negotiation, emphasis is on determining the optimal strategy. This is done by analyzing data and formalizing negotiation as a *game* between participants. Those approaches are often based on a set of formal hypotheses (like game rules, payoffs corresponding to different situations, goals of negotiating agents). This allows to prove that a given strategy is (or is not) the optimal one for a participant. It is also assumed that the participants are rational, in the sense that they make decisions which are in accordance with their knowledge and their goals. This guarantees that negotiating parties behave in certain ways (Varian, 1995). However, classical game-theoretic approaches have some significant limitations from the computational perspective (Dash, Jennings, and Parkes, 2003). Specifically, most of these approaches assume that agents have unbounded computational resources and that the space of outcomes is completely known. In most realistic environments, however, these assumptions fail due to the limited processing and communication capabilities of the information systems.

6.2.2 Heuristic-based approaches

As a response to the above limitations of game-theoretic approaches (mostly strong hypotheses about agent rationality and unbounded computational resources), a number of heuristic-based negotiation approaches have been developed. Heuristics are rules of thumb that produce good enough (rather than optimal) outcomes. On the contrary, those systems are more efficient and demand for less resources. Of course, the inconvenience is that every particular heuristic demands for empirical evaluation and adjustment of parameters (e.g. Faratin, 2000). When compared with game-theoretic approaches, these methods offer approximations.

6.2.3 Argumentation-based approaches

Although game theoretic and heuristic based approaches both have desirable features and are widely studied by researches, they share some limitations. In most game-theoretic and heuristic models, agents exchange proposals (i.e. potential agreements or potential deals). This, for example, can be a promise to purchase a given object at a specified price. However, agents are not allowed to exchange any additional information other than what is expressed in the proposal itself. This can be problematic, for example, in situations where agents have limited information about the environment, or where their rational choices depend on those of other agents. Another limitation of conventional approaches to automated negotiation is that the agent's preference relation on the set of offers is supposed to stay fixed during the interaction.

Consider the following example.

Two professors, say Pr_1 and Pr_2, want to employ a new research assistant on a European project. Three candidates, Carla, John and Mary are interested in the position. Unfortunately, the two professors have conflicting preferences. Professor Pr_1 prefers Carla to John and John to Mary (i.e. Carla $\succeq^1$ John $\succeq^1$ Mary). However, professor Pr_2 prefers John to Carla and Carla to Mary (i.e. John $\succeq^2$ Carla $\succeq^2$ Mary). The following negotiation may take place between the two agents:

Pr_1: I suggest to recruit Carla

Pr_2: No, I prefer John.

If we suppose that the only object of negotiation is the candidate, then further negotiation is hard, since there is no obvious way to make a compromise. The idea of argumentation-based negotiation (ABN) is to allow negotiating parties to exchange *arguments* which contain information that can change other agents' beliefs and, consequently, his/her preferences on the set of options. The previous dialogue can continue if arguments are exchanged. For example:

Pr_1: I suggest to recruit Carla

Pr_2: No, I prefer John. He is working on my research topic.

Pr_1: But, you know that Carla has a better publication record than John. Moreover, recently she did a very interesting work on your topic.

Pr_2: Really, I didn't know that. So let's give her the position then.

In this dialogue, Pr_2 received a strong argument in favor of Carla which leads him to change his preference between John and Carla.

In other words, since game-theoretic and heuristic approaches assume that agents' preferences on the set of options are fixed, the only direction of negotiation is finding some compromise. One agent cannot directly influence another agent's preference model, or any of its internal mental attitudes (e.g. beliefs, desires, goals, etc.) that generate its preference model. A rational agent would only modify its preferences upon receipt of new information. Traditional automated negotiation mechanisms do not facilitate the exchange of this information. That is why exchange of arguments can be beneficial for negotiation. In the context of negotiation, an argument is intended to influence another agent in complex ways. Thus, in addition to accepting a proposal, rejecting it, or proposing another possible deal, an agent can justify his/her choice and/or criticise another agent's offer and/or arguments. By understanding why its counterpart cannot accept a particular deal, an agent may be in a better position to make an alternative offer that has a higher chance of being acceptable.

Sycara (1990) was among the first to emphasize the importance of using argumentation techniques in negotiation. Since then, several works were done including those by Parsons and Jennings (1996), Reed (1998), Kraus, Sycara, and Evenchik (1998), Tohmé (1997), Amgoud, Dimopoulos, and Moraitis (2007), Amgoud, Parsons, and Maudet (2000b), Amgoud and Prade (2004), or Kakas and Moraitis (2006).

6.3 A formal analysis of the role of argumentation in negotiation dialogues

As said before, several proposals were made in the literature for modeling argumentation-based negotiation. Most of them were interested in proposing protocols which show how arguments and offers can be generated, evaluated and exchanged in a negotiation dialogue. Unfortunately, except the termination of each dialogue generated under those protocols, nothing is said on their *quality*. In particular, it is not clear what kind of *solutions* (or *outcomes*) are reached by their dialogues. The first reason is that the notion of optimal solution is not defined for argument-based negotiations. Indeed, there is no study on the types of outcomes that may be reached in such negotiations. It is also worth mentioning that before the work done by Amgoud, Dimopoulos, and Moraitis (2007), it was not formally shown that new arguments may influence the preferences of an agent. In that paper, each agent is equipped with a *theory* which is an argumentation-based decision making system that computes a preference relation on the set of offers. It was shown that the *theory* of an agent may evolve when new arguments are received, and consequently the initial preference relation may change.

However, it is not clear how this evolution of agents' theories may have an impact on the outcome of a negotiation. In other words, when is this theories' evolution beneficial for a negotiation and for the agents?

The goal of the rest of the chapter is twofold. It characterizes for the first time the possible outcomes of ABN dialogues. Different kinds of outcomes (solutions) are identified: accepted solution, optimal solution, local solution, Pareto optimal solution and ideal solution. Accepted, local and Pareto optimal solutions are the best outcomes at a given step of a dialogue while optimal and ideal solutions are the best solutions in general and are time-independent. The second contribution of this chapter consists in studying to what extent and under which conditions, argumentation may be beneficial in a negotiation dialogue. We show that when an ideal solution exists, argumentation pushes negotiation towards this solution. Even when such a solution does not exist, arguing may be beneficial, since it can allow agents to make decisions under more information (i.e. less uncertainty). Our study is undertaken at an abstract level since we do not take into account protocols and strategical issues. Thus, our results are true under any protocol and using any strategy.

6.3.1 Negotiation framework

In a negotiation dialogue several agents may be involved. In what follows, in order to simplify notation, we restrict ourselves to the case of only two agents denoted by $Ag1$ and $Ag2$. However, it is easy to see that all the results can be expressed in the case of n agents. These agents are assumed to share some background in order to understand each other. They use the same logical language $\mathcal{L}$ and the same definition of an argument. Thus, both agents recognize any argument in the set of all arguments $\texttt{Arg}(\mathcal{L})$ (we denote by $\texttt{Arg}(\mathcal{L})$ the set of all arguments that can be constructed from the language $\mathcal{L}$). Similarly, we suppose that each agent recognizes any conflict in $\mathcal{R}(\mathcal{L})$ (the set of all attacks on $\texttt{Arg}(\mathcal{L})$) and that they use the same definition of attack, i.e. they use the same attack relation (for instance, they both use "undercut"). In addition, each negotiating agent is equipped with an argumentation-based decision making framework. For the purpose of this chapter, we will suppose that every agent has a decision framework as the one which was described in Section 5.2. This framework is used to build and evaluate practical and epistemic arguments (recall that those sets are denoted $\mathcal{A}_o$ and $\mathcal{A}_b$ and are supposed to be disjunct), to evaluate offers, to compare pairs of offers, and finally to select the best offer. We suppose that an offer is a possible value of the negotiation object and that elements of $\mathcal{O}$ represent the possible offers. We will use notation $\mathcal{O}(\mathcal{L})$ for the set of all offers that can be defined from a language $\mathcal{L}$ and $\mathcal{O}$ for a set of offers of a particular negotiation we will study.

Thus, the theory of agent i is $\texttt{AF}^i = (\mathcal{O}^i, \mathcal{A}^i, \mathcal{R}^i, \geq^i, \mathcal{H}^i)$ where: $\mathcal{O}^i$ is a *finite* subset of $\mathcal{O}(\mathcal{L})$, $\mathcal{A}^i$ the set of his/her arguments, $\mathcal{R}^i = \mathcal{R}(\mathcal{L})|_{\mathcal{A}^i}$ the attack relation between them, with $\mathcal{R}^i = \mathcal{R}^i_b \cup \mathcal{R}^i_o \cup \mathcal{R}^i_m$, $\geq^i$ the agent's preference relation with $\geq^i = \geq^i_b \cup \geq^i_o \cup \geq^i_m$, and $\mathcal{H}^i$ the function which relates options with practical arguments, as in the previous chapter. In what follows, we assume that agents have the same set of offers; furthermore, we suppose that it does not change during negotiation. We will use $\mathcal{O}$ to denote that set. However, the two agents may not necessarily have the same arguments in favor of an offer, thus, we will use two distinct functions, $\mathcal{H}^i(\mathcal{L}) : \texttt{Arg}_o(\mathcal{L}) \rightarrow \mathcal{O}(\mathcal{L})$, with $i \in \{1, 2\}$, which for every practical argument returns the offer it supports. This is because a practical argument e can be in favor of one offer for $Ag1$ and in favor of another one for $Ag2$. For a practical argument e, we will write $\texttt{Conc}^i(e) = o$ iff $o \in \mathcal{O}$ and $e \in \mathcal{H}^i(o)$ (i.e. if e is in favour of o for Agi).

We assume that exchanged arguments are not self-defeating. A similar study can be conducted without this hypothesis, which would not change much of the chapter. Moreover, we assume that when a new argument is added to the original set $\mathcal{A}$ of arguments, other arguments cannot be built using the information underlying the new argument and that underlying arguments of $\mathcal{A}$. This is since we are conducting an abstract and general study without entering in arguments' structure, thus, it is impossible to know which arguments could be generated using information from other arguments.

Note that by using a decision framework which has only arguments in favour of offers (and not against them), we are slightly restricting the generality of our approach. However, since the goal of this work is not to develop another argumentation-based decision making framework, we are not interested in adding arguments against the offers (this is left for future work).

Since we suppose that agents must keep their arguments in some sort of memory, then both $\mathcal{A}^1$ and $\mathcal{A}^2$ are supposed to be *finite* subsets of $\texttt{Arg}(\mathcal{L})$.

Note that the preference relation between arguments is expressed on the whole set $\texttt{Arg}(\mathcal{L})$. This means that an agent is *able* to express a preference between any pair of arguments. Formally, we suppose that $\geq^i (\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$ and $\geq^i = \geq^i (\mathcal{L})|_{\mathcal{A}^i}$.

6.3.1.1 Negotiation dialogues

In order to analyze the role of argumentation in negotiation, we need a minimal definition of a negotiation dialogue, that is a definition that sheds light on the basic elements that are exchanged during such a dialogue. In order to stay as general as possible, we do not focus on protocols; indeed, our definition can be extended by rules of any possible protocol. The basic element of a negotiation dialogue is the notion of move through which agents exchange offers of $\mathcal{O}$ and/or arguments of $\texttt{Arg}(\mathcal{L})$.

Definition 6.3.1 (Move)**.** *Let θ be a symbol that denotes that neither an argument nor an offer is sent. A* move *is a tuple $m = \langle p, a, o \rangle$ such that:*

1. $p \in \{Ag1, Ag2\}$,
2. $a \in \mathcal{A}^1 \cup \mathcal{A}^2 \cup \{\theta\}$,
3. $o \in \mathcal{O} \cup \{\theta\}$, *and*
4. $(a \neq \theta)$ *or* $(o \neq \theta)$.

The function Player *(resp.* Argument*,* Offer*) returns the player (resp. the argument, the offer) of the move. Let $\mathcal{M}$ be the set of all moves that can be built from $\langle \{Ag1, Ag2\}, \mathcal{A}^1 \cup \mathcal{A}^2, \mathcal{O} \rangle$.*

The fourth condition of the above definition states that at each step of the dialogue, an agent utters an argument, an offer or both. This means that the set of possible moves is finite.

A negotiation dialogue is a sequence of moves.

Definition 6.3.2 (Negotiation)**.** *A* negotiation dialogue *d between two agents $Ag1$ and $Ag2$ is a finite and non-empty sequence $\langle m_1, \ldots, m_l \rangle$ of moves. d is* argumentative *iff $\exists i \in \{1, \ldots, l\}$ s.t. $\texttt{Argument}(m_i) \neq \theta$. d is* non-argumentative *iff it is not argumentative.*

It is common in negotiation dialogues that agents propose less preferred offers in case their best options are all rejected by the other party. Such offers are called *concessions*. However, for the purpose of this study, we do not need to formally define this notion.

6.3.1.2 Impact of new arguments on an agent theory

So far, we have supposed that each agent has an abstract model for decision making. It takes as input a set of offers, a set of arguments (some of them supporting offers), an attack relation among arguments, a preference relation between arguments and a function which relates arguments with offers. The model computes a total preorder on the set of offers and thus, the best offer(s). We will now review all the possible changes of a negotiation framework.

Let $\mathtt{AF}_0^i$ be the *initial* theory of agent i, i.e. his/her theory before a dialogue $d = \langle m_1, \ldots, m_l \rangle$ starts. At each step t of d, a new theory $\mathtt{AF}_t^i$ is computed. Assume that $\mathtt{AF}_{t-1}^i = (\mathcal{O}_{t-1}^i, \mathcal{A}_{t-1}^i, \mathcal{R}_{t-1}^i, \geq_{t-1}^i, \mathcal{H}_{t-1}^i)$. If no argument is sent, then $\mathtt{AF}_t^i = \mathtt{AF}_{t-1}^i$. Else, let $e = \mathtt{Argument}(m_t)$. Then, $\mathtt{AF}_{t-1}^i = (\mathcal{O}_t^i, \mathcal{A}_t^i, \mathcal{R}_t^i, \geq_t^i, \mathcal{H}_t^i)$ is defined as:

- $\mathcal{O}_t^i = \mathcal{O}_{t-1}^i = \mathcal{O}$,
- $\mathcal{A}_t^i = \mathcal{A}_{t-1}^i \cup \{e\}$
- $\mathcal{R}_t^i = \mathcal{R}^i(\mathcal{L})|_{\mathcal{A}_t^i}$,
- $\geq_t^i = \geq^i (\mathcal{L})|_{\mathcal{A}_t^i}$,
- $\mathcal{H}_t^i = \mathcal{H}^i(\mathcal{L})|_{\mathcal{A}_t^i \cap \mathtt{Arg}_o(\mathcal{L})}$

Let us now describe different changes that may occur during negotiation. Recall that we supposed that the set of offers is fixed during the negotiation. However, let us note that a more general case (which is left for future work) is to consider even the case when the set $\mathcal{O}$ of offers can change during the negotiation.

Changing the set of options: By receiving a new argument, an agent may learn that there exists another option which is not considered in the set $\mathcal{O}$. Let us illustrate this case by a simple example.

Example 6.3.1. *Two agents are negotiating a price and date for a delivery of 500 kg of strawberries. Thus* $\mathcal{O} = \{(0.80€, July\ 13, 500kg), (0.90€, July\ 13, 500kg), (0.90€, July\ 14, 500kg), (1.00€, July\ 14, 500kg), (0.90€, July\ 15, 500kg), (1.00€, July\ 15, 500kg)\}$. *The negotiation process is not going very well. Then, the seller proposes to accept the delivery date and price per kg proposed by the buyer,* $(0.90€, July\ 14)$, *but under the condition to deliver 600 kg. Thus, the set of offers is enlarged by this new offer, and becomes* $\mathcal{O}' = \{(0.80€, July\ 13, 500kg), (0.90€, July\ 13, 500kg), (0.90€, July\ 14, 500kg), (1.00€, July\ 14, 500kg), (0.90€, July\ 15, 500kg), (1.00€, July\ 15, 500kg), (0.90€, July\ 14\ , 600kg\}$.

The following property characterizes a situation where the new offer becomes accepted by a receiving agent (the agent who receives the offer supported by an argument).

Proposition 6.3.1. Let $\texttt{AF}^i = (\mathcal{O}^i, \mathcal{A}^i, \mathcal{R}^i, \geq^i, \mathcal{H}^i)$ be the theory of agent i. Let $e \in \texttt{Arg}_o(\mathcal{L})$ be such that $\texttt{Conc}^i(e) \notin \mathcal{O}^i$. If $\forall e' \in \mathcal{A}^i \cap \texttt{Arg}_o(\mathcal{L})$, $e >^i e'$ and $\mathcal{R}^i_m = \emptyset$, then $\texttt{Conc}^i(e)$ will be acceptable (under preferred, grounded as well as under stable semantics if stable extensions exist) after this offer and argument have been received.

A new offer can be rejected in the extended theory if it is attacked by an epistemic argument which is sceptically accepted in the original theory.

Proposition 6.3.2. Let $\texttt{AF}^i = (\mathcal{O}^i, \mathcal{A}^i, \mathcal{R}^i, \geq^i, \mathcal{H}^i)$ be the theory of agent i. Let $e \in \texttt{Arg}_o(\mathcal{L})$ be such that $\texttt{Conc}^i(e) \notin \mathcal{O}^i$. If $\exists a \in \mathcal{A}^i \cap \texttt{Arg}_b(\mathcal{L})$ such that a is sceptically accepted in $\texttt{AF}^i$ and $(a, e) \in \mathcal{R}(\mathcal{L})$, then $\texttt{Conc}^i(e)$ is rejected (under preferred, grounded and stable semantics) after the new offer and argument has been received.

Note again that that this situation does not occur in our framework since we assumed that the two agents have the same set of offers which remain fixed during negotiation.

Changing the set of epistemic arguments: Receiving a new epistemic argument allows an agent to *revise his/her beliefs*. Consequently, the output of the theory may change.

Example 6.3.2. *Let* $\mathcal{O} = \{o_1, o_2\}$, $\mathcal{A}_b = \emptyset$, $\mathcal{H}(o_1) = \{e_1\}$, $\mathcal{H}(o_2) = \{e_2\}$ *and* $e_1 >_p e_2$. *This theory has one stable/preferred extension* $\mathcal{E} = \{e_1\}$. *Thus, option* o_1 *is acceptable while* o_2 *is rejected. Consequently,* $o_1 \succ o_2$. *Assume now that this agent receives an epistemic argument* a *such that* $a\mathcal{R}_m e_1$ *and* $\neg(a\mathcal{R}_m e_2)$. *The new theory has one stable/preferred/grounded extension which is* $\mathcal{E}' = \{a, e_2\}$. *Thus,* o_2 *is acceptable and* o_1 *is rejected; consequently,* $o_2 \succ o_1$.

Changing the set of practical arguments: A new practical argument may also have an impact on the outcome of a theory. If the new argument is not already in $\mathcal{A}^i \cap \texttt{Arg}_o(\mathcal{L})$, it induces a *revision* of agent's theory.

Example 6.3.3. *Let* $\mathcal{O} = \{o_1, o_2\}$, $\mathcal{A}_b = \emptyset$, $\mathcal{H}(o_1) = \{e_1\}$, $\mathcal{H}(o_2) = \emptyset$. *This theory has one stable/preferred/grounded extension* $\mathcal{E} = \{e_1\}$. *Thus, option*

o_1 is acceptable while o_2 is non-supported. Consequently, $o_1 \succ o_2$. Assume now that this agent receives a practical argument e_2 in favor of o_2 and $e_2 >_p e_1$. The new theory has one stable/preferred/grounded extension which is $\mathcal{E}' = \{e_2\}$. Thus, o_2 is acceptable and o_1 is rejected. This means that $o_2 \succ o_1$.

Recall that in the previous chapter, we have shown under which conditions an offer may move from acceptance to rejection and vice versa when a new piece of information arrives.

Changing the attack relation: When the set of arguments changes, the attack relation may change as well since new attacks may appear between the new argument and the existing ones. Note that a new argument never leads to a new attack between two existing arguments since all the possible attacks should already have been captured by the attack relation of the agent's theory.

Changing the preference relation between arguments: Recall that we supposed that preference relation is *static* and cannot change. For example, it is not possible for an agent to prefer a to b, and after receiving a new argument c, not prefer a to b. In order to allow the revision of preferences, we need a theory in which preferences are themselves subject to debate and are conclusions of arguments. An example of such model is the one proposed by Prakken and Sartor (1997).

It is easy to see that in the particular case of non-argumentative dialogues, the output of a theory does not change.

Proposition 6.3.3. Let $\succeq^i_0$ be the output of the theory of agent i before a dialogue. For every non-argumentative negotiation dialogue $d = \langle m_1, \ldots, m_l \rangle$, $\succeq^i_t = \succeq^i_0$, for any $t \in \{1, \ldots, l\}$.

This result confirms the intuition that non-argumentative approaches for negotiation (i.e. game-theoretic and heuristic-based approaches) do not model any change of the preorder on the set of offers. Allowing agents to exchange arguments which can influence them to change their beliefs and goals, is a step towards more realistic and more flexible negotiation frameworks.

6.3.2 Negotiation outcomes

In the previous subsection, we have seen that exchanging arguments allows for a rich and flexible negotiation framework. Intuitively, by exchanging arguments, negotiation can be ameliorated. However, it has never been formally shown that argumentation positively influences negotiation outcome. In order to do so, we first need to define a quality of a negotiation outcome.

In other words, we address the question: what is a "good" outcome in an ABN dialogue? In this subsection, we propose to define two categories of solutions: time-dependent solutions and global ones. Time-dependent solutions are the outcomes at a given step of a dialogue. Global solutions are defined without reference to a specific step of negotiation, i.e. they are time-independent. In what follows, we discuss each type of solution from an agent point of view and from a dialogue point of view.

6.3.2.1 Outcomes from agents perspective

From the point of view of a single agent, the best solutions at a given step of a dialogue are those having the best status (i.e. acceptable) at that step.

Definition 6.3.3 (Accepted solution for an agent)**.** *Let $d = \langle m_1, \ldots, m_l \rangle$ be a negotiation dialogue and $\mathtt{AF}_t^i$ the theory of agent i at step $t \leq l$. An offer $o \in \mathcal{O}$ is an* accepted solution *for agent i at the step t iff o is acceptable in $\mathtt{AF}_t^i$. We will use notation $\mathcal{O}_a(\mathtt{AF})$ for the set of acceptable offers w.r.t. agent theory $\mathtt{AF}$.*

The status of accepted solutions may change during a negotiation. Indeed, it may be the case that at step t, an offer is acceptable for an agent while it becomes rejected at step $t + 1$. Thus, such solutions are time-dependent. *Optimal solutions*, however, do not depend on a dialogue step. They are offers that an agent would choose if (s)he had access to all arguments owned by the other agent (or *agents* in a more general case). New arguments allow agents to revise their mental states; thus, the best decision for an agent is the one (s)he makes under 'complete' information (i.e. under minimal uncertainty).

Definition 6.3.4 (Optimal solution for an agent)**.** *Let $Ag1$ and $Ag2$ be two agents and $\mathtt{AF}^1 = (\mathcal{O}, \mathcal{A}_0^1, \mathcal{R}_0^1, \geq_0^1, \mathcal{H}_0^1)$ and $\mathtt{AF}^2 = (\mathcal{O}, \mathcal{A}_0^2, \mathcal{R}_0^2, \geq_0^2, \mathcal{H}_0^2)$ their initial theories. Let $\mathcal{A}^u = \mathcal{A}_0^1 \cup \mathcal{A}_0^2$. An offer $o \in \mathcal{O}$ is an* optimal solution *for agent i iff o is acceptable in $(\mathcal{O}, \mathcal{A}^u, \mathcal{R}(\mathcal{L})|_{\mathcal{A}^u}, \geq^i (\mathcal{L})|_{\mathcal{A}^u}, \mathcal{H}^i(\mathcal{L})|_{\mathcal{A}^u \cap \mathtt{Arg}_o(\mathcal{L})})$.*

Note that an optimal solution may differ from one agent to another even if the agents have the same sets of arguments supporting the same offers.

This is due to the fact that each agent i uses his/her own preference relation $\geq^i$ on arguments. This corresponds to the fact that in real life, from the same data, people do not necessarily take the same decision.

Example 6.3.4. *Let $Ag1$ and $Ag2$ negotiate about a restaurant to choose from two possible alternatives. They both agree that the first one is cheaper and the second one has better meals, but one agent may prefer the first and the other the second restaurant.*

Proposition 6.3.4. If o is an optimal solution for Agi, then there exists a dialogue $d = \langle m_1, \ldots, m_l \rangle$, such that o is an accepted solution for Agi at the end of the dialogue d.

6.3.2.2 Types of negotiation outcomes

Let us now analyze the different types of solutions of negotiation dialogues. Three types of solutions are distinguished. The first one, called *local solution*, is an offer which is accepted for both agents at a given step of a negotiation.

Definition 6.3.5 (Local solution of a negotiation)**.** *Let $d = \langle m_1, \ldots, m_t \rangle$ be a negotiation dialogue. An offer o is a* local solution *at the step l of d, with $1 \leq l \leq t$ iff o is accepted in both $\mathtt{AF}^1_l$ and $\mathtt{AF}^2_l$.*

Local solutions do not always exist, and when they exist, the protocol should be efficient in order to reach them.

There are cases where non-argumentative dialogues have no local solutions. It is particularly the case when at the beginning of the dialogue the two agents have no common accepted offer.

Theorem 6.3.1. Let $\mathtt{AF}^1$ and $\mathtt{AF}^2$ be the initial theories of the two agents such that $\mathcal{O}_a(\mathtt{AF}^1) \cap \mathcal{O}_a(\mathtt{AF}^1) = \emptyset$. There does not exist a non-argumentative dialogue d s.t. d has a local solution at some step.

The following result characterizes the case where there exists a local solution. In order to reach it, the agents should exchange the appropriate sequence of arguments.

Proposition 6.3.5. Let $Ag1$ and $Ag2$ be agents and $\mathtt{AF}^1 = (\mathcal{O}, \mathcal{A}^1, \mathcal{R}^1, \geq^1, \mathcal{H}^1)$ and $\mathtt{AF}^2 = (\mathcal{O}, \mathcal{A}^2, \mathcal{R}^2, \geq^2, \mathcal{H}^2)$ their initial theories. There exists a local solution iff $\exists \mathcal{A}'^1 \subseteq \mathcal{A}^1$ and $\exists \mathcal{A}'^2 \subseteq \mathcal{A}^2$ s.t.
$\mathcal{O}_a((\mathcal{O}, \mathcal{A}^1 \cup \mathcal{A}'^2, \mathcal{R}(\mathcal{L})|_{\mathcal{A}^1 \cup \mathcal{A}'^2}, \geq^1 (\mathcal{L})|_{\mathcal{A}^1 \cup \mathcal{A}'^2}, \mathcal{H}^1(\mathcal{L})|_{(\mathcal{A}^1 \cup \mathcal{A}'^2) \cap \mathtt{Arg}_o(\mathcal{L})})) \cap$
$\mathcal{O}_a((\mathcal{O}, \mathcal{A}'^1 \cup \mathcal{A}^2, \mathcal{R}(\mathcal{L})|_{\mathcal{A}'^1 \cup \mathcal{A}^2}, \geq^2 (\mathcal{L})|_{\mathcal{A}'^1 \cup \mathcal{A}^2}, \mathcal{H}^2(\mathcal{L})|_{(\mathcal{A}'^1 \cup \mathcal{A}^2) \cap \mathtt{Arg}_o(\mathcal{L})})) \neq \emptyset$.

The next result studies the situation when agents do not have to agree on everything but they agree on the arguments related to a given part of the negotiation, which is separated from other problems. If the first agent owns more information than the second, then there exists a dialogue in which the second will agree with the first one.

Theorem 6.3.2. Let $Ag1$ and $Ag2$ be agents and $\texttt{AF}^1 = (\mathcal{O}, \mathcal{A}^1, \mathcal{R}^1, \geq^1, \mathcal{H}^1)$ and $\texttt{AF}^2 = (\mathcal{O}, \mathcal{A}^2, \mathcal{R}^2, \geq^2, \mathcal{H}^2)$ their initial theories. Let $\mathcal{A} \subseteq \mathcal{A}^1 \cup \mathcal{A}^2$ be a set s.t. $\geq^1 |_{\mathcal{A}} = \geq^2 |_{\mathcal{A}}$ and let $\mathcal{A}$ be not attacked w.r.t. $\mathcal{R}'$ by arguments of $(\mathcal{A}^1 \cup \mathcal{A}^2) \setminus \mathcal{A}$. If $\mathcal{A}^1 \cap \mathcal{A} \supseteq \mathcal{A}^2 \cap \mathcal{A}$ and $\exists o \in \mathcal{O}$, $\exists a \in \mathcal{H}^1(o) \cap \mathcal{H}^2(o) \cap \mathcal{A}$ s.t. a is sceptically accepted in $\texttt{AF}^1$, then there exists a dialogue $d = \langle m_1, \ldots, m_l \rangle$ s.t. o is a local solution at step $t \leq l$ of d.

Another kind of time-dependent solution is a Pareto optimal solution. It takes into account the possible concessions that agents may make during a dialogue. In game-theoretic and heuristic-based approaches for negotiation, agents look for such solutions.

Definition 6.3.6 (Pareto optimal solution)**.** *Let $d = \langle m_1, \ldots, m_l \rangle$ be a negotiation dialogue. An offer $o \in \mathcal{O}$ is a Pareto optimal solution at step t iff $\nexists o' \in \mathcal{O}$ s.t. $(o' \succ^1_t o$ and $o' \succeq^2_t o)$ or $(o' \succeq^1_t o$ and $o' \succ^2_t o)$, where $\succeq^i_t$ is the preference relation on $\mathcal{O}$ returned by Agi at the step t.*

Roughly speaking, the protocols that have been developed in the literature for generating ABN dialogues lead to local solutions. Examples of such protocols are the one proposed by Amgoud, Dimopoulos, and Moraitis (2007) and its extended version (Hadidi, Dimopoulos, and Moraitis, 2010). Indeed, in those protocols, agents make concessions when they cannot defend their best offers.

It is easy to check that any local solution is also a Pareto optimal solution. However, the reverse is not true.

Proposition 6.3.6. If an offer is a local solution at the given step of a dialogue, then it is a Pareto optimal solution at that step of the dialogue.

The last kind of solution is the so-called *ideal solution*. It is an offer which is optimal for both agents.

Definition 6.3.7 (Ideal solution of a negotiation)**.** *An offer $o \in \mathcal{O}$ is an* ideal solution *for a negotiation iff it is optimal for both $Ag1$ and $Ag2$.*

We can show that if an ideal solution exists, then there exists at least one dialogue in which this solution is local.

Proposition 6.3.7. If an offer $o \in \mathcal{O}$ is an ideal solution, then there exists a dialogue d such that o is a local solution at the end of d.

It is natural to expect that two agents who share arguments and who agree on the preferences between those arguments can find an ideal solution.

Theorem 6.3.3. Let $\texttt{AF}^1 = (\mathcal{O}, \mathcal{A}^1, \mathcal{R}^1, \geq^1, \mathcal{H}^1)$ and $\texttt{AF}^2 = (\mathcal{O}, \mathcal{A}^2, \mathcal{R}^2, \geq^2, \mathcal{H}^2)$ be the theories of the two agents s.t. $\geq^1 (\mathcal{L}) = \geq^2 (\mathcal{L})$, $\mathcal{H}^1(\mathcal{L}) = \mathcal{H}^2(\mathcal{L})$ and $\mathcal{A}^1 \supseteq \mathcal{A}^2$. If o is an accepted solution for $Ag1$ before the beginning of a dialogue, then o is an ideal solution.

6.3.3 Added value of argumentation

The main goal of this chapter is to shed light on the role argumentation may play in negotiation dialogues. The idea is to study whether argumentation may improve or decrease the *quality of the outcome* of a dialogue, and under which conditions. It is clear that in real life, arguing does not necessarily lead to an agreement. In other words, it may be the case that two agents exchange arguments and at the end, the negotiation fails. Does this mean that arguing was not necessary in this case or it was rather harmful for the dialogue? In order to answer these questions, we need to compare the best outcomes that may be reached by non-argumentative dialogues with those reached by argumentative ones. In this section, we show that argumentation may improve the quality of the outcome. Indeed, in the best case, arguing leads to an ideal solution. When this is not possible, it can at least improve the choices made by the agents.

Let $\texttt{AF}^1$ and $\texttt{AF}^2$ be the initial theories of the two agents. We distinguish four situations which are the different combinations between local and ideal solutions.

Case 1. In the first case, there does not exist a local solution before a dialogue while there exists an ideal solution. In such a situation, argumentation will improve the outcome of a negotiation since it leads towards reaching such a solution. In the extreme case, it is sufficient for agents to exchange all their non-common arguments.

Theorem 6.3.4. Let $\texttt{AF}^1$ and $\texttt{AF}^2$ be the initial theories of the two agents. Let X be the set of ideal solutions and let $X \neq \emptyset$. For all $o \in X$, there exists an argumentative dialogue where o is a local solution at at the end of d.

Since before a dialogue starts, there is no local solution (i.e. there is no offer which is accepted for both agents), the agents should exchange arguments in order to have a chance to reach the ideal solution. This means that any non-argumentative dialogue will not lead to an ideal solution.

Theorem 6.3.5. Let $\mathtt{AF}^1$ and $\mathtt{AF}^2$ be the initial theories of the two agents s.t. $\mathcal{O}_a(\mathtt{AF}^1) \cap \mathcal{O}_a(\mathtt{AF}^2) = \emptyset$ and let $X \neq \emptyset$ be the set of ideal solutions. There does not exist a non-argumentative dialogue having $o \in X$ as a local solution at its end.

An important question is: what about Pareto optimal solutions? We show that it may happen that a non-argumentative dialogue ends with a Pareto optimal solution which is not an ideal one.

Example 6.3.5. *Assume that $\mathcal{O} = \{o_1, o_2, o_3\}$, $o_1 \succ^1 o_3 \succ^1 o_2$ and $o_2 \succ^2 o_3 \succ^2 o_1$. It is clear that there is no local solution while o_3 is a Pareto optimal one. If we assume that o_2 is the ideal solution, then it is clear that any non-argumentative dialogue will miss o_2.*

Conclusion: in this case, argumentative negotiations lead to an ideal solution (of course provided that the protocols are defined in an efficient way) while non-argumentative ones never find an ideal solution. Thus, argumentative dialogues yield a strictly better outcome than non-argumentative ones.

Case 2. Let us study the case where there exists at least one local solution before any dialogue and there exists an ideal solution. It is clear that if agents exchange appropriate offers, then a local solution may be reached even with non-argumentative dialogues.

Proposition 6.3.8. Let $\mathtt{AF}^1$ and $\mathtt{AF}^2$ be the initial theories of the two agents s.t. $\mathcal{O}_a(\mathtt{AF}^1) \cap \mathcal{O}_a(\mathtt{AF}^2) \neq \emptyset$. There exists a non-argumentative dialogue whose outcome is a member of $\mathcal{O}_a(\mathtt{AF}^1) \cap \mathcal{O}_a(\mathtt{AF}^2)$.

Note that the solution reached by non-argumentative dialogues may be a non-ideal one. Thus, an exchange of arguments may help to improve the quality of the output, i.e. to pass from a local solution to an ideal one. Moreover, according to Theorem 6.3.4, there exits an argumentative dialogue which leads for sure to an ideal solution. Thus, an argumentative dialogue will lead to an outcome which is at least as good as the outcome that may be reached by a non-argumentative dialogue. The following example illustrates this issue.

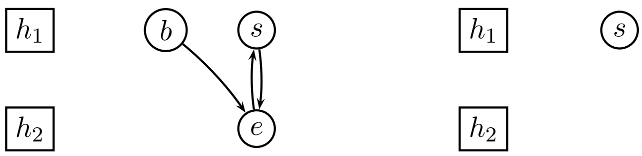
Example 6.3.6. *Assume that* $\mathcal{O} = \{o_1, o_2, o_3\}$*,* $o_1 \succ^1 o_3 \succ^1 o_2$ *and* $o_1 \succ^2 o_2 \succ^2 o_3$*. It is clear that* o_1 *is a local solution before any dialogue, and thus it can be reached with a simple exchange of offers. Assume now that* o_2 *is an ideal solution. Thus,* o_2 *is clearly better than* o_1 *since* o_2 *is a choice that both agents would make under "complete" information.*

Conclusion: in this case, argumentative negotiations lead to an ideal solution (of course provided that the protocols are defined in an efficient way) while non-argumentative ones sometimes find an ideal solution and sometimes not. Thus, argumentative dialogues yields a better or equal outcome than non-argumentative ones.

Case 3. Let us now consider the case where there is no ideal solution but there exists a local solution at the step 0, and let us show how sending and requesting arguments can be beneficial in this case.

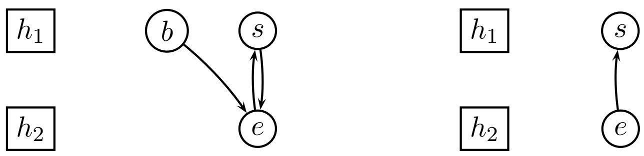
Example 6.3.7. *Let us consider the case where* $Ag1$ *wants to sell a house to* $Ag2$*. Let* $\mathcal{O} = \{h_1, h_2\}$*, where* h_1 *and* h_2 *represent two houses. The argument* b *represents the fact that the seller has a bonus if he sells* h_1*,* s *means that* h_1 *has a swimming-pool, and* e *means that* o_2 *is energy efficient. The preferences of the seller are* $b >^1 s$*,* $b >^1 e$*. The arguments and defeats of the seller are depicted on the left, and those of the buyer on the right side of Figure 6.1. Thus, for both* $Ag1$ *and* $Ag2$*,* o_1 *is acceptable and* o_2 *is rejected. The buyer has only one argument, but his potential preferences (formally captured by* $\geq^2 (\mathcal{L})$*, but we write* $\geq^2$ *to simplify notation) would be* $e >^2 s >^2 b$*.*

Figure 6.1: Buying a house: step 0.



Thus, the optimal solution for $Ag1$ *is* o_1*, and the optimal solution for* $Ag2$ *is* o_2*. There is no ideal solution. If no arguments are exchanged, agents agree on* o_1 *and the negotiation ends.* o_1 *is also a Pareto optimal solution at this step. However, if* $Ag2$ *requests information about energetic efficiency* *of houses, and* $Ag1$ *sends the argument* e*, then there is no local solution at the step* 1*.*

Figure 6.2: Buying a house: step 1.



At the first sight, argumentation ruined a possible deal. On the contrary, we believe that argumentation ameliorated the quality of negotiation, since $Ag2$ *was misled at the beginning. It is better not to conclude a deal than to accept a bad offer. There are several points to be made here.*

First, at the beginning, if energetic efficiency is the most important criterion for $Ag2$*, then he must request (any good protocol should allow this) informations about energetic efficiency of the houses (and a "proof" for them) even if they are not provided by the seller. If* $Ag1$ *does not provide them,* $Ag2$ *should be suspicious and hesitate or consult someone else. If* $Ag1$ *provides informations about energetic efficiency (in our case, by sending the argument* e*), then* $Ag2$ *will have a more realistic picture about the decision to make.*

Second, since there is no ideal solution, either one agent will deceive the other one, or both agents will make concessions. We argue that, from the "global" perspective, agents are in better situation at the step 1 *than at the step* 0*; since there is no ideal solution, the solution at the step* 0 *(since it is optimal for* $Ag1$*) is not a concession. At the step* 1*, agents are closer to a concession than at the step* 0*. Thus, argumentation helps them to find an offer acceptable for both of them.*

Conclusion: in the case where there is no ideal solution, sending and requesting arguments can help agents to make better decisions. If this means braking a deal, then it was certainly not a good one for at least one of the agents. Thus, we believe that in this case, argumentation leads to more quality solutions.

However, since the negotiating parties often have conflicting interests, an agent should take into account that another one may try to mislead him. An agent must request missing information which is relevant to making decision in question. If such information is not provided, or cannot be justified, the agent must be able to construct new arguments to model that fact. For example, if $Ag1$ refused to provide the argument about energetic efficiency, $Ag2$ should refuse further negotiations or search for the data (e.g. by asking a payed independent expert).

Provided a reasonable good protocol, arguing in this case may either lead to a Pareto optimal solution (if agents accept to make concessions) or to a failure. We argue that even a failure of negotiations is better for an agent than a bad solution. Indeed, the aim of a negotiation is not to reach any solution but to reach a solution which is good for both agents.

Case 4. The last case corresponds to the situation where there is no ideal solution and no local solution at the step 0. Non-argumentative dialogues may only find Pareto optimal solutions if agents accept to make concessions. However, those solutions may be bad for both agents as illustrated by the following example.

Example 6.3.8. *Assume that* $\mathcal{O} = \{o_1, o_2, o_3\}$. *The initial theory of* $Ag1$ *returns* $o_1 \succ^1 o_3 \succ^1 o_2$ *and the theory of* $Ag2$ *returns* $o_2 \succ^2 o_1 \succ^2 o_3$. *It is clear that there is no local solution. The offers* o_1 *or* o_2 *are Pareto optimal solutions, and may be accepted in a non-argumentative dialogue. Assume now that when agents exchange all of their arguments, the new theories of the two agents return respectively* $o_2 \succ^1 o_3 \succ^1 o_1$ *and* $o_3 \succ^2 o_2 \succ^2 o_1$. *This means that* o_1 *is the worst offer for both agents. Thus, if the two agents have sufficient information, they will never opt for* o_1.

Conclusion: in this case there, exchanging arguments may prevent agents from accepting a bad compromise and push them towards a better one.

6.4 Conclusion

In this chapter we have studied argumentation-based negotiation. We first presented existing works, classified in game-theoretic, heuristic-based and those based on argumentation. Argumentation has been integrated into negotiation dialogues in the early nineties by Sycara (1990). In that work, the author emphasized the advantages of using argumentation in negotiation dialogues, and a specific framework was introduced. In Kraus, Sycara, and Evenchik (1998), the different types of arguments that are used in a negotiation dialogue, such as threats and rewards, were discussed. Moreover, a particular framework for negotiation was proposed. Additional frameworks were also proposed (Parsons and Jennings, 1996; Tohmé, 1997). Even if all these frameworks are based on different logics, and use different definitions of arguments, they all have at their heart an exchange of offers and arguments. However, none of those proposals explain when arguments can be used within a negotiation, and how they should be dealt with by the agent that receives them. Thus the protocol for handling arguments was missing. Another limitation of the above frameworks is the fact that the argumentation frameworks they use are quite poor, since they use a very simple acceptability semantics. Amgoud, Parsons, and Maudet (2000b) suggested a negotiation framework that fills that gap. A protocol that handles the arguments was also proposed. However, the notion of concession is not modeled in that framework, and it is not clear what is the status of the outcome of the dialogue. Moreover, it is not explained how an agent chooses the offer to propose at a given step of the dialogue. Some authors have focused mainly on this decision problem (Kakas and Moraitis, 2006). They have proposed an argumentation-based decision framework that is used by agents in order to choose the offer to propose or to accept during the dialogue. In that work, agents are supposed to have a belief base and a goal base. Amgoud, Dimopoulos, and Moraitis (2007) proposed a more general setting. Indeed, the authors proposed an abstract argument-based decision model, and have shown how it is updated when an agent receives a new argument. Finally, they proposed a simple protocol allowing agents to exchange offers and arguments. Hadidi, Dimopoulos, and Moraitis (2010) proposed a slightly different version of that protocol. However, in both papers nothing is said about the quality of the outcome that may be returned under those protocols.

To the best of our knowledge the only work that attempted to show that argumentation is beneficial in negotiation is the work by Rahwan, Pasquier, Sonenberg, and Dignum (2007). In that paper, agents need resources in order to reach their goals. Thus, they negotiate with each other by exchanging resources and their goals following an extended version of the bargaining protocol. The paper shows that an exchange of goals may increase the utility of the outcome. Our work is more general in the sense that we do not focus on a particular negotiation object (like resources). Our notion of argument is much more general, and our analysis is made independently from any protocol. Finally, in our paper we have identified the different types of outputs and we have studied the role of argumentation whatever the negotiation object is.

To summarize, despite the huge number of works on argument-based approach for negotiation, there is no work which formally studies the impact of arguments on a negotiation dialogue as well as the role that is played by argumentation. We believe that our work is the first attempt in formalizing and identifying these issues.

Prediction is very difficult,
especially if it's about the future.
Niels Bohr

7 Conclusion and perspectives

This chapter concludes the thesis and presents several possible directions for future work.

7.1 Conclusion

The main contributions of this thesis are:

- the study of equivalence in logic-based argumentation,
- the study of the role(s) played by preferences in argumentation frameworks,
- the study of dynamics of argumentation-based decision frameworks, and
- the study of the impact of argumentation on the quality of negotiation outcomes.

The first contribution of this thesis is defining and studying different notions of **equivalence** in argumentation. Despite the obvious benefit of developing equivalence criteria for argumentation frameworks, this question has not received much attention. Until now, the only work on equivalence in argumentation (Oikarinen and Woltran, 2010) is conducted for abstract argumentation frameworks, which means that the structure of arguments is supposed to be not known. Only the notion of strong equivalence is addressed in that paper. But even the results concerning strong equivalence showed that if there are no self-attacking arguments, two argumentation frameworks are equivalent only if they coincide. We showed that when the structure of arguments is taken into account, similarities arise which are undetectable on the abstract level. We have proposed different equivalence criteria, investigated their links and shown under which conditions two frameworks are equivalent w.r.t. each of the proposed criteria. The notion of equivalence is then used in order to compute the *core(s)* of an argumentation framework. A core of a framework is an equivalent sub-framework. We showed that instead of using an argumentation framework which may be infinite, it is sufficient to consider one of its cores, which are usually finite.

The second part of the thesis concerns the use of **preferences** in argumentation. We have investigated the roles that preferences may play in an argumentation framework. Two particular *roles* have been identified: i) to privilege strong arguments when computing the standard solutions of a framework, and ii) to refine those standard solutions. We have shown that the two roles are completely independent and require different procedures for modeling them. Besides, we have shown that the existing works have tackled only the first role. Moreover, the proposed approaches suffer from a drawback which consists of returning conflicting extensions. We have proposed an approach which solves this problem and which presents two novelties: First, it takes into account preferences at a semantic level, i.e. it defines new acceptability semantics which are grounded on attacks and preferences between arguments. Second, a semantics is defined as a dominance relation that compares any pair of subsets of arguments.

The third part illustrates preference-based argumentation framework (PAF) in case of decision making and negotiation.

We have studied an instantiation of our PAF which rank-orders options in a **decision making** problem, where options are supported by arguments, which have different strengths and attack each other. Arguments supporting beliefs and those supporting options are distinguished. Our particular attention is drawn to the dynamics of this model. More precisely, we have shown how the ordering on options changes in light of a new argument. We have provided conditions under which an accepted option becomes rejected and vice versa. Our study is undertaken under two acceptability semantics: grounded semantics and preferred one. These results may be used in negotiation dialogues, namely to determine strategies. Indeed, at a given step of a dialog, an agent may choose which argument to send to another agent in order to change the status of an option. Our results may also help to understand which arguments are useful and which ones are useless in a given situation.

We have also used our PAF in order to show the benefits of arguing in **negotiation dialogues**. Even if it has been claimed by many researches that exchanging arguments may positively influence the quality of negotiation outcome, this was never formally shown on an abstract level. To accomplish this goal, it is necessary to define different types of solutions and to compare them. We have used an abstract framework for argument-based negotiation, defined the different types of solutions that may be reached in such dialogues, and we have formally shown that that arguing is beneficial during a negotiation. Our work is very general, since it does not depend on a particular notion of an argument and our analysis is made independently from any protocol.

7.2 Future work

In Chapter 3, we have studied equivalence between argumentation frameworks. All the results are shown under stable semantics. Our future work will include conducting this study for other semantics. Even if the main ideas will stay the same, there will certainly be changes when arguments are evaluated using different semantics (at least in proofs).

Chapter 4 is devoted to the study of the role of preferences in argumentation. We have proposed different relations which generalise stable, preferred and grounded semantics. Then, we have studied all the relations that can generalise stable semantics. An extension of our work would be to characterize the different dominance relations that generalise preferred semantics and those which generalise grounded semantics. A similar work can be done on those semantics proposed by Baroni et al. (2005), ideal semantics (Dung et al., 2007) and semi-stable semantics (Caminada, 2006b). Another direction of future work consists of studying how the new semantics can be used in a decision making context in order to rank order a set of alternatives. Namely, if we are able to compare extensions then we obtain more information than provided by traditional approaches. We believe that this information can be used in decision making, since better extensions will strengthen options supported by arguments of those extensions more than some weaker extensions and their arguments.

In Chapter 5, we have used an argument-based decision making framework which contains only arguments *in favour* of options. Future works will include the study of dynamics of an argumentation-based decision making framework which contains both arguments in favour and against options. In the proposed model, a preference relation between offers is defined on the basis of the partition of the set of offers to acceptable, negotiable, non-supported and rejected. The future work will be to refine this relation. Amgoud and Prade (2009) have proposed different criteria for comparing decisions which can be used to extend the decision making framework.

In Chapter 6, we have studied a negotiation framework based on exchange of arguments. Our future work concerns several points. The first one is to relax the assumption that the set of possible offers is the same to both agents. Indeed, it is more natural to assume that agents may have different sets of offers. Another urgent work would be to study the case where the preference relations between arguments may evolve. This means that the decision model should be able to reason about preferences. Also, we supposed that when an agent receives an argument, no new arguments are generated from this knowledge and the knowledge already owned by the agent. Finding a way to relax this hypothesis will also be a part of future work. The structure of arguments will have to be specified in order to do this.

A
Appendix

A.1 Proofs for results in Chapter 3

Proposition 3.2.1. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation system built from Σ. If Σ is finite and $\mathcal{R}$ satisfies $C2$, then $(\mathcal{A}, \mathcal{R})$ has a finite number of extensions.

Proof. Let $S_1, \ldots, S_n \subseteq \Sigma$ be all the consistent subsets of Σ. We will use the notation $\mathcal{A}_i = \{a \in \mathcal{A} \mid \texttt{Supp}(a) = S_i\}$, with $i \in \{1, \ldots, n\}$. (Note that some of the sets in $\mathcal{A}_1, \ldots, \mathcal{A}_n$ may be empty, but that is not important for the proof.) We will now prove that for every stable extension $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$, for any $i \in \{1, \ldots, n\}$, for any $a, a' \in \mathcal{A}_i$ we have $a \in \mathcal{E}$ iff $a' \in \mathcal{E}$. Let us suppose that $a \in \mathcal{E}$ and $a' \notin \mathcal{E}$. Since $\mathcal{E}$ is a stable extension, then $\exists b \in \mathcal{E}$ s.t. $b\mathcal{R}a'$. Since $\mathcal{R}$ satisfies $C2$ and $\texttt{Supp}(a) = \texttt{Supp}(a')$, then $b\mathcal{R}a$, which contradicts the fact that $\mathcal{E}$ is a stable extension. Therefore, if $a \in \mathcal{E}$ then $a' \in \mathcal{E}$. This means that for any $i \in \{1, \ldots, n\}$, any extension either contains all elements of $\mathcal{A}_i$ or neither of them. Formally, $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $\forall i \in \{1, \ldots, n\}$, we have $\mathcal{E} \cap \mathcal{A}_i = \mathcal{A}_i$ or $\mathcal{E} \cap \mathcal{A}_i = \emptyset$. Consequently, there is at most 2^n different extensions. □

Theorem 3.2.1. Let $\mathcal{F}$ and $\mathcal{F}'$ be two argumentation frameworks built on the same logic $(\mathcal{L}, \mathsf{CN})$. Table A.1 summarises the dependencies in the following form: $(\mathcal{F} \equiv_c \mathcal{F}') \Rightarrow (\mathcal{F} \equiv_{c'} \mathcal{F}')$.

Proof. Since there are 18 criteria available, there are 324 cases of this theorem. That is why we do not provide all the proofs. However, we provide proofs of several implications to show the reasoning behind them and we believe that the reader can use the similar reasoning to prove the other parts. Some counter-examples are also provided. Throughout the proof, we use notation $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$.

EQi/EQj	1	11	12	13	2	21	22	23	3	31	32	33	4	4b	5	5b	6	6b
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
11		+		+						+		+				+	+	+
12			+	+							+	+			+	+		+
13				+								+				+		+
2					+	+	+	+					+	+				
21						+		+						+				
22							+	+					+	+				
23								+						+				
3									+	+	+	+			+	+		
31										+		+				+		
32											+	+			+	+		
33												+				+		
4													+	+				
4b														+				
5															+	+		
5b																+		
6																	+	+
6b																		+

Table A.1: Links between criteria. For two criteria, c in row i, and c' in column j, sign $+$ means that c implies c', more precisely, if two argumentation frameworks are equivalent w.r.t. c then they are equivalent w.r.t. c'.

Example A.1.1. *Let* $\mathcal{L} = \{r_1, r_2, r_3, r_4, r_5, c\}$ *with* CN *being defined as follows: for all* $X \subseteq \mathcal{L}$,

$$\mathsf{CN}(X) = \begin{cases} \mathcal{L} \setminus \{c\}, & \text{if } c \notin X \text{ and } X \neq \emptyset \\ \mathcal{L}, & \text{if } c \in X \\ \emptyset, & \text{if } X = \emptyset \end{cases}$$

and $\mathsf{CN}\{r_1\} = \mathsf{CN}\{r_2\} = \mathsf{CN}\{r_3\} = \mathsf{CN}\{r_4\} = \mathsf{CN}\{r_5\} = \mathcal{L} \setminus \{c\}$. *Let* $a_1 = (\{r_1\}, r_1)$, $a_2 = (\{r_1\}, r_2)$, $a_3 = (\{r_1\}, r_3)$, $a_4 = (\{r_1\}, r_4)$, $a_5 = (\{r_1\}, r_5)$. *Let* $\mathcal{A} = \{a_1, a_2, a_3\}$, $\mathcal{R} = \{(a_2, a_3), (a_3, a_2)\}$, $\mathcal{A}' = \{a_4, a_5\}$ *and* $\mathcal{R}' = \{(a_4, a_5), (a_5, a_4)\}$. $\mathtt{Sc}(\mathcal{F}) = \{a_1\}$, $\mathtt{Sc}(\mathcal{F}') = \emptyset$. $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$ *since a bijection verifying conditions of EQ11 can be defined as:* $f : \mathtt{Ext}(\mathcal{F}) \rightarrow \mathtt{Ext}(\mathcal{F}')$, $f(\{a_1, a_2\}) = \{a_4\}$, $f(\{a_1, a_3\}) = \{a_5\}$.

This example illustrates the fact that EQ11 does not imply EQ1, EQ12, EQ2, EQ21, EQ22, EQ23, ...

We will now show that EQ11 implies EQ31. Let $a \in \mathtt{Cr}(\mathcal{F})$. We will prove that $\exists a' \in \mathtt{Cr}(\mathcal{F}')$ s.t. $a \approx_1 a'$. Since $a \in \mathtt{Cr}(\mathcal{F})$ then $\exists \mathcal{E} \in \mathtt{Ext}(\mathcal{F})$ s.t. $a \in \mathcal{E}$. Let f be a bijection from EQ11 and let $\mathcal{E}' = f(\mathcal{E})$. From EQ11, $\mathcal{E} \sim_1 \mathcal{E}'$, thus $\exists a' \in \mathcal{E}'$ s.t. $a \approx_1 a'$. This means that $\forall x \in \mathtt{Cr}(\mathcal{F})$, $\exists x' \in \mathtt{Cr}(\mathcal{F}')$ such that $x \approx_1 x'$. To prove that $\forall a' \in \mathtt{Cr}(\mathcal{F}')$, $\exists a \in \mathtt{Cr}(\mathcal{F})$ such that $a \approx_1 a'$ is similar. Thus, $\mathtt{Cr}(\mathcal{F}) \sim_1 \mathtt{Cr}(\mathcal{F}')$.

Note that EQ11 does not imply EQ4b in the general case as illustrated by Example A.1.1.

If EQ11 is true then EQ6 is true: Let $\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$ and $\mathcal{E}' = f(\mathcal{E})$. We can check that $\mathtt{Base}(\mathcal{E}) = \mathtt{Base}(\mathcal{E}')$. This means that $\forall \mathcal{E} \in \mathtt{Ext}(\mathcal{F})$,

$\exists\mathcal{E}' \in \mathtt{Ext}(\mathcal{F}')$ s.t. $\mathtt{Base}(\mathcal{E}) = \mathtt{Base}(\mathcal{E}')$ and $\forall\mathcal{E}' \in \mathtt{Ext}(\mathcal{F}')$, $\exists\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$ s.t. $\mathtt{Base}(\mathcal{E}) = \mathtt{Base}(\mathcal{E}')$.

Example A.1.2. *Let* $(\mathcal{L}, \mathsf{CN})$ *be the logic from Example A.1.1. Let* $a_1 = (\{r_1\}, r_1)$, $a_2 = (\{r_2\}, r_1)$, $a_3 = (\{r_3\}, r_1)$, $a_4 = (\{r_4\}, r_1)$, $a_5 = (\{r_5\}, r_1)$. *Let* $\mathcal{A} = \{a_1, a_2, a_3\}$, $\mathcal{R} = \{(a_2, a_3), (a_3, a_2)\}$, $\mathcal{A}' = \{a_4, a_5\}$ *and* $\mathcal{R}' = \{(a_4, a_5), (a_5, a_4)\}$. $\mathtt{Sc}(\mathcal{F}) = \{a_1\}$, $\mathtt{Sc}(\mathcal{F}') = \emptyset$.

Example A.1.2 shows that EQ12 does not imply EQ1, EQ11, EQ2, EQ21, EQ22, EQ23, ...

Let us prove that EQ12 implies EQ6b. Let $B \in \mathtt{Bases}(\mathcal{F})$. Then, $\exists\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$ s.t. $B = \mathtt{Base}(\mathcal{E})$. Let f be a bijection from EQ12 and let $\mathcal{E}' = f(\mathcal{E})$; then we have $\mathtt{Base}(\mathcal{E}') \cong B$.

We will now show that EQ33 implies EQ5b. Let $x \in \mathtt{Output}_{cr}(\mathcal{F})$. This means that $\exists a \in \mathcal{A}$ s.t. $a \in \mathtt{Cr}(\mathcal{F})$ and $\mathtt{Conc}(a) = x$. Since $\mathcal{F} \equiv_{EQ33} \mathcal{F}'$, then $\exists a' \in \mathcal{A}'$ s.t. $a' \in \mathtt{Cr}(\mathcal{F}')$ and $a \approx_3 a'$. Let $x' = \mathtt{Conc}(a')$. From $a \approx_3 a'$, we have $x \equiv x'$. Since x was arbitrary then $\forall x \in \mathtt{Output}_{cr}(\mathcal{F})$ $\exists x' \in \mathtt{Output}_{cr}(\mathcal{F}')$ s.t. $x \equiv x'$. To show that $\forall x' \in \mathtt{Output}_{cr}(\mathcal{F}')$ $\exists x \in \mathtt{Output}_{cr}(\mathcal{F})$ s.t. $x \equiv x'$ is similar. Thus, $\mathtt{Output}_{cr}(\mathcal{F}) \cong \mathtt{Output}_{cr}(\mathcal{F}')$, which means that $\mathcal{F} \equiv_{EQ5b} \mathcal{F}'$.

Example A.1.3. *Let* $(\mathcal{L}, \mathsf{CN})$ *be propositional logic and let* $\mathcal{A} = \{(\{x \wedge y\}, x)\}$, $\mathcal{A}' = \{(\{x \wedge z\}, x)\}$, $\mathcal{R} = \emptyset$, $\mathcal{R}' = \emptyset$. $\mathtt{Output}_{sc}(\mathcal{F}) = \mathtt{Output}_{sc}(\mathcal{F}') = \{x\}$.

Example A.1.3 shows that EQ4 does not imply EQ1, EQ11, EQ12, EQ13, EQ2, EQ21, EQ22, EQ23, EQ3, EQ31, EQ32, EQ33, EQ6, EQ6b.

□

Proposition 3.2.3. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ verifies $C1'$ and $C2$. For all $a, a', b, b' \in \mathcal{A}$, $(a \approx_1 a'$ and $b \approx_1 b') \Rightarrow (a\mathcal{R}b$ iff $a'\mathcal{R}b')$.

Proof. Let $a \approx_1 a'$ and $b \approx_1 b'$ and let $a\mathcal{R}b$. Since $\mathtt{Supp}(b) = \mathtt{Supp}(b')$ then from C2 we have that $a\mathcal{R}b'$. From C1' and $\mathtt{Conc}(a) \equiv \mathtt{Conc}(a')$, we obtain $a'\mathcal{R}b'$. To show that $a'\mathcal{R}b'$ implies $a\mathcal{R}b$ is similar. □

Proposition 3.2.4. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1'$ and $C2$. For all $a, a' \in \mathcal{A}$, if $a \approx_1 a'$, then $\forall\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$, $a \in \mathcal{E}$ iff $a' \in \mathcal{E}$.

Proof. $\Rightarrow$ Let $\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$, $a \approx_1 a'$ and $a \in \mathcal{E}$. We will prove that $a' \in \mathcal{E}$. Let $c \in \mathcal{A}$. Proposition 3.2.3 implies that $a\mathcal{R}c$ iff $a'\mathcal{R}c$ and $c\mathcal{R}a$ iff $c\mathcal{R}a'$. From these facts, we conclude that $\mathcal{E} \cup \{a'\}$ is conflict-free, since in the case of contrary, if for $b \in \mathcal{E}$, we had $a'\mathcal{R}b$ or $b\mathcal{R}a'$, Proposition 3.2.3 would imply $a\mathcal{R}b$ or $b\mathcal{R}a$, which is impossible. Since $\mathcal{E}$ is a stable extension then it is a maximal conflict-free set. This is why the case $a' \notin \mathcal{E}$ is not possible; consequently $a' \in \mathcal{E}$.

$\Leftarrow$ If $a \notin \mathcal{E}$, then $a' \notin \mathcal{E}$. The contrary would, from $a' \in \mathcal{E}$ (like in the first part of the proof) imply that $a \in \mathcal{E}$, contradiction. □

Proposition 3.2.5. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, and let $\mathcal{R}$ and $\mathcal{R}'$ verify $C1'$ and $C2$, and $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$. For all $a \in \mathcal{A}$ and for all $a' \in \mathcal{A}'$, if $a \approx_1 a'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a', \mathcal{F}')$.

Proof. If $\mathcal{F}$ has no extensions, then all arguments in $\mathcal{F}$ and $\mathcal{F}'$ are rejected. Thus, in the rest of the proof, we study the case when $\texttt{Ext}(\mathcal{F}) \neq \emptyset$. We will first prove that for any extension $\mathcal{E}$ of $\mathcal{F}$, $a \in \mathcal{E}$ iff $a' \in f(\mathcal{E})$, where $f : \texttt{Ext}(\mathcal{F}) \to \texttt{Ext}(\mathcal{F}')$ is a bijection which satisfies EQ11. Let $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$, let $a \in \mathcal{E}$ and let $a' \in \mathcal{A}'$ with $a \approx_1 a'$. Let $\mathcal{E}' = f(\mathcal{E})$; we will prove that $a' \in \mathcal{E}'$. From EQ11, one obtains $\exists a'' \in \mathcal{E}'$ s.t. $a \approx_1 a''$. (Note that we do not know whether $a' = a''$ or not.) We will prove that $\{a'\} \cup \mathcal{E}'$ is conflict-free. Let us suppose the contrary. This means that $\exists x \in \mathcal{E}'$ s.t. $x\mathcal{R}'a'$ or $a'\mathcal{R}'x$. From $x\mathcal{R}'a'$ and C2, we have $x\mathcal{R}'a''$ which contradicts the fact that $\mathcal{E}'$ is a stable extension. Else, from $a'\mathcal{R}'x$, condition C1' implies $a''\mathcal{R}'x$ which is not possible neither. We conclude that $\{a'\} \cup \mathcal{E}$ is conflict-free. Since $\mathcal{E}'$ is a stable extension, it attacks any argument $y \notin \mathcal{E}'$. Since $\mathcal{E}'$ does not attack a', then $a' \in \mathcal{E}'$.

This means that we showed that for any $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$, if $a \in \mathcal{E}$ then $a' \in f(\mathcal{E})$. Let $a \notin \mathcal{E}$ and let us prove that $a' \notin f(\mathcal{E})$. Suppose the contrary, i.e. suppose that $a' \in f(\mathcal{E})$. Since we made exactly the same hypothesis on $\mathcal{F}$ and $\mathcal{F}'$, by using the same reasoning as in the first part of the proof, we can prove that $a \in \mathcal{E}$, contradiction. This means that $a' \notin f(\mathcal{E})$. So, we proved that for any extension $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$, we have $a \in \mathcal{E}$ iff $a' \in f(\mathcal{E})$.

If a is sceptically accepted, then for any $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $a \in \mathcal{E}$. Let $\mathcal{E}' \in \texttt{Ext}(\mathcal{F}')$. Then, from EQ11, there exists $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$ s.t. $\mathcal{E}' = f(\mathcal{E})$. Since $a \in \mathcal{E}$, then $a' \in \mathcal{E}'$. If a is not sceptically accepted, then $\exists \mathcal{E} \in \texttt{Ext}(\mathcal{F})$ s.t. $a \notin \mathcal{E}$. It is clear that $\mathcal{E}' = f(\mathcal{E})$ is an extension of $\mathcal{F}'$ and that $a' \notin \mathcal{E}'$. Thus, in this case a' is not sceptically accepted in $\mathcal{F}'$.

Let a be credulously accepted in $\mathcal{F}$ and let $\mathcal{E} \in \mathcal{F}$ be an extension s.t. $a \in \mathcal{E}$. Then, $a' \in f(\mathcal{E})$, thus a' is credulously accepted in $\mathcal{F}'$. It is easy to see that the case when a is not credulously accepted in $\mathcal{F}$ and a' is credulously accepted in $\mathcal{F}'$ is not possible.

If a is rejected in $\mathcal{F}$, then a is not credulously accepted, thus a' is not credulously accepted which means that it is rejected. □

Theorem 3.2.2. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1'$ and $C2$. If $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ21, EQ23, EQ4b\}$.

Proof. Let us prove that EQ21 is verified. If $\texttt{Ext}(\mathcal{F}) = \emptyset$, then from EQ11, $\texttt{Ext}(\mathcal{F}') = \emptyset$. In this case, EQ21 is trivial, since $\texttt{Sc}(\mathcal{F}) = \texttt{Sc}(\mathcal{F}') = \emptyset$. Else, let $\texttt{Ext}(\mathcal{F}) \neq \emptyset$.

Let $\texttt{Sc}(\mathcal{F}) = \emptyset$. We will prove that $\texttt{Sc}(\mathcal{F}') = \emptyset$. Suppose the contrary and let $a' \in \texttt{Sc}(\mathcal{F}')$. Let $\mathcal{E}' \in \texttt{Ext}(\mathcal{F}')$. Argument a' is sceptically accepted, thus $a' \in \mathcal{E}'$. Let f be a bijection from EQ11, and let us denote $\mathcal{E} = f^{-1}(\mathcal{E}')$. From $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$, we obtain $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$. Furthermore, $\mathcal{E} \sim_1 \mathcal{E}'$, and, consequently, $\exists a \in \mathcal{E}$ s.t. $a \approx_1 a'$. Proposition 3.2.5 implies that a is sceptically accepted in $\mathcal{F}$, contradiction.

Let $\texttt{Sc}(\mathcal{F}) \neq \emptyset$ and let $a \in \texttt{Sc}(\mathcal{F})$. Since EQ11 is verified, and a is in at least one extension, then $\exists a' \in \mathcal{A}'$ s.t. $a' \approx_1 a$. Since EQ11 is verified then, from Proposition 3.2.5, a' is sceptically accepted in $\mathcal{F}'$. Thus $\forall a \in \texttt{Sc}(\mathcal{F})$, $\exists a' \in \texttt{Sc}(\mathcal{F}')$ s.t. $a' \approx_1 a$. To prove that $\forall a' \in \texttt{Sc}(\mathcal{F}')$, $\exists a \in \texttt{Sc}(\mathcal{F})$ s.t. $a \approx_1 a'$ is similar.
Since EQ21 implies EQ23 and EQ4b in the general case, as shown in Theorem 3.2.1, then we conclude that $\mathcal{F}$ and $\mathcal{F}'$ must be equivalent w.r.t. EQ21, EQ23 and EQ4b. □

Proposition 3.2.6. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1$ and $C2'$. For all $a, a', b, b' \in \mathcal{A}$, $(a \approx_2 a'$ and $b \approx_2 b') \Rightarrow (a\mathcal{R}b$ iff $a'\mathcal{R}b')$.

Proof. Similar to Proposition 3.2.3. □

Proposition 3.2.7. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1$ and $C2'$. For all $a, a' \in \mathcal{A}$, if $a \approx_2 a'$ then $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $a \in \mathcal{E}$ iff $a' \in \mathcal{E}$.

Proof. Similar to Proposition 3.2.4. □

Proposition 3.2.8. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1$ and

$C2'$, and $\mathcal{F} \equiv_{EQ12} \mathcal{F}'$. For all $a \in \mathcal{A}$ and for all $a' \in \mathcal{A}'$, if $a \approx_2 a'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a', \mathcal{F}')$.

Proof. Similar to Proposition 3.2.5. □

Theorem 3.2.3. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1$ and $C2'$. If $\mathcal{F} \equiv_{EQ12} \mathcal{F}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ22, EQ23, EQ4, EQ4b\}$.

Proof. Similar to Theorem 3.2.2. □

Proposition 3.2.9. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1'$ and $C2'$. For all $a, a', b, b' \in \mathcal{A}$, $(a \approx_3 a'$ and $b \approx_3 b') \Rightarrow (a\mathcal{R}b$ iff $a'\mathcal{R}b')$.

Proof. Similar to Proposition 3.2.3. □

Proposition 3.2.10. Let $(\mathcal{A}, \mathcal{R})$ be an argumentation framework s.t. $\mathcal{R}$ enjoys $C1'$ and $C2'$. For all $a, a' \in \mathcal{A}$, if $a \approx_3 a'$ then $\forall \mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $a \in \mathcal{E}$ iff $a' \in \mathcal{E}$.

Proof. Similar to Proposition 3.2.4. □

Proposition 3.2.11. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1'$ and $C2'$, and $\mathcal{F} \equiv_{EQ13} \mathcal{F}'$. For all $a \in \mathcal{A}$ and for all $a' \in \mathcal{A}'$, if $a \approx_3 a'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a', \mathcal{F}')$.

Proof. Similar to Proposition 3.2.5. □

Theorem 3.2.4. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from the same logic $(\mathcal{L}, \mathsf{CN})$, $\mathcal{R}$ and $\mathcal{R}'$ verify $C1'$ and $C2'$. If $\mathcal{F} \equiv_{EQ13} \mathcal{F}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ23, EQ4b\}$.

Proof. Similar to Theorem 3.2.2. □

Theorem 3.2.5. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\texttt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \texttt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1'$ and $C2$ and $\mathcal{A} \sim_1 \mathcal{A}'$, then $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$.

Proof. Let us first suppose that $\texttt{Ext}(\mathcal{F}) \neq \emptyset$ and let us define the function $f' : 2^{\mathcal{A}} \to 2^{\mathcal{A}'}$ as follows: $f'(B) = \{a' \in \mathcal{A}' \mid \exists a \in B \text{ s.t. } a' \approx_1 a\}$.

Let f be the restriction of f' to $\texttt{Ext}(\mathcal{F})$. We will prove that the image of this function is $\texttt{Ext}(\mathcal{F}')$ and that f is a bijection between $\texttt{Ext}(\mathcal{F})$ and $\texttt{Ext}(\mathcal{F}')$ which verifies EQ11.

- First, we will prove that for any $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$, $f(\mathcal{E}) \in \texttt{Ext}(\mathcal{F}')$. Let $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$ and let $\mathcal{E}' = f(\mathcal{E})$. We will prove that $\mathcal{E}'$ is conflict-free. Let $a', b' \in \mathcal{E}'$. There must exist $a, b \in \mathcal{E}$ s.t. $a \approx_1 a'$ and $b \approx_1 b'$. Since $\mathcal{E}$ is an extension, $\neg(a\mathcal{R}b)$ and $\neg(b\mathcal{R}a)$. By applying Proposition 3.2.3 on $(\texttt{Arg}(\mathcal{L}), \mathcal{R}(\mathcal{L}))$, we have that $\neg(a'\mathcal{R}'b')$ and $\neg(b'\mathcal{R}'a')$. Let $x' \in \mathcal{A}' \setminus \mathcal{E}'$. Then $\exists x \in \mathcal{A}$ s.t. $x \approx_1 x'$. Note also that it must be that $x \notin \mathcal{E}$. Since $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$, then $\exists y \in \mathcal{E}$ s.t. $y\mathcal{R}x$. From $\mathcal{A} \sim_1 \mathcal{A}'$, $\exists y' \in \mathcal{E}'$ s.t. $y' \approx_1 y$. From Proposition 3.2.3, $y'\mathcal{R}'x'$.

- We have shown that the image of f is the set $\texttt{Ext}(\mathcal{F}')$. We will now prove that $f : \texttt{Ext}(\mathcal{F}) \to \texttt{Ext}(\mathcal{F}')$ is injective. Let $\mathcal{E}_1, \mathcal{E}_2 \in \texttt{Ext}(\mathcal{F})$ with $\mathcal{E}_1 \neq \mathcal{E}_2$ and $\mathcal{E}' = f(\mathcal{E}_1) = f(\mathcal{E}_2)$. We will show that if $\mathcal{E}_1 \sim_1 \mathcal{E}_2$ then $\mathcal{E}_1 = \mathcal{E}_2$. Let us suppose that $\mathcal{E}_1 \sim_1 \mathcal{E}_2$ and $\mathcal{E}_1 \neq \mathcal{E}_2$. Without loss of generality, let $\exists x \in \mathcal{E}_1 \setminus \mathcal{E}_2$. Then, from $\mathcal{E}_1 \sim_1 \mathcal{E}_2$, $\exists x' \in \mathcal{E}_2$, s.t. $x' \approx_1 x$. Then, since $x \in \mathcal{E}_1$ and $x \notin \mathcal{E}_2$, from Proposition 3.2.4 we obtain that $x' \in \mathcal{E}_1$ and $x' \notin \mathcal{E}_2$. Contradiction with $x' \in \mathcal{E}_2$. Thus, we proved that $\mathcal{E}_1 \sim_1 \mathcal{E}_2$ implies $\mathcal{E}_1 = \mathcal{E}_2$. Consequently, it must be that $\neg(\mathcal{E}_1 \sim_1 \mathcal{E}_2)$. Without loss of generality, $\exists a_1 \in \mathcal{E}_1 \setminus \mathcal{E}_2$ s.t. $\nexists a_2 \in \mathcal{E}_2$ s.t. $a_1 \approx_1 a_2$. Let $a' \in \mathcal{A}'$ s.t. $a' \approx_1 a_1$. Recall that $\mathcal{E}' = f(\mathcal{E}_2)$. Thus, $\exists a_2 \in \mathcal{E}_2$ s.t. $a_2 \approx_1 a'$. Contradiction.

- We show that $f : \texttt{Ext}(\mathcal{F}) \to \texttt{Ext}(\mathcal{F}')$ is surjective. Let $\mathcal{E}' \in \texttt{Ext}(\mathcal{F}')$, and let us show that $\exists \mathcal{E} \in \texttt{Ext}(\mathcal{F})$ s.t. $\mathcal{E}' = f(\mathcal{E})$. Let $\mathcal{E} = \{a \in \mathcal{A} \mid \exists a' \in \mathcal{E}' \text{ s.t. } a \approx_1 a'\}$. From Proposition 3.2.3 we see that $\mathcal{E}$ is conflict-free. For any $b \in \mathcal{A} \setminus \mathcal{E}$, $\exists b' \in \mathcal{A}' \setminus \mathcal{E}'$ s.t. $b \approx_1 b'$. Since $\mathcal{E}' \in \texttt{Ext}(\mathcal{F}')$, then $\exists a' \in \mathcal{E}'$ s.t. $a'\mathcal{R}'b'$. Now, $\exists a \in \mathcal{E}$ s.t. $a \approx_1 a'$; from Proposition 3.2.3, $a\mathcal{R}b$. Thus, $\mathcal{E}$ is a stable extension in $\mathcal{F}$.

- We will now show that $f : \texttt{Ext}(\mathcal{F}) \to \texttt{Ext}(\mathcal{F}')$ verifies the condition of EQ11. Let $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$ and $\mathcal{E}' = f(\mathcal{E})$. Let $a \in \mathcal{E}$. Then, $\exists a' \in \mathcal{A}'$ s.t. $a' \approx_1 a$. From the definition of f, it must be that $a' \in \mathcal{E}'$. Similarly, if $a' \in \mathcal{E}'$, then must be an argument $a \in \mathcal{A}$ s.t. $a \approx_1 a'$, and again from the definition of the function f, we conclude that $a \in \mathcal{E}$.

From all above, we conclude that $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$. Let us take a look at the case when $\texttt{Ext}(\mathcal{F}) = \emptyset$. We will show that $\texttt{Ext}(\mathcal{F}') = \emptyset$. Suppose not and let $\mathcal{E}' \in \texttt{Ext}(\mathcal{F}')$. Let us define $\mathcal{E} = \{a \in \mathcal{A} \mid \exists a' \in \mathcal{E}' \text{ s.t. } a \approx_1 a'\}$. From Proposition 3.2.3, $\mathcal{E}$ must be conflict-free. The same proposition shows that for any $b \in \mathcal{A} \setminus \mathcal{E}$, $\exists a \in \mathcal{E}$ s.t. $a\mathcal{R}b$. Thus, $\mathcal{E}$ is a stable extension in $\mathcal{F}$. Contradiction with the hypothesis that $\texttt{Ext}(\mathcal{F}) = \emptyset$. □

Corollary 3.2.1. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\texttt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \texttt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1'$ and $C2$ and $\mathcal{A} \sim_1 \mathcal{A}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ13, EQ21, EQ23, EQ31, EQ33, EQ4b, EQ5b, EQ6, EQ6b\}$.

Proof. From Theorem 3.2.5, we conclude that $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$. Equivalences w.r.t. EQ13, EQ31, EQ33, EQ5b, EQ6 and EQ6b are consequences of EQ11, as shown in Theorem 3.2.1. Theorem 3.2.2 yields a conclusion that EQ21, EQ23 and EQ4b are verified. □

Theorem 3.2.6. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\texttt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \texttt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1$ and $C2'$ and $\mathcal{A} \sim_2 \mathcal{A}'$, then $\mathcal{F} \equiv_{EQ12} \mathcal{F}'$.

Proof. Similar to Theorem 3.2.5. □

Corollary 3.2.2. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\texttt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \texttt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1$ and $C2'$ and $\mathcal{A} \sim_2 \mathcal{A}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ13, EQ22, EQ23, EQ32, EQ33, EQ4, EQ4b, EQ5, EQ5b, EQ6b\}$.

Proof. Similar to Corollary 3.2.1. □

Theorem 3.2.7. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\texttt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$. $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \texttt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1'$ and $C2'$ and $\mathcal{A} \sim_3 \mathcal{A}'$, then $\mathcal{F} \equiv_{EQ13} \mathcal{F}'$.

Proof. Similar to Theorem 3.2.5. □

Corollary 3.2.3. Let $(\mathcal{L}, \mathsf{CN})$ be a fixed logic, $\texttt{Arg}(\mathcal{L})$ a set of arguments and $\mathcal{R}(\mathcal{L}) \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks s.t. $\mathcal{A}, \mathcal{A}' \subseteq \texttt{Arg}(\mathcal{L})$ and $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}, \mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{R}(\mathcal{L})$ satisfies $C1'$ and $C2'$ and $\mathcal{A} \sim_3 \mathcal{A}'$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ23, EQ33, EQ4b, EQ5b, EQ6b\}$.

Proof. Similar to Corollary 3.2.1. □

Corollary 3.3.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from a Tarskian logic $(\mathcal{L}, \mathsf{CN})$, s.t. $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ verifies C1' and C2, $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$ and $\mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{F} \equiv_{EQ11S}$, then $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ13S, EQ21S, EQ23S, EQ31S, EQ33S, EQ4bS, EQ5bS, EQ6S, EQ6bS\}$.

Proof. Let $B \subseteq \mathtt{Arg}(\mathcal{L})$. Since $\mathcal{F} \equiv_{EQ11S} \mathcal{F}'$, then $\mathcal{F} \oplus \mathcal{B} \equiv_{EQ11} \mathcal{F}' \oplus \mathcal{B}$. From Corollary 3.2.1, $\mathcal{F} \oplus \mathcal{B} \equiv_x \mathcal{F}' \oplus \mathcal{B}$, with $x \in \{EQ13, EQ21, EQ23, EQ31, EQ33, EQ4b, EQ5b, EQ6, EQ6b\}$. Since $\mathcal{B}$ was arbitrary, we conclude that $\mathcal{F} \equiv_x \mathcal{F}'$ with $x \in \{EQ13S, EQ21S, EQ23S, EQ31S, EQ33S, EQ4bS, EQ5bS, EQ6S, EQ6bS\}$. □

Theorem 3.3.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be two argumentation frameworks built from a Tarskian logic $(\mathcal{L}, \mathsf{CN})$, s.t. $\mathcal{R}(\mathcal{L}) \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ verifies C1' and C2, $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$ and $\mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$. If $\mathcal{A} \sim_1 \mathcal{A}'$, then $\mathcal{F} \equiv_{EQ11S} \mathcal{F}'$.

Proof. Let $\mathcal{B} \subseteq \mathtt{Arg}(\mathcal{L})$. Since $\mathcal{A} \sim_1 \mathcal{A}'$ then clearly $\mathcal{A} \cup \mathcal{B} \sim_1 \mathcal{A}' \cup \mathcal{B}$. From Theorem 3.2.5, we obtain that $\mathcal{F} \oplus \mathcal{B} \equiv_{EQ11} \mathcal{F}' \oplus \mathcal{B}$. Thus, $\mathcal{F} \equiv_{EQ11S} \mathcal{F}'$. □

Theorem 3.4.1. Let $\mathcal{F}$ be an argumentation framework and $\mathcal{F}'$ one of its cores. Then: $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$.

Proof. The result is obtained by applying Theorem 3.2.5 on $\mathcal{F}$ and $\mathcal{F}'$. □

Corollary 3.4.1. Let $\mathcal{F}$ be an argumentation framework and $\mathcal{F}'$ one of its cores. Then:

- $\mathtt{Sc}(\mathcal{F}) \sim_1 \mathtt{Sc}(\mathcal{F}')$
- $\mathtt{Cr}(\mathcal{F}) \sim_1 \mathtt{Cr}(\mathcal{F}')$
- $\mathtt{Output}_{sc}(\mathcal{F}) \cong \mathtt{Output}_{sc}(\mathcal{F}')$
- $\mathtt{Output}_{cr}(\mathcal{F}) \cong \mathtt{Output}_{cr}(\mathcal{F}')$
- $\mathtt{Bases}(\mathcal{F}) = \mathtt{Bases}(\mathcal{F}')$

Proof. From Theorem 3.4.1, $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$. From Theorems 3.2.1 and 3.2.2, we have $\mathcal{F} \equiv_x \mathcal{F}'$, with $x \in \{EQ21, EQ31, EQ4b, EQ5b, EQ6\}$, which ends the proof. □

Proposition 3.4.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and let $a, a' \in \mathcal{A}$ be two arguments such that $\texttt{Supp}(a) = \texttt{Supp}(a')$. Then: $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a', \mathcal{F})$.

Proof. We will prove that for every stable extension $\mathcal{E}$, we have $a \in \mathcal{E}$ iff $a' \in \mathcal{E}$. Let us suppose that $a \in \mathcal{E}$ and $a' \notin \mathcal{E}$. Since $\mathcal{E}$ is a stable extension, then $\exists b \in \mathcal{E}$ s.t. $b\mathcal{R}a'$. From C2, we have that $b\mathcal{R}a$ which contradicts the fact that $\mathcal{E}$ is a stable extension. The case $a \notin \mathcal{E}$ and $a' \in \mathcal{E}$ is symmetric. This means that each extension of $\mathcal{F}$ either contains both a and a or does not contain any of those two arguments. Consequently, the statuses of those arguments must coincide. □

Proposition 3.4.2. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ its core.

- If $a \in \mathcal{A}'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a, \mathcal{F}')$,
- If $a \notin \mathcal{A}'$ then $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(b, \mathcal{F}')$, where $b \in \mathcal{A}'$ is an arbitrary argument s.t. $\texttt{Supp}(a) = \texttt{Supp}(b)$.

Proof.

- From Proposition 3.4.1, $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$. Now, from Proposition 3.2.5, $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a, \mathcal{F}')$.
- From the first part of the proposition, $\texttt{Status}(b, \mathcal{F}) = \texttt{Status}(b, \mathcal{F}')$. From Proposition 3.4.1, $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(b, \mathcal{F})$. Thus, it must be $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(b, \mathcal{F}')$.

□

Theorem 3.4.2. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework built over a knowledge base Σ (i.e. let $\mathcal{A} \subseteq \texttt{Arg}(\Sigma)$). If $\texttt{Cncs}(\Sigma)/\equiv$ is finite, then any core of $\mathcal{F}$ is finite.

Proof. Let $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ be a core of $\mathcal{F}$ and let us prove that $\mathcal{F}'$ is finite. Since Σ is finite, then $\{\texttt{Supp}(a) \mid a \in \mathcal{A}'\}$ must be finite. If for all $H \in \{\texttt{Supp}(a) \mid a \in \mathcal{A}'\}$, the set $\{a \in \mathcal{A}' \mid \texttt{Supp}(a) = H\}$, is finite, then the set $\mathcal{A}'$ is clearly finite. Else, there exists $H_0 \in \{\texttt{Supp}(a) \mid a \in \mathcal{A}'\}$, s.t. the set $\mathcal{A}_{H_0} = \{a \in \mathcal{A}' \mid \texttt{Supp}(a) = H_0\}$ is infinite. By the definition of $\mathcal{A}'$, one obtains that $\forall a, b \in \mathcal{A}_{H_0}$, $\texttt{Conc}(a) \not\equiv \texttt{Conc}(b)$. It is clear that $\forall a \in \mathcal{A}_{H_0}$, $\texttt{Conc}(a) \in \texttt{Cncs}(\Sigma)$. This implies that there are infinitely many different formulae having pairwise non-equivalent conclusions in $\texttt{Cncs}(\Sigma)$, formally, the set $\texttt{Cncs}(\Sigma)/\equiv$ is infinite, contradiction. □

Lemma A.1.1. Let $(\mathcal{A}_c, \mathcal{R}_c)$ be a core of $\mathcal{F}_\downarrow = (\mathcal{A}_\downarrow = \texttt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow = \mathcal{R}(\mathcal{L})|_{\mathcal{A}_\downarrow})$ and let $\mathcal{A}_1$ be an arbitrary set which contains $\mathcal{A}_c$, i.e. $\mathcal{A}_c \subseteq \mathcal{A}_1 \subseteq \texttt{Arg}(\Sigma)$. We define $\mathcal{R}_1 = \mathcal{R}|_{\mathcal{A}_1}$, as expected, and $\mathcal{F}_1 = (\mathcal{A}_1, \mathcal{R}_1)$. Let $S_1, \dots, S_n$ be all the maximal consistent subsets of Σ, and let $\mathcal{E}_1 = \texttt{Arg}(S_1) \cap \mathcal{A}_1, \dots, \mathcal{E}_n = \texttt{Arg}(S_n) \cap \mathcal{A}_1$. Then, $\texttt{Ext}(\mathcal{F}_1) = \{\mathcal{E}_1, \dots, \mathcal{E}_n\}$.

Proof. We will first prove that for any maximal consistent subset S_i of Σ, the set $\mathcal{E}_i = \texttt{Arg}(S_i) \cap \mathcal{A}_1$ is a stable extension of $\mathcal{F}_1$. It is easy to see that if S_i is consistent then $\texttt{Arg}(S_i)$ is conflict-free. Let us prove that $\mathcal{E}_i$ attacks any argument in $\mathcal{A}_1 \setminus \mathcal{E}_i$. Let $a' \in \mathcal{A}_1 \setminus \mathcal{E}_i$. Since $a' \notin \mathcal{E}_i$, then $\exists h \in \texttt{Supp}(a')$ s.t.
$h \notin S_i$. Since $\texttt{Supp}(a') \subseteq \Sigma$ and S_i is a maximal consistent subset of Σ, it follows that $S_i \cup \{h\}$ is inconsistent. Then, there exists a minimal set $C \subseteq S_i$ s.t.
$C \cup \{h\}$ is inconsistent. Let $a = (C, \neg h)$. Then, since a uses only atoms from Σ (since $h \in \Sigma$) and since $(\mathcal{A}_c, \mathcal{R}_c)$ is a core of $\mathcal{F}_\downarrow$ then $\exists a_1 \in \mathcal{A}_c$ s.t. $a_1 \approx_1 a$. Since $\texttt{Supp}(a_1) \subseteq S_i$ then $a_1 \in \mathcal{E}_i$. Also, $a_1 \mathcal{R}_1 a'$. Hence, $\mathcal{E}_i$ is a stable extension of $\mathcal{F}_1$.

We will now prove that for any $\mathcal{E}' \in \texttt{Ext}(\mathcal{F}_1)$, there exists a maximal consistent subset of Σ, denoted S', s.t.
$\mathcal{E}' = \texttt{Arg}(S') \cap \mathcal{A}_1$. To show this, we will show that: 1) $\texttt{Base}(\mathcal{E}')$ is consistent, 2) $\texttt{Base}(\mathcal{E}')$ is a maximal consistent set in Σ, 3) $\mathcal{E}' = \texttt{Arg}(\texttt{Base}(\mathcal{E}')) \cap \mathcal{A}_1$.

1. Let $S' = \texttt{Base}(\mathcal{E}')$. Suppose that S' is an inconsistent set and let $C \subseteq S'$ be a minimal inconsistent subset of S'. Let $C = \{f_1, \dots, f_k\}$, and let us construct the following argument: $a = (C \setminus \{f_1\}, \neg f_1)$. Since $\mathcal{E}'$ is conflict-free, then $a \notin \mathcal{E}'$ and $\nexists a_1 \in \mathcal{E}'$ s.t. $a_1 \approx_1 a$. Since $\mathcal{A}_c \subseteq A_1$, then there exists an argument $a_1 \in \mathcal{A}_1$ s.t. $a_1 \approx_1 a$. This means that, $a_1 \in \mathcal{A}_1 \setminus \mathcal{E}'$. Since $\mathcal{E}'$ is a stable extension, $\mathcal{E}'$ must attack a_1. Formally, $\exists a' \in \mathcal{E}'$ s.t. $a' \mathcal{R}_1 a_1$. So, $\texttt{Conc}(a') \equiv \neg f_2$ or $\texttt{Conc}(a') \equiv \neg f_3$, …, or $\texttt{Conc}(a') \equiv \neg f_k$. Without loss of generality, let $\texttt{Conc}(a') \equiv \neg f_k$. Since $f_k \in S'$, then there exists at least one argument a_k in $\mathcal{E}'$ s.t.
$f_k \in \texttt{Supp}(a_k)$. Consequently, $\mathcal{E}'$ is not conflict-free, since a' attacks at least one argument in $\mathcal{E}'$.

2. Let $S' = \texttt{Base}(\mathcal{E}')$ and suppose that S' is not a maximal consistent set in Σ. According to (1) S' is consistent, hence $\exists f \in \Sigma \setminus S'$ s.t.
$S' \cup \{f\}$ is consistent. Thus, for the argument $b = (\{f\}, f)$, we have that $\exists b_1 \in \mathcal{A}_1 \setminus \mathcal{E}'$ s.t. $b_1 \approx b$, but no argument in $\mathcal{E}'$ attacks b_1. (This is since $\neg f$ cannot be inferred from S'; consequently, no argument can

be constructed from S' having its conclusion logically equivalent to $\neg f$.)

3. It is easy to see that for any set of arguments $\mathcal{E}'$, we have $\mathcal{E}' \subseteq \mathtt{Arg}(\mathtt{Base}(\mathcal{E}'))$. Since $\mathcal{S}' = \mathtt{Base}(\mathcal{E}')$ is a consistent set, then set of arguments $\mathtt{Arg}(\mathtt{Base}(\mathcal{E}')) \cap \mathcal{A}_1$ must be conflict-free. From the fact that $\mathcal{E}'$ is a stable extension of $\mathcal{F}_1$, we conclude that the case $\mathcal{E}' \subsetneq \mathtt{Arg}(\mathtt{Base}(\mathcal{E}')) \cap \mathcal{A}_1$ is not possible (since every stable extension is a *maximal* conflict-free set).

We will now show that if S, S' are two different maximal consistent subsets of Σ, $\mathcal{E} = \mathtt{Arg}(S) \cap \mathcal{A}_1$ and $\mathcal{E}' = \mathtt{Arg}(S') \cap \mathcal{A}_1$, then $\mathcal{E} \neq \mathcal{E}'$. Without loss of generality, let $f \in S \setminus S'$. Let $a_f \in \mathcal{A}_1$ be an argument s.t.
$\mathtt{Supp}(a_f) = \{f\}$ and $\mathtt{Conc}(a_f) \equiv f$. Such an argument must exist since $\mathcal{A}_1$ contains $\mathcal{A}_c$, and $(\mathcal{A}_c, R_c)$ is a core of $\mathcal{F}_\downarrow$. It is clear that $a \in \mathcal{E} \setminus \mathcal{E}'$, which shows that $\mathcal{E} \neq \mathcal{E}'$. This ends the proof. □

Theorem 3.4.3. Let $\mathcal{F} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ and $\mathcal{F}_\downarrow = (\mathtt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow)$. For all $a \in \mathtt{Arg}(\Sigma)_\downarrow$, $\mathtt{Status}(a, \mathcal{F}) = \mathtt{Status}(a, \mathcal{F}_\downarrow)$.

Proof. Let $S_1, \ldots, S_n$ be all the maximal consistent subsets of Σ. Since $(\mathtt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow)$ and $(\mathtt{Arg}(\Sigma), \mathcal{R})$ both contain at least one core of $(\mathtt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow)$ (in fact, they both contain *all* cores of this set) then Lemma A.1.1 implies that extensions of $(\mathtt{Arg}(\Sigma), \mathcal{R})$ are exactly $\mathtt{Arg}(S_i)$, and extensions of $(\mathtt{Arg}(\Sigma)_\downarrow, \mathcal{R}_\downarrow)$ are exactly $\mathtt{Arg}(S_i) \cap \mathtt{Arg}(\Sigma)_\downarrow$, when $1 \leq i \leq n$. Thus, the two frameworks have the same number of extensions and any argument of $\mathtt{Arg}(\Sigma)_\downarrow$ is in the same number of extensions in them. Consequently, its status must be the same in both frameworks. □

Theorem 3.4.4. Let $\mathcal{F} = (\mathtt{Arg}(\Sigma), \mathcal{R})$ be an argumentation framework built over Σ. For all $a \in \mathtt{Arg}(\Sigma) \setminus \mathtt{Arg}(\Sigma)_\downarrow$, $\mathtt{Status}(a, \mathcal{F}) = \mathtt{Status}(b, \mathcal{F})$ where $b \in \mathtt{Arg}(\Sigma)_\downarrow$ and $\mathtt{Supp}(a) = \mathtt{Supp}(b)$.

Proof. Let $a \in \mathtt{Arg}(\Sigma) \setminus \mathtt{Arg}(\Sigma)_\downarrow$ and $b \in \mathtt{Arg}(\Sigma)_\downarrow$ and let $\mathtt{Supp}(a) = \mathtt{Supp}(b)$. From Proposition 3.4.1, $\mathtt{Status}(a, \mathcal{F}) = \mathtt{Status}(b, \mathcal{F})$. □

Proposition 3.4.3. It holds that $|\mathtt{Arg}(\Sigma)_\downarrow / \approx_1| \leq 2^n \cdot 2^{2^m}$, where $n = |\Sigma|$ and $m = |\mathtt{Atoms}(\Sigma)|$.

Proof. There are at most 2^n different supports of arguments. It is well-known that there are at most 2^{2^m} logically non-equivalent Boolean functions of m variables. Thus, for any support H, there are at most 2^{2^m} different non-equivalent arguments, where m is the number of different atoms in Σ. □

Theorem 3.4.5. Let $\mathcal{F} = (\mathcal{A} = \texttt{Arg}(\Sigma), \mathcal{R})$ be an argumentation framework built over a knowledge base Σ, let $\mathcal{F}' = (\mathcal{A}' = \texttt{Arg}(\Sigma)_{\downarrow}, \mathcal{R}')$, with $\mathcal{R}' = \mathcal{R}|_{\mathcal{A}'}$, and let $\mathcal{G}$ be a core of $\mathcal{F}'$. Then, $\texttt{Output}_{sc}(\mathcal{F}) = \{x \in \mathcal{L} \text{ s.t.} \texttt{Output}_{sc}(\mathcal{G}) \vdash x\}$.

Proof. Let $\mathcal{G} = (\mathcal{A}_g, \mathcal{R}_g)$.
$\Rightarrow$ Let $h \in \texttt{Output}_{sc}(\mathcal{F})$. This means that $\exists a \in \mathcal{A}$ s.t. $a \in \texttt{Sc}(\mathcal{F})$. Let $a = (H, h)$ and let $H = \{f_1, \ldots, f_k\}$. Since a is an argument, then H is consistent and no formula in H can be deduced from other formulae in H. Then, $a' = (H, f_1 \wedge \ldots \wedge f_k)$ must also be an argument. Note that its conclusion contains only atoms from Σ, thus $a' \in \mathcal{A}'$. Consequently, there must exist an argument $a_g \in \mathcal{A}_g$ s.t.
$a_g \approx_1 a'$. $\mathcal{G}$ is a core of $\mathcal{F}'$, thus they are equivalent w.r.t. EQ11 (Theorem 3.4.1). Since equivalent arguments have the same status in equivalent frameworks (Proposition 3.2.5) then a_g is sceptically accepted in $\mathcal{G}$. So, $\texttt{Output}_{sc}(\mathcal{G}) \vdash f_1 \wedge \ldots \wedge f_k$. Consequently, $\texttt{Output}_{sc}(\mathcal{G}) \vdash h$.

$\Leftarrow$ Let f be a propositional formula that can be deduced from $\texttt{Output}_{sc}(\mathcal{G})$. Let $S_1, \ldots, S_n$ be all the maximal consistent subsets of Σ. According to Lemma A.1.1, $\exists a \in \mathcal{A}_g$ s.t.$\texttt{Supp}(a) \subseteq S_1 \cap \ldots \cap S_n$ and $\texttt{Conc}(a) = f$. Let us denote $H = \texttt{Supp}(a)$. Obviously, $H \vdash f$. Furthermore, $H \subseteq S_1 \cap \ldots \cap S_n$. From those two facts, we conclude that it must exist an argument $a' \in \texttt{Arg}(\Sigma)$ s.t.
$\texttt{Supp}(a') \subseteq H$ and $\texttt{Conc}(a') = f$. From Lemma A.1.1, a' is sceptically accepted in $\mathcal{F}$. Thus, $f \in \texttt{Output}_{sc}(\mathcal{F})$.

□

Theorem 3.5.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework which contains a core of $\mathcal{G} = (\mathcal{A}_g = \texttt{Arg}(\Sigma), \mathcal{R}_g = \mathcal{R}(\mathcal{L})|_{\mathcal{A}_g})$ and let $\mathcal{E} \subseteq \texttt{Arg}(\Sigma)$. Then:

- $\mathcal{F} \equiv_{EQ11} \mathcal{F} \oplus \mathcal{E}$
- $\forall a \in \mathcal{A}, \texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a, \mathcal{F} \oplus \mathcal{E})$
- $\forall e \in \mathcal{E} \setminus \mathcal{A}, \texttt{Status}(e, \mathcal{F} \oplus \mathcal{E}) = \texttt{Status}(a, \mathcal{F})$, where $a \in \mathcal{A}$ is any argument s.t. $\texttt{Supp}(a) = \texttt{Supp}(e)$.

Proof. Let $\mathcal{F}' = \mathcal{F} \oplus \mathcal{E}$ with $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ and let $\mathcal{H} = (\mathcal{A}_h, \mathcal{R}_h)$ be a core of $\mathcal{G}$ s.t. $\mathcal{A}_h \subseteq \mathcal{A}$. We will first show that $\mathcal{H}$ is a core of both $\mathcal{F}$ and $\mathcal{F}'$.

- Let us first show that $\mathcal{H}$ is a core of $\mathcal{F}$. We will show that all conditions of Definition 3.4.1 are verified.

– We have already seen why $\mathcal{A}_h \subseteq \mathcal{A}$.
– We will show that $\forall a \in \mathcal{A}$, $\exists! a' \in \mathcal{A}_h$ s.t. $a' \approx_1 a$. Let $a \in \mathcal{A}$. Since $a \in \mathcal{A}_g$ and $\mathcal{H}$ is a core of $\mathcal{G}$, then $\exists! a' \in \mathcal{A}_h$ s.t. $a' \approx_1 a$.
– Since $\mathcal{R} = \mathcal{R}(\mathcal{L})|_{\mathcal{A}}$ and $\mathcal{R}_h = \mathcal{R}(\mathcal{L})|_{\mathcal{A}_h}$ then from $\mathcal{A}_h \subseteq \mathcal{A}$ we obtain that $\mathcal{R}_h = \mathcal{R}|_{\mathcal{A}_h}$.

Thus, $\mathcal{H}$ is a core of $\mathcal{F}$. Let us now show that $\mathcal{H}$ is also a core of $\mathcal{F}'$:

– Since $\mathcal{A}_h \subseteq \mathcal{A}$ and $\mathcal{A} \subseteq \mathcal{A}'$ then $A_h \subseteq \mathcal{A}'$.
– Let $a \in \mathcal{A}'$. Since $a \in \mathcal{A}_g$ and $\mathcal{H}$ is a core of framework $\mathcal{G}$, then $\exists! a' \in \mathcal{A}_h$ s.t. $a' \approx_1 a$.
– Since $\mathcal{R}' = \mathcal{R}(\mathcal{L})|_{\mathcal{A}'}$, $\mathcal{R}_h = \mathcal{R}(\mathcal{L})|_{\mathcal{A}_h}$ and $\mathcal{A}_h \subseteq \mathcal{A}'$, then we obtain that $\mathcal{R}_h = \mathcal{R}'|_{\mathcal{A}_h}$.

We have shown that $\mathcal{H}$ is a core of $\mathcal{F}$ and of $\mathcal{F}'$. From Theorem 3.4.1, $\mathcal{F} \equiv_{EQ11} \mathcal{H}$ and $\mathcal{F}' \equiv_{EQ11} \mathcal{H}$. Since $\equiv_{EQ11}$ is an equivalence relation, then $\mathcal{F} \equiv_{EQ11} \mathcal{F}'$. Let $a \in \mathcal{A}$. From Proposition 3.2.5, $\texttt{Status}(a, \mathcal{F}) = \texttt{Status}(a, \mathcal{F}')$.

Let $e \in \mathcal{A}' \setminus \mathcal{A}$ and let $a \in \mathcal{A}$ be an argument such that $\texttt{Supp}(a) = \texttt{Supp}(e)$. From Proposition 3.4.1, we obtain that $\texttt{Status}(e, \mathcal{F}') = \texttt{Status}(a, \mathcal{F}')$. Since we have seen that $\texttt{Status}(a, \mathcal{F}') = \texttt{Status}(a, \mathcal{F})$, then $\texttt{Status}(e, \mathcal{F}') = \texttt{Status}(a, \mathcal{F})$.

□

A.2 Proofs for results in Chapter 4

Proposition 4.3.1. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. If a dominance relation $\succeq$ satisfies postulate $P1$, then each element of the set $\succeq_{max}$ is conflict-free w.r.t. $\mathcal{R}$.

Proof. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. Assume that $\succeq$ is a dominance relation which satisfies postulate $P1$. Let us show that each element of the set $\succeq_{max}$ is conflict-free w.r.t. $\mathcal{R}$.

Assume that $\mathcal{E} \in \succeq_{max}$. Thus, $\forall \mathcal{E}' \in \mathcal{P}(\mathcal{A})$, $\mathcal{E} \succeq \mathcal{E}'$. In particular, $\mathcal{E} \succeq \emptyset$. Since $\emptyset \in \mathcal{CF}(\mathcal{T})$, then from Postulate $P1$, $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$. □

Theorem 4.3.1. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework and $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$. Let $\texttt{Ext}(\mathcal{F})$ be the set containing all the stable extensions of $\mathcal{F}$. The equality $\texttt{Ext}(\mathcal{F}) = \succeq_{max}$ holds iff $\forall \mathcal{E} \in \mathcal{P}(\mathcal{A})$,

1. if $\mathcal{E} \notin \mathcal{CF}(\mathcal{F})$ then $\exists \mathcal{E}' \in \mathcal{P}(\mathcal{A})$ s.t. not($\mathcal{E} \succeq \mathcal{E}'$), and

2. if $\mathcal{E} \in \mathcal{CF}(\mathcal{F})$ and $\forall a' \notin \mathcal{E}$, $\exists a \in \mathcal{E}$ s.t. $a\mathcal{R}a'$, then $\forall \mathcal{E}' \in \mathcal{P}(\mathcal{A})$, $\mathcal{E} \succeq \mathcal{E}'$, and

3. if $\mathcal{E} \in \mathcal{CF}(\mathcal{F})$ and $\exists a' \in \mathcal{A} \setminus \mathcal{E}$ s.t. $\nexists a \in \mathcal{E}$ s.t. $a\mathcal{R}a'$, then $\exists \mathcal{E}' \in \mathcal{P}(\mathcal{A})$ s.t. not($\mathcal{E} \succeq \mathcal{E}'$).

Proof. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AF and $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$.

$\Rightarrow$ Assume that $\texttt{Ext}(\mathcal{F}) = \succeq_{max}$ and let us prove that the three conditions are satisfied.

1. Assume that $\mathcal{E} \in \mathcal{P}(\mathcal{A})$ and $\mathcal{E} \notin \mathcal{CF}(\mathcal{F})$. So, $\mathcal{E} \notin \texttt{Ext}(\mathcal{F})$, consequently, $\mathcal{E} \notin \succeq_{max}$. Thus, $\exists \mathcal{E}' \in \mathcal{P}(\mathcal{A})$ s.t. $\neg(\mathcal{E} \succeq \mathcal{E}')$.

2. Assume that $\mathcal{E} \in \mathcal{CF}(\mathcal{F})$ and that $\forall x' \notin \mathcal{E}$, $\exists x \in \mathcal{E}$ s.t. $x\mathcal{R}x'$. Thus, $\mathcal{E}$ is a stable extension of $(\mathcal{A}, \mathcal{R})$, which means that $\mathcal{E} \in \succeq_{max}$. Consequently, $\forall \mathcal{E}' \in \mathcal{P}(\mathcal{A}), \mathcal{E} \succeq \mathcal{E}'$.

3. Assume that $\mathcal{E} \in \mathcal{CF}(\mathcal{F})$ and $\exists x' \in \mathcal{A} \setminus \mathcal{E}$ s.t. $\nexists x \in \mathcal{E}$ and $x\mathcal{R}x'$. It is obvious that $\mathcal{E}$ is not a stable extension of $(\mathcal{A}, \mathcal{R})$, thus $\mathcal{E} \notin \texttt{Ext}(\mathcal{F})$. Since $\texttt{Ext}(\mathcal{F}) = \succeq_{max}$, it follows that $\mathcal{E} \notin \succeq_{max}$. Thus, $\exists \mathcal{E}' \in \mathcal{P}(\mathcal{A})$ s.t. not ($\mathcal{E} \succeq \mathcal{E}'$).

$\Leftarrow$ Assume that relation $\succeq$ satisfies the three conditions and let us prove that $\texttt{Ext}(\mathcal{F}) = \succeq_{max}$.

- Let $\mathcal{E}$ be a stable extension of $(\mathcal{A}, \mathcal{R})$ and let $\mathcal{E}' \in \mathcal{P}(\mathcal{A})$. From the second condition, $\mathcal{E} \succeq \mathcal{E}'$. Thus, $\mathcal{E} \in \succeq_{max}$.

- Assume that $\mathcal{E} \in \succeq_{max}$ and let us prove that $\mathcal{E} \in \texttt{Ext}(\mathcal{F})$. Thus, for all $\mathcal{E}' \in \mathcal{P}(\mathcal{A})$, $\mathcal{E} \succeq \mathcal{E}'$. From the first condition, it follows that $\mathcal{E}$ is conflict-free. Assume that $\mathcal{E} \notin \texttt{Ext}(\mathcal{F})$. Thus, $\exists x \notin \mathcal{E}$ and $\mathcal{E}$ does not attack x. From the third condition, $\exists \mathcal{E}' \in \mathcal{P}(\mathcal{A})$ s.t. not $\mathcal{E} \succeq \mathcal{E}'$. This contradicts the fact that $\mathcal{E} \in \succeq_{max}$.

□

Proposition 4.3.2. The relation $\succeq_s$ satisfies postulates $P1, P2$ and $P3$.

Proof. Let us show that the relation $\succeq_s$ satisfies postulates $P1, P2$ and $P3$. From the first condition of Definition 4.3.5, it is clear that postulate $P1$ is satisfied.

Let x and x' be two arguments. Since we assumed throughout the paper that there are no self-attacking arguments, then $\{x\}$ and $\{x'\}$ are conflict-free. Assume now that $x\mathcal{R}x'$, $\neg(x'\mathcal{R}x)$ and $\neg(x' > x)$. From the second condition of Definition 4.3.5, it follows that $\{x\} \succ_s \{x'\}$. Thus, $\succeq_s$ satisfies postulate $P2$.

Assume now that $x\mathcal{R}x'$ and $x' > x$. From the second condition of Definition 4.3.5, it follows that $\{x'\} \succeq_s \{x\}$. Also, $\neg(\{x\} \succeq_s \{x'\})$. Thus, $\{x'\} \succ_s \{x\}$. Consequently, postulate $P3$ is satisfied by $\succeq_s$. □

Theorem 4.3.2. The relation $\succeq_s$ generalises stable semantics.

Proof. Let us show that the relation $\succeq_s$ generalises stable semantics. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. Assume that $\nexists a, b \in \mathcal{A}$ s.t. $a\mathcal{R}b$ and $b > a$.

$\Rightarrow$ Assume that $\mathcal{E}' \in \succeq^s_{max}$ and let us show that $\mathcal{E}'$ is a stable extension of $(\mathcal{A}, \mathcal{R})$.

- Since $\mathcal{E}' \in \succeq^s_{max}$ then it is conflict-free.
- We will now prove that $\mathcal{E}'$ defends all its elements. Let us suppose that $(\exists a \in \mathcal{E}')$ $(\exists x \in \mathcal{A})$ s.t. $(x, a) \in \mathcal{R} \wedge (\nexists y \in \mathcal{E}')$ $(y, x) \in \mathcal{R}$. Since $\mathcal{E}'$ is conflict-free, then $x \notin \mathcal{E}'$. Let $\mathcal{E} = \{x\} \cup \{t \in \mathcal{E}' \mid (x, t) \notin \mathcal{R} \wedge (t, x) \notin \mathcal{R}\}$. It is clear the $\mathcal{E}$ is conflict-free since $\mathcal{E}$ is the union of two conflict-free sets which do not attack one another. Since $\mathcal{E}' \in \succeq_{max}$ then $\mathcal{E}' \succeq_s \mathcal{E}$. In particular, since $x \in \mathcal{E} \setminus \mathcal{E}'$, then $(\exists x' \in \mathcal{E}' \setminus \mathcal{E})$ s.t. $((x', x) \in \mathcal{R} \wedge (x, x') \notin >) \vee (x', x) \in >$. Since $(\nexists y \in \mathcal{E}')$ $(y, x) \in \mathcal{R}$, then it must be the case that $(x', x) \notin \mathcal{R}$ and $(x', x) \in >$. Since $x' \in \mathcal{E}'$ and $x' \notin \mathcal{E}$ then, with respect to definition of $\mathcal{E}$, from $x' \notin \mathcal{E}$ we have that $(x, x') \in \mathcal{R}$ or $(x', x) \in \mathcal{R}$. Since we have just seen that $(x', x) \notin \mathcal{R}$, it must be that $(x, x') \in \mathcal{R}$. Recall that we have $(x', x) \in >$. But we supposed that $(\nexists z, z' \in \mathcal{A})$ s.t. $(z, z') \in \mathcal{R}$ and $(z', z) \in >$. Contradiction. Thus, $\mathcal{E}'$ defends its arguments.
- We have just shown that $\mathcal{E}'$ is admissible, i.e. it is conflict-free and it defends all its arguments. We will now prove that $\mathcal{E}'$ attacks all arguments in $\mathcal{A} \setminus \mathcal{E}'$. Let $x \notin \mathcal{E}'$ be an argument and suppose that $(\nexists y \in \mathcal{E}')$ $(y, x) \in \mathcal{R}$. Either x attacks some argument of $\mathcal{E}'$ or not. If it is the case, i.e. if $(\exists a \in \mathcal{E}')$ s.t. $(x, a) \in \mathcal{R}$ then, since $\mathcal{E}'$ defends all its elements, it holds that $(\exists y \in \mathcal{E}')$ s.t. $(y, x) \in \mathcal{R}$. Contradiction. So, it must be that $(\nexists a \in \mathcal{E}')$ s.t. $(x, a) \in \mathcal{R}$. This means that $\mathcal{E} = \mathcal{E}' \cup \{x\}$ is conflict-free. According to Proposition 4.3.3, it holds that $\neg(\mathcal{E}' \succeq_s \mathcal{E})$. Contradiction with the fact that $\mathcal{E}' \in \succeq^s_{max}$.

So, $\mathcal{E}'$ is conflict-free and it attacks all arguments in $\mathcal{A} \setminus \mathcal{E}'$. This means that $\mathcal{E}'$ is a stable extension of the framework $(\mathcal{A}, \mathcal{R})$.

$\Leftarrow$ Let $\mathcal{E}'$ be a stable extension of the framework $(\mathcal{A}, \mathcal{R})$ and let us prove that $\mathcal{E}' \in \succeq^s_{max}$.

- Since $\mathcal{E}'$ is stable then it is conflict-free.
- We will prove that for an arbitrary conflict-free set of arguments $\mathcal{E}$ it holds that $\mathcal{E}' \succeq_s \mathcal{E}$. Let $\mathcal{E} \subseteq \mathcal{A}$ be a conflict-free set. If $\mathcal{E} \setminus \mathcal{E}' = \emptyset$ the proof is over. If it is not the case, let $x \in \mathcal{E} \setminus \mathcal{E}'$. Since $x \notin \mathcal{E}'$ and $\mathcal{E}'$ is a stable extension, then $(\exists x' \in \mathcal{E}')$ s.t. $(x', x) \in \mathcal{R}$. We supposed that $(\nexists z, z' \in \mathcal{A})$ s.t. $(z, z') \in \mathcal{R}$ and $(z', z) \in >$. Thus, $(x, x') \notin >$. Since $x \in \mathcal{E} \setminus \mathcal{E}'$ was arbitrary, it holds that $\mathcal{E}' \succeq_s \mathcal{E}$.
- From Proposition 4.3.3, it follows that $\mathcal{E}' \in \succeq^s_{max}$.

□

Proposition 4.3.4. The relation $\succeq_p$ satisfies postulates $P1, P2$ and $P3$.

Proof. Let us show that the relation $\succeq_p$ satisfies postulates P1, P2 and P3. The definition of $\succeq_p$ implies that P1 is ensured. Let us now suppose that for $x, x' \in \mathcal{A}$ we have $x\mathcal{R}x'$, $\neg(x'\mathcal{R}x)$ and $\neg(x' > x)$. Since there are no self-attacking arguments, both $\{x\}$ and $\{x'\}$ are conflict-free. From Definition 4.3.6, we obtain $\{x\} \succeq_p \{x'\}$ and $\neg(\{x'\} \succeq_p \{x\})$. Thus, $P2$ is verified. Let $x\mathcal{R}x'$ and $x' > x$. From the same definition, this time we have that $\neg(\{x\} \succeq_p \{x'\})$ and $\{x'\} \succeq_p \{x\}$. In other words, $\{x'\} \succ_p \{x\}$, which means that $P3$ is verified. □

Theorem 4.3.3. The relation $\succeq_p$ generalises preferred semantics.

Proof. We will prove that preferred extensions of $(\mathcal{A}, \mathcal{R})$ are exactly maximal elements of relation $\succeq_p$. Since we supposed that $(\nexists x, y \in \mathcal{A})$ s.t. $(x, y) \in \mathcal{R}$ $\wedge$ $(y, x) \in >$ then $\mathcal{E}' \succeq_p \mathcal{E}$ iff $(\forall x' \in \mathcal{E}')$ $(\forall x \in \mathcal{E})$ if $(x, x') \in \mathcal{R}$ then $(\exists y' \in \mathcal{E}')$ s.t. $(y, x) \in \mathcal{R}$.

$\Rightarrow$ Let $\mathcal{E}'$ be a preferred extension of $(\mathcal{A}, \mathcal{R})$.

- Since $\mathcal{E}'$ is a preferred extension then it is conflict-free.
- Let us prove that $\mathcal{E}' \in \succeq^p_{max}$. Suppose the contrary. This means that one of the following is true:

1. $(\exists \mathcal{E} \subseteq \mathcal{A})$ s.t. $\mathcal{E}$ is conflict-free and $\neg(\mathcal{E}' \succeq_p \mathcal{E})$
2. $(\exists \mathcal{E} \subseteq \mathcal{A})$ s.t. $\mathcal{E}$ is conflict-free $\wedge\ \mathcal{E}' \subsetneq \mathcal{E} \wedge (\forall \mathcal{E}'' \subseteq \mathcal{A})\ \mathcal{E} \succeq_p \mathcal{E}''$

Let (1) be the case. Since $\neg(\mathcal{E}' \succeq_p \mathcal{E})$ then $(\exists x' \in \mathcal{E}')(\exists x \in \mathcal{E})$ s.t. $(x, x') \in \mathcal{R} \wedge (\nexists y' \in \mathcal{E}')$ s.t. $(y', x) \in \mathcal{R}$. This leads to the conclusion that $\mathcal{E}'$ does not defend its arguments, thus it cannot be a preferred extension. Contradiction. So, it must be that (2) holds. Since $\mathcal{E}'$ is preferred and $\mathcal{E}' \subsetneq \mathcal{E}$ then $\mathcal{E}$ is not admissible. From the fact that $\mathcal{E}$ is conflict-free, one concludes that it does not defend its arguments. Thus, $(\exists x'' \in \mathcal{E}'' \setminus \mathcal{E}')$ s.t. $(\exists y \in \mathcal{A})$ s.t. $(y, x'') \in \mathcal{R} \wedge (\nexists z'' \in \mathcal{E}'')$ s.t. $(z'', y) \in \mathcal{R}$. Hence, $\neg(\mathcal{E}'' \succeq_p \{y\})$. Contradiction.

$\Leftarrow$ Let $\mathcal{E}' \in \succeq^p_{max}$. We will prove that $\mathcal{E}'$ is a preferred extension of Dung's argumentation framework $(\mathcal{A}, \mathcal{R})$.

- Since $\mathcal{E}' \in \succeq^p_{max}$ then it is conflict-free.
- Let us prove that $\mathcal{E}'$ defends all its arguments. Suppose not. This means that $(\exists y \in \mathcal{A})$ s.t. $(y, x') \in \mathcal{R} \wedge (\nexists z' \in \mathcal{E}')$ s.t. $(z', y) \in \mathcal{R}$. This means that $\neg(\mathcal{E}' \succeq_p \{y\})$. Contradiction.
- We have just seen that $\mathcal{E}'$ is admissible. Let us prove that $\mathcal{E}'$ is a preferred extension of $(\mathcal{A}, \mathcal{R})$. Suppose the contrary, i.e. $(\exists \mathcal{E} \subseteq \mathcal{A})$ s.t. $\mathcal{E}$ is a preferred extension and $\mathcal{E}' \subsetneq \mathcal{E}$. Since $\mathcal{E}' \in \succeq^p_{max}$ then $\mathcal{E} \notin \succeq^p_{max}$. On the other hand, since $\mathcal{E}$ is a preferred extension, then $\mathcal{E} \in \succeq^p_{max}$, as we have proved in the first part of this theorem. Contradiction.

□

Theorem 4.3.4. For any $(\mathcal{A}, \mathcal{R}, \geq)$, it holds that $\succeq^s_{max} \subseteq \succeq^p_{max}$.

Proof. We will prove that for any $(\mathcal{A}, \mathcal{R}, \geq)$, every pref-stable extension of this framework is a pref-preferred extension of that framework. In order to simplify the notation, we will write $x\mathcal{D}y$ instead of $(x\mathcal{R}y$ and not $(y > x))$ or $(y\mathcal{R}x$ and $x > y)$. It has been proved by Amgoud and Vesic (2010a) that if $\mathcal{E} \in \succeq^s_{max}$ then $\mathcal{E}$ is a Dung's stable extension of the framework $(\mathcal{A}, \mathcal{D})$. Then, the results by Dung (1995) imply that that $\mathcal{E}$ is a preferred extension of the framework $(\mathcal{A}, \mathcal{D})$. Now, we only have to prove that $\mathcal{E} \in \succeq^p_{max}$. It is obvious that $\mathcal{E} \in \mathcal{CF}$. Let $\mathcal{E}' \subseteq \mathcal{A}$. One can easily see that $\mathcal{E} \succeq_p \mathcal{E}'$. Let us prove that $\nexists \mathcal{E}''$ s.t. $\mathcal{E} \subsetneq \mathcal{E}''$ and $\forall \mathcal{E}'$, $\mathcal{E}'' \succeq_p \mathcal{E}'$. Suppose the contrary; this would mean that $\mathcal{E}''$ is admissible in $(\mathcal{A}, \mathcal{D})$ which contradicts the fact that $\mathcal{E}$ is a preferred extension of $(\mathcal{A}, \mathcal{D})$. Thus, it must be that $\mathcal{E} \in \succeq^p_{max}$, which ends the proof. □

Proposition 4.3.5. The equality $| \succeq^g_{max} | = 1$ holds.

Proof. In the proof of this proposition, we will use Properties 50 and 51 from the paper by Baroni and Giacomin (2007) which imply that for any argumentation framework $(\mathcal{A}, \mathcal{R})$, for any $x \in \mathcal{A}$, we have that $x \in \texttt{GE}$ iff $sd'(x, \texttt{GE})$, where GE is the standard notation for grounded extension which will be used throughout the proof and sd' is the notion of strong defense as defined in Definition 13 of the paper by Baroni and Giacomin (2007). Note that for any $a \in \mathcal{A}$, for any $\mathcal{A} \subseteq \mathcal{E}$, we have that $sd(a, \mathcal{E})$ iff ($\forall b \in \mathcal{A}$ if $b\mathcal{D}a$ then $\exists c \in \mathcal{E} \setminus \{a\}$ s.t. $c\mathcal{D}b$ and $sd(c, \mathcal{E} \setminus \{a\})$), where we use $x\mathcal{D}y$ as abbreviation for ($x\mathcal{R}y$ and not $(y > x)$) or ($y\mathcal{R}x$ and $x > y$). This proof will be based on the fact that we have $sd(a, \mathcal{E})$ in $(\mathcal{A}, \mathcal{R}, \geq)$ if and only if we have $sd'(a, \mathcal{E})$ in $(\mathcal{A}, \mathcal{D})$. Thus, when we write $sd(a, \mathcal{E})$, we refer to framework $(\mathcal{A}, \mathcal{R}, \geq)$, and when we use the function sd' and write $sd'(a, \mathcal{E})$, we refer to the corresponding framework $(\mathcal{A}, \mathcal{D})$. By using this equivalence, we will prove that any set $\mathcal{E} \subseteq \mathcal{A}$ is a pref-grounded extension of $(\mathcal{A}, \mathcal{R}, \geq)$ iff $\mathcal{E}$ is the grounded extension of $(\mathcal{A}, \mathcal{D})$.

$\Rightarrow$ Let $\mathcal{E}$ be the grounded extension of $(\mathcal{A}, \mathcal{D})$. It is obvious that $\mathcal{E} \in \mathcal{CF}$. Let $\mathcal{E}' \subseteq \mathcal{A}$. Since $\mathcal{E}$ is a grounded extension of $(\mathcal{A}, \mathcal{D})$, then from the results by Baroni and Giacomin (2007), we have $x \in \mathcal{E} \Rightarrow sd'(x, \mathcal{E})$. This means that we have $sd(x, \mathcal{E})$ in $(\mathcal{A}, \mathcal{R}, \geq)$. Thus, $sd(x, \mathcal{E}, \mathcal{E}')$ for any $\mathcal{E}'$, which means that $\forall \mathcal{E}'$, $\mathcal{E} \succeq_g \mathcal{E}'$. Let us prove that $\nexists \mathcal{E}'$ s.t. $\mathcal{E}' \in \mathcal{CF}$ and $\mathcal{E} \subsetneq \mathcal{E}'$ and $\forall \mathcal{E}''$, $\mathcal{E}' \succeq_g \mathcal{E}''$. Suppose the contrary. Suppose also that $\forall x \in \mathcal{E}'$, $sd(x, \mathcal{E}')$. This means that $\forall x \in \mathcal{E}'$, $sd'(x, \mathcal{E}')$ in $(\mathcal{A}, \mathcal{D})$. Thus, from Proposition 51 (Baroni and Giacomin, 2007), $\mathcal{E}' \subseteq \mathcal{E}$, since $\mathcal{E}$ is the grounded extension of $(\mathcal{A}, \mathcal{D})$. Contradiction, so it must be that $\exists x \in \mathcal{E}'$ s.t. $\neg sd(x, \mathcal{E}')$. Thus, $\exists y \in \mathcal{A}$ s.t. $\neg(\mathcal{E} \succeq_g \{y\})$. Contradiction, so we proved that $\mathcal{E} \in \succeq^g_{max}$.

$\Leftarrow$ Let $\mathcal{E} \in \succeq^g_{max}$. It is clear that $\forall x \in \mathcal{E}$, $sd(x, \mathcal{E})$ in $(\mathcal{A}, \mathcal{R}, \geq)$. Thus, $\forall x \in \mathcal{E}$, $sd'(x, \mathcal{E})$ in $(\mathcal{A}, \mathcal{D})$. From Proposition 51 (Baroni and Giacomin, 2007) we obtain $\mathcal{E} \subseteq \texttt{GE}$, where GE is the grounded extension of $(\mathcal{A}, \mathcal{D})$. Let us suppose that $\mathcal{E} \subsetneq \texttt{GE}$. In the first part of this proof, we have shown that the grounded extension of $(\mathcal{A}, \mathcal{D})$ is in $\succeq^g_{max}$. Contradiction, since we have supposed that $\mathcal{E} \in \succeq^g_{max}$ and we have $\mathcal{E} \subsetneq \texttt{GE}$. Thus, $\mathcal{E} = \texttt{GE}$.

This shows that $\mathcal{E} \in \succeq^g_{max}$ iff $\mathcal{E}$ is the grounded extension of the framework $(\mathcal{A}, \mathcal{D})$. Since it has been shown by Dung (1995) that every argumentation framework (without preferences) has exactly one grounded extension, we conclude that $\succeq^g_{max}$ has exactly one element. ☐

Proposition 4.3.6. The relation $\succeq_g$ satisfies postulates $P1, P2$ and $P3$.

Proof. It is easy to see that P1 is satisfied. Let $x\mathcal{R}x'$, $\neg(x'\mathcal{R}x)$ and $\neg(x' > x)$. From the definition of pref-grounded semantics, we have that $\{x\} \succeq_g \{x'\}$ since $sd(x, \{x\}, \{x'\})$. On the other hand, the fact that $\neg sd(x', \{x'\}, \{x\})$ implies that $\neg(\{x'\} \succeq_g \{x\})$. Thus, P2 is verified. Let us now prove that $\succeq_g$ verifies P3. Let $x\mathcal{R}x'$ and $x' > x$. In this case, we obtain $\neg sd(x, \{x\}, \{x'\})$ and $sd(x', \{x'\}, \{x\})$, which means that $\{x'\} \succ_g \{x\}$. □

Theorem 4.3.5. The relation $\succeq_g$ generalises grounded semantics.

Proof. Let $(\mathcal{A}, \mathcal{R}, \geq)$ be a PAF s.t. $\nexists x, y \in \mathcal{A}$ s.t. $x\mathcal{R}y$ and $y > x$. We show that the grounded extension of $(\mathcal{A}, \mathcal{R})$ is the only maximal element w.r.t. $\succeq_g$. Since there are no critical attacks, we can simplify Definition 4.3.7 which becomes: $sd(x, \mathcal{E}', \mathcal{E})$ iff $(\forall y \in \mathcal{E})$ (if $(y, x) \in \mathcal{R}$ then $(\exists z \in \mathcal{E}' \setminus \{x\})$ s.t. $((z, y) \in \mathcal{R} \wedge sd(z, \mathcal{E}' \setminus \{x\}, \mathcal{E})))$. In this particular case when no attacked argument is strictly preferred to its attacker, our definition of $sd(x, \mathcal{E})$ becomes exactly the same as Definition 13 in the work by Baroni and Giacomin (2007). Thus, using Proposition 50 and Proposition 51 of the same paper, we conclude that $x \in \texttt{GE}$ iff $sd(x, \texttt{GE})$, where GE is the grounded extension of the framework $(\mathcal{A}, \mathcal{R})$.
$\Rightarrow$ Let $\mathcal{E}'$ be the grounded extension of $(\mathcal{A}, \mathcal{R})$.

- Since $\mathcal{E}'$ is the grounded extension then it is conflict-free.
- We will prove that for an arbitrary conflict-free set $\mathcal{E} \subseteq \mathcal{A}$ it holds that $\mathcal{E}' \succeq_g \mathcal{E}$. Let $\mathcal{E} \subseteq \mathcal{A}$ be conflict-free. Since $\mathcal{E}'$ is the grounded extension then $x \in \mathcal{E}' \Rightarrow sd(x, \mathcal{E}')$. On the other hand, $(\forall x \in \mathcal{E}')$ $sd(x, \mathcal{E}')$ implies that $sd(x, \mathcal{E}', \mathcal{E})$. Thus, $\mathcal{E}' \succeq_g \mathcal{E}$. Since $\mathcal{E}$ was arbitrary, then $(\forall \mathcal{E} \subseteq \mathcal{A})$, $(\mathcal{E}' \succeq_g \mathcal{E}))$.
- We will now prove that $(\nexists \mathcal{E} \subseteq \mathcal{A})$ s.t. $\mathcal{E}$ is conflict-free and $\mathcal{E}' \subsetneq \mathcal{E}$ and $(\forall \mathcal{E}'' \subseteq \mathcal{A})$, $\mathcal{E} \succeq_g \mathcal{E}''$. Suppose the contrary. Suppose also that $(\forall x \in \mathcal{E})$ $sd(x, \mathcal{E})$. If this is the case, according to Proposition 51 in the paper by Baroni and Giacomin (2007), $\mathcal{E} \subseteq \texttt{GE}$. Contradiction. So, it must be that $(\exists x \in \mathcal{E})$ s.t. $\neg sd(x, \mathcal{E})$. Thus, $(\exists y \in \mathcal{A})$ s.t. $\neg sd(x, \mathcal{E}, \{y\})$. Consequently, $\neg(\mathcal{E} \succeq_g \{y\})$. Contradiction. So, we have proved that $\mathcal{E}' \in \succeq^g_{max}$.

$\Leftarrow$ Let $\mathcal{E}' \in \succeq^g_{max}$ and let us prove that $\mathcal{E}' = \texttt{GE}$. Since $(\forall x \in \mathcal{A})$ $\mathcal{E}' \succeq_g \{x\}$ then $(\forall x' \in \mathcal{E}')$ $sd(x', \mathcal{E}')$. From the fact that $(\forall x' \in \mathcal{E}')$ $sd(x', \mathcal{E}')$ and Proposition 51 (Baroni and Giacomin, 2007) we have that $\mathcal{E}' \subseteq \texttt{GE}$. Let us now prove that $\mathcal{E}' = \texttt{GE}$. Suppose not, i.e. suppose that $\mathcal{E}' \subsetneq \texttt{GE}$. We have proved in the first part of this theorem that $\texttt{GE} \in \succeq^g_{max}$. Contradiction, since we have supposed that $\mathcal{E}' \in \succeq^g_{max}$ and we have $\mathcal{E}' \subsetneq \texttt{GE}$. □

Theorem 4.3.6. For any $(\mathcal{A}, \mathcal{R}, \geq)$, if $\mathcal{E} \in \succeq^g_{max}$ then $\mathcal{E} \subseteq \bigcap_{\mathcal{E}_i \in \succeq^p_{max}} \mathcal{E}_i$.

Proof. Let us suppose that $\mathcal{E}$ is the pref-grounded extension of $(\mathcal{A}, \mathcal{R}, \geq)$. By using the same reasoning as in the proof of Proposition 4.3.5, we conclude that $\mathcal{E}$ is the grounded extension of the framework $(\mathcal{A}, \mathcal{D})$, where $x\mathcal{D}y$ is defined as $((x\mathcal{R}y$ and not $(y > x))$ or $(y\mathcal{R}x$ and $x > y))$. Dung (1995) has shown that the grounded extension of any argumentation framework is a subset of the intersection of all preferred extensions of that framework. Thus, in order to prove this property, it is sufficient to show that $\forall \mathcal{E}$, if $\mathcal{E} \in \succeq^p_{max}$, then $\mathcal{E}$ is a preferred extension of $(\mathcal{A}, \mathcal{D})$, since this will imply that the intersection of preferred extensions of $(\mathcal{A}, \mathcal{R}, \geq)$ is a subset of the intersection of pref-preferred extensions of $(\mathcal{A}, \mathcal{D})$.

Let $\mathcal{E} \in \succeq^p_{max}$. Obviously, $\mathcal{E} \in \mathcal{CF}$. Let us prove that $\mathcal{E}$ is admissible in $(\mathcal{A}, \mathcal{D})$. Let $a \in \mathcal{E}, a' \notin \mathcal{E}$ and $a'\mathcal{D}a$. Since we supposed that $\mathcal{E} \in \succeq^p_{max}$, then $\mathcal{E} \succeq_p \{a'\}$. Consequently, $\exists b \in \mathcal{E}$ s.t. $b\mathcal{D}a'$, so $\mathcal{E}$ is admissible in $(\mathcal{A}, \mathcal{D})$. Let us suppose that $\exists \mathcal{E}' \subseteq \mathcal{A}$, s.t. $\mathcal{E} \subsetneq \mathcal{E}'$ and $\mathcal{E}'$ is admissible in $(\mathcal{A}, \mathcal{D})$. Then, $\forall \mathcal{E}''$, we have $\mathcal{E}' \succeq_p \mathcal{E}''$. Consequently, from the second item of Definition 4.3.2, we have that $\mathcal{E} \notin \succeq^p_{max}$, contradiction. Thus, it must be that $\mathcal{E}$ is a preferred extension of $(\mathcal{A}, \mathcal{D})$. Since every pref-preferred extension of $(\mathcal{A}, \mathcal{R}, \geq)$ is a preferred extension of $(\mathcal{A}, \mathcal{D})$, then $\bigcap_{\mathcal{E}_i \in \succeq^p_{max}} \mathcal{E}_i \subseteq \bigcap_{\mathcal{E}_j \textit{ is a preferred extension of } (\mathcal{A},\mathcal{D})} \mathcal{E}_j$, which ends the proof of this property. □

Proposition 4.3.7. Let $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$. If $\succeq$ satisfies postulates $P4$ and $P5$, then it also satisfies postulates $P2$ and $P3$.

Proof. Let $x, x' \in \mathcal{A}$. Since there are no self-attacking arguments, then $\{x\}, \{x'\} \in \mathcal{CF}$. Let $x\mathcal{R}x'$, $\neg(x'\mathcal{R}x)$ and $\neg(x' > x)$. From the first part of Postulate 5 we have that $\{x\} \succeq \{x'\}$. From Postulate 4, we have $\neg(\{x'\} \succeq \{x\})$. Thus, Postulate 2 is verified. Let $x\mathcal{R}x'$ and $x' > x$. From Postulate 5, $\{x'\} \succeq \{x\}$. Furthermore, Postulate 4 implies $\neg(\{x\} \succeq \{x'\})$. In sum, $\{x'\} \succ \{x\}$, which means that Postulate 3 is verified. □

Proposition 4.3.8. $\succeq_s$ is a pref-stable relation.

Proof. To show that $\succeq_s$ is a pref-stable relation, we show that it satisfies postulates $P4, P5, P6$. Postulate 6 is satisfied since from the second item of the same definition, when comparing two sets $\mathcal{E}$ and $\mathcal{E}'$, common elements are not taken into account. The second condition of the definition of $\succeq_s$ is exactly the negation of the condition of Postulate 4. Since Postulate 5 implies the second item of this definition, then it is verified. $\square$

Theorem 4.3.7. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\succeq, \succeq' \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$. If $\succeq$ and $\succeq'$ are pref-stable relations, then $\succeq_{max} = \succeq'_{max}$.

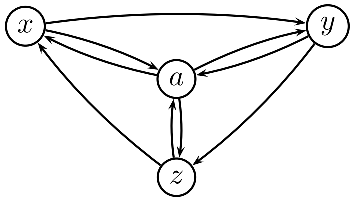
Proof. We prove that all pref-stable relations return the same set of extensions.
$\Rightarrow$ Let $\mathcal{E} \in \succeq_{max}$. We will prove that $\mathcal{E} \succeq'_{max}$. From Postulate 1, $\mathcal{E} \in \mathcal{CF}$. Let $\mathcal{E}' \subseteq \mathcal{A}$. If $\mathcal{E}'$ is not conflict-free then, from Postulate 1, $\mathcal{E} \succeq' \mathcal{E}'$. Else, from Postulate 6, $\mathcal{E} \succeq' \mathcal{E}'$ iff $\mathcal{E} \backslash \mathcal{E}' \succeq' \mathcal{E}' \backslash \mathcal{E}$. Let $\mathcal{E}_1 = \mathcal{E} \backslash \mathcal{E}'$ and $\mathcal{E}_2 = \mathcal{E}' \backslash \mathcal{E}$. $\mathcal{E}_1$ and $\mathcal{E}_2$ are disjunct conflict-free sets. If condition of Postulate 5 is satisfied for $\mathcal{E}_1$ and $\mathcal{E}_2$, then $\mathcal{E}_1 \succeq' \mathcal{E}_2$. Let us study the case when this condition is not satisfied. Condition of Postulate 4 is not satisfied since $\mathcal{E} \in \succeq_{max}$. Thus, it must be that $(\exists x' \in \mathcal{E}_2)$ s.t. $(\nexists x \in \mathcal{E}_1)((x, x') \in \mathcal{R} \wedge (x', x) \notin >) \vee ((x', x) \in \mathcal{R} \wedge (x, x') \in >)$ and $(\exists x \in \mathcal{E}_1)(x, x') \in >$. Let $X = \{x \in \mathcal{E}_1 | (x, x') \in >\}$. X is conflict-free. From Postulate 4, $\neg(\mathcal{E}_1 \setminus X \succeq \{x'\})$. Postulate 6 implies that $\neg(\mathcal{E}_1 \backslash X \cup (X \cup (\mathcal{E} \cap \mathcal{E}')) \succeq \{x'\} \cup (X \cup (\mathcal{E} \cap \mathcal{E}')))$, i.e. $\neg(\mathcal{E} \succeq \{x'\} \cup (X \cup (\mathcal{E} \cap \mathcal{E}')))$. Contradiction with $\mathcal{E} \in \succeq_{max}$. Thus, condition of Postulate 5 is satisfied for $\mathcal{E}_1$ and $\mathcal{E}_2$, and $\mathcal{E}_1 \succeq' \mathcal{E}_2$. Consequently, $\mathcal{E} \succeq' \mathcal{E}'$. This means that $\mathcal{E} \in \succeq'_{max}$.

$\Leftarrow$ In the first part of proof, we showed that for all pref-stable relations $\succeq_1, \succeq_2$, it holds that if $\mathcal{E} \in \succeq^1_{max}$ then $\mathcal{E} \in \succeq^2_{max}$. Contraposition of this rule gives specifies that if $\mathcal{E} \notin \succeq^2_{max}$ then $\mathcal{E} \notin \succeq^1_{max}$. Since this was proved for arbitrary relations which satisfy $P1, P4, P5$ and $P6$, we conclude: if $\mathcal{E} \notin \succeq'_{max}$ then $\mathcal{E} \notin \succeq_{max}$. $\square$

Proposition 4.3.9. There exists no transitive relation which generalises stable semantics and satisfies postulates $P1$ and $P5$.

Proof. Let us suppose that there exists a transitive relation which satisfies $P1$ and $P5$ and which generalises stable semantics. Let us now consider the framework depicted in Figure A.1. Suppose that attacks are as depicted and that $\geq = \{(w, w) \mid w \in \mathcal{A}\}$. From $P1$, we have that for any $\mathcal{E}' \notin \mathcal{CF}$, it holds that $\{x\} \succeq \mathcal{E}'$. From $P5$, $\{x\} \succeq \{a\}$, $\{a\} \succeq \{x\}$, $\{x\} \succeq \{y\}$, $\{y\} \succeq \{z\}$, $\{x\} \succeq \emptyset$. From those relations and transitivity of $\succeq$, we have $\{x\} \succeq \{x\}$ and $\{x\} \succeq \{z\}$. Thus, $\{x\} \in \succeq_{max}$. This contradicts the fact that $\succeq$ generalises stable semantics, since $\{x\}$ is not a stable extension of the framework $(\mathcal{A}, \mathcal{R})$. $\square$

Figure A.1: No transitive relation generalises stable semantics and verifies $P1$ and $P5$.



Theorem 4.3.8. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF. Any pref-stable relation $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$ generalises stable semantics.

Proof. We will show that extensions of $(\mathcal{A}, \mathcal{R})$ coincide with maximal elements of $\succeq$ for any preference-based argumentation framework $\mathcal{T}$, such that $(\nexists a, b \in \mathcal{A})(a, b) \in \mathcal{R} \wedge (b, a) \in >$. Let $\mathtt{Ext}(\mathcal{F})$ denote stable extensions of $\mathcal{F} = (\mathcal{A}, \mathcal{R})$.

$\Rightarrow$ Let $\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$. We prove that $\mathcal{E} \in \succeq_{max}$. Let $\mathcal{E}' \in \mathcal{P}(\mathcal{A})$. If $\mathcal{E}' \notin \mathcal{CF}$ then, from Postulate 1, $\mathcal{E} \succeq \mathcal{E}'$. Let $\mathcal{E}' \in \mathcal{CF}$. Since $\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$ then $(\forall x' \in \mathcal{E}' \setminus \mathcal{E})(\exists x \in \mathcal{E} \setminus \mathcal{E}')(x, x') \in \mathcal{R}$. We supposed $(\nexists a, b \in \mathcal{A})(a, b) \in \mathcal{R} \wedge (b, a) \in >$. Thus, from Postulate 5, $\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}$. Now, Postulate 6 implies $\mathcal{E} \succeq \mathcal{E}'$. Since $\mathcal{E}'$ was arbitrary, then $\mathcal{E} \in \succeq_{max}$.

$\Leftarrow$ Let $\mathcal{E} \in \succeq_{max}$. We will show that $\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$. From Postulate 1, $\mathcal{E} \in \mathcal{CF}$. Let $x' \notin \mathcal{E}$. Since $\mathcal{E} \in \succeq_{max}$ then it must be $\mathcal{E} \succeq \{x'\}$. From Postulate 4, $(\exists x \in \mathcal{E})(x, x') \in \mathcal{R} \vee (x, x') \in >$. If $(\exists x \in \mathcal{E})(x, x') \in \mathcal{R}$, the proof is over. Let us suppose the contrary. Then $(\nexists x \in \mathcal{E})(x, x') \in \mathcal{R}$. Let $X = \{x \in \mathcal{E} | x > x'\}$. From Postulate 4, $\neg(\mathcal{E} \setminus X \succeq \{x'\})$. This fact and Postulate 6 imply $\neg(\mathcal{E} \succeq (X \cup \{x'\}))$. Contradiction with $\mathcal{E} \in \succeq_{max}$. Thus, $\mathcal{E} \in \mathtt{Ext}(\mathcal{F})$. □

Proposition 4.3.10. $\succeq_{gn}$ is a pref-stable relation.

Proof. It is easy to show that relation $\succeq_{gn}$ satisfies $P1, P4, P5$ and $P6$. Postulate 1 is satisfied since from the first item of the definition of $\succeq_{gn}$, any conflict-free set is preferred to any conflicting set. Postulate 6 is satisfied since from the second item of the same definition, when comparing two sets $\mathcal{E}$ and $\mathcal{E}'$, common elements are not taken into account. Postulate 4 implies that the second item of Definition 4.3.10 is not satisfied. Postulate 5 is trivially verified. □

Proposition 4.3.11. $\succeq_{sp}$ is a pref-stable relation.

Proof. Let us show that $\succeq_{sp}$ satisfies $P1, P4, P5$ and $P6$. We see from the first item of Definition 4.3.11 that all (conflict-free and non conflict-free) sets are better than non conflict-free sets. A non conflict-free set, however, cannot be better than conflict-free set. Thus, Postulate 1 is satisfied. Postulates 4, 5 and 6 are verified for the same reasons as in the case of relation $\succeq_{gn}$. □

Theorem 4.3.9. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}, \mathcal{E}' \in \mathcal{P}(\mathcal{A})$. Let $\succeq$ be a pref-stable relation.

- If $\mathcal{E} \succeq_{gn} \mathcal{E}'$ then $\mathcal{E} \succeq \mathcal{E}'$.
- If $\mathcal{E} \succeq \mathcal{E}'$ then $\mathcal{E} \succeq_{sp} \mathcal{E}'$.

Proof. We will show that for any relation $\succeq$ which satisfies $P1, P4, P5$ and $P6$, we have that if $\mathcal{E} \succeq_{gn} \mathcal{E}'$ then $\mathcal{E} \succeq \mathcal{E}'$ and if $\mathcal{E} \succeq \mathcal{E}'$ then $\mathcal{E} \succeq_{sp} \mathcal{E}'$.

- Let $\mathcal{E} \succeq_{gn} \mathcal{E}'$. This means that $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$. If $\mathcal{E}' \notin \mathcal{CF}(\mathcal{T})$, then from Postulate 1, $\mathcal{E} \succeq \mathcal{E}'$. We study the case when $\mathcal{E}' \in \mathcal{CF}(\mathcal{T})$. From Postulate 6, we have $\mathcal{E} \succeq \mathcal{E}'$ iff $\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}$. From Definition 4.3.10 and Postulate 5, $\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}$. Thus, $\mathcal{E} \succeq \mathcal{E}'$.
- If $\mathcal{E}, \mathcal{E}' \notin \mathcal{CF}(\mathcal{T})$ then, Definition 4.3.11 implies $\mathcal{E} \succeq_{sp} \mathcal{E}'$. Case $\mathcal{E} \notin \mathcal{CF}(\mathcal{T})$, $\mathcal{E}' \in \mathcal{CF}(\mathcal{T})$ is not possible because of Postulate 1. If $\mathcal{E} \in \mathcal{CF}(\mathcal{T}), \mathcal{E}' \notin \mathcal{CF}(\mathcal{T})$, then from Definition 4.3.11, $\mathcal{E} \succeq_{sp} \mathcal{E}'$. In the non-trivial case, when $\mathcal{E}, \mathcal{E}' \in \mathcal{CF}(\mathcal{T})$, from Postulate 6, $\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E}$. Suppose that $\neg(\mathcal{E} \setminus \mathcal{E}' \succeq_{sp} \mathcal{E}' \setminus \mathcal{E})$. Now, Definition 4.3.11 implies $(\exists x' \in \mathcal{E}' \setminus \mathcal{E})(\nexists x \in \mathcal{E} \setminus \mathcal{E}')$ s.t. $(x, x') \in >$ or $(x, x') \in \mathcal{R} \wedge (x', x) \notin >$. From this fact and Postulate 4, it holds that $\neg(\mathcal{E} \setminus \mathcal{E}' \succeq \mathcal{E}' \setminus \mathcal{E})$. Contradiction.

□

Theorem 4.3.10. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\succeq$ be a pref-stable relation.
$\mathcal{E} \in \succeq_{max}$ iff:

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$, and
- $\forall a' \in \mathcal{A} \setminus \mathcal{E}$, $\exists a \in \mathcal{E}$ such that $(a\mathcal{R}a'$ and $\text{not}(a' > a))$ or $(a'\mathcal{R}a$ and $a > a')$.

Proof. We will now prove that a set is a pref-stable extension iff it is conflict-free and its arguments win in all conflicts with exterior ones. Throughout the proof, we will use notation $\succeq^{gn}_{max}$ to refer to the set of maximal elements w.r.t. relation $\succeq_{gn}$.

Since both relations $\succeq$ and $\succeq_{gn}$ verify Postulates 1, 4, 5 and 6, then from Theorem 4.3.7, $\succeq_{max}=\succeq^{gn}_{max}$. This means that it is sufficient to prove that $\mathcal{E} \in\succeq^{gn}_{max}$ iff the two conditions of theorem are satisfied.

$\Rightarrow$ Let $\mathcal{E} \in\succeq^{gn}_{max}$. Since $\mathcal{E}$ is a pref-extension, according to Proposition 4.3.1, $\mathcal{E} \in \mathcal{CF}$. Let $x' \in \mathcal{A} \setminus \mathcal{E}$. We supposed that $(\nexists a \in \mathcal{A})$ s.t. $(a, a) \in \mathcal{R}$, so it must be that $\{x'\}$ is conflict-free. Since $\mathcal{E} \in\succeq^{gn}_{max}$, it holds that $\mathcal{E} \succeq_{gn} \{x'\}$. Since $\mathcal{E}$ and $\{x'\}$ are conflict-free, Definition 4.3.10 implies $(\exists x \in \mathcal{E})$ s.t. $(((x, x') \in \mathcal{R} \wedge (x', x) \notin>) \vee ((x', x) \in \mathcal{R} \wedge (x, x') \in>))$.

$\Leftarrow$ Let $\mathcal{E}$ be a conflict-free set and let $(\forall x' \in \mathcal{A}\setminus\mathcal{E})$ $(\exists x \in \mathcal{E})$ s.t. $(((x, x') \in \mathcal{R}$ $\wedge$ $(x', x) \notin>) \vee ((x', x) \in \mathcal{R} \wedge (x, x') \in>))$. Let us prove that $\mathcal{E} \in\succeq^{gn}_{max}$.

- Since $\mathcal{E} \in \mathcal{CF}$ then for every non conflict-free set $\mathcal{E}'$ it holds that $\mathcal{E} \succeq_{gn} \mathcal{E}'$.
- Let $\mathcal{E}' \subseteq \mathcal{A}$ be an arbitrary conflict-free set of arguments. If $\mathcal{E}' \subseteq \mathcal{E}$, the second condition of theorem is trivially satisfied. Else, let $x' \in \mathcal{E}' \setminus \mathcal{E}$. From what we supposed, we have that $(\exists x \in \mathcal{E} \setminus \mathcal{E}')$ s.t. $((x, x') \in \mathcal{R}$ $\wedge$ $(x', x) \notin>)$ or $((x', x) \in \mathcal{R} \wedge (x, x') \in>)$. Thus, $\mathcal{E} \succeq_{gn} \mathcal{E}'$.

From those two items, we have that $\mathcal{E} \in\succeq^{gn}_{max}$. ☐

Theorem 4.3.11. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF, let $\succeq_p$ be a relation from Definition 4.3.6 and let $\succeq^{p}_{max}$ be the set of maximal elements of $\mathcal{T}$ w.r.t. that relation. Then, $\mathcal{E} \in\succeq^{p}_{max}$ iff:

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$, and
- $(\forall a' \in \mathcal{E}')$ $(\forall a \in \mathcal{A} \setminus \mathcal{E}')$ if $(((a, a') \in \mathcal{R} \wedge (a', a) \notin>)$ or $((a', a) \in \mathcal{R}$ $\wedge$ $(a, a') \in>))$ then $(\exists b' \in \mathcal{E}')$ s.t. $((b', a) \in \mathcal{R}$ and $(a, b') \notin>)$ or $((a, b') \in \mathcal{R}$ and $b' > a)$, and
- $\mathcal{E}'$ is a maximal set (w.r.t. set inclusion) which satisfies previous two items.

Proof. $\Rightarrow$ Let us suppose that $\mathcal{E}' \in\succeq^{p}_{max}$ and let us prove that the three conditions stated in this theorem are satisfied.

- $\mathcal{E}'$ is conflict-free.
- Let $a \notin \mathcal{E}'$ be an arbitrary argument which does not belong to $\mathcal{E}'$. Since $\{a\}$ is conflict-free, then $\mathcal{E}' \succeq_p \{a\}$. This means that $(\forall a' \in \mathcal{E}')$ if $(((a, a') \in \mathcal{R} \wedge (a', a) \notin >)$ or $((a', a) \in \mathcal{R} \wedge (a, a') \in >))$ then $(\exists b' \in \mathcal{E}')$ s.t. $((b', a) \in \mathcal{R}$ and $(a, b') \notin >)$ or $((a, b') \in \mathcal{R}$ and $b' > a)$.
- Let us suppose that $(\exists \mathcal{E}'' \subseteq \mathcal{A})$ s.t. $\mathcal{E}''$ is conflict-free and $\mathcal{E}' \subsetneq \mathcal{E}''$ and $(\forall a' \in \mathcal{E}'')$ $(\forall a \in \mathcal{A} \setminus \mathcal{E}'')$ if $(((a, a') \in \mathcal{R} \wedge (a', a) \notin >)$ or $((a', a) \in \mathcal{R}$ $\wedge$ $(a, a') \in >))$ then $(\exists b' \in \mathcal{E}'')$ s.t. $((b', a) \in \mathcal{R}$ and $(a, b') \notin >)$ or $(a, b') \in \mathcal{R}$ and $b' > a)$. We will prove that this means that $\mathcal{E}' \notin \succeq^p_{max}$.
  - $\mathcal{E}''$ is conflict-free, trivial.
  - We will now prove that $(\forall \mathcal{E} \subseteq \mathcal{A})$ if $\mathcal{E}$ is conflict-free then $\mathcal{E}'' \succeq_p \mathcal{E}$. Let $\mathcal{E} \subseteq \mathcal{A}$ and let $a \in \mathcal{E}$. Let $a \notin \mathcal{E}''$. Then, from what we supposed, we have that if $(((a, a') \in \mathcal{R} \wedge (a', a) \notin >)$ or $((a', a) \in \mathcal{R} \wedge (a, a') \in >))$ then $(\exists b' \in \mathcal{E}')$ s.t. $((b', a) \in \mathcal{R}$ and $(a, b') \notin >)$ or $((a, b') \in \mathcal{R}$ and $b' > a)$. This means that $\mathcal{E}'' \succeq_p \mathcal{E}$. Let $a \in \mathcal{E}''$. Since $\mathcal{E}''$ is conflict-free, then the condition in question is trivially satisfied. In this case, also $\mathcal{E}'' \succeq_p \mathcal{E}$.

  The two previous items imply that $\mathcal{E}' \notin \succeq^p_{max}$. Contradiction.

$\Leftarrow$ Let us suppose that $\mathcal{E}' \subseteq \mathcal{A}$ satisfies three conditions given in the theorem and let us prove that $\mathcal{E}' \in \succeq^p_{max}$.

- $\mathcal{E}'$ is conflict-free. Trivial.
- Let $\mathcal{E} \subseteq \mathcal{A}$ be an arbitrary conflict-free set of arguments. Let us prove that $\mathcal{E}' \succeq_p \mathcal{E}$. Let $a \in \mathcal{E}$ and $(((a, a') \in \mathcal{R} \wedge (a', a) \notin >)$ or $((a', a) \in \mathcal{R} \wedge (a, a') \in >))$. Since $\mathcal{E}'$ is conflict-free then $a \notin \mathcal{E}'$. From the second item it holds that $(\exists b' \in \mathcal{E}')$ s.t. $((b', a) \in \mathcal{R}$ and $(a, b) \notin >)$ or $((a, b') \in \mathcal{R}$ and $b' > a)$. Therefore, $\mathcal{E}' \succeq_p \mathcal{E}$.
- Let us suppose that there exists $\mathcal{E}'' \subseteq \mathcal{A}$ such that $\mathcal{E}''$ is conflict-free and $\mathcal{E}' \subsetneq \mathcal{E}''$ and $(\forall \mathcal{E} \subseteq \mathcal{A})$, $\mathcal{E}'' \succeq_p \mathcal{E}$. We will prove that this is in contradiction with the third item of the theorem.
  - It is obvious that $\mathcal{E}''$ is conflict-free.
  - Let $a \in \mathcal{A} \setminus \mathcal{E}''$. Since $(\forall \mathcal{E} \subseteq \mathcal{A})$ $\mathcal{E}'' \succeq_p \mathcal{E}$, then $\mathcal{E}'' \succeq_p \{a\}$. This means that $(\forall a' \in \mathcal{E}')$ if $(((a, a') \in \mathcal{R} \wedge (a', a) \notin >)$ or $((a', a) \in \mathcal{R}$ $\wedge$ $(a, a') \in >))$ then $(\exists b' \in \mathcal{E}'')$ s.t. $((b', a) \in \mathcal{R}$ and $(a, b') \notin >)$ or $((a, b') \in \mathcal{R}$ and $b' > a)$.

Since $\mathcal{E}''$ satisfies first and second item of this theorem and $\mathcal{E}' \subsetneq \mathcal{E}''$ then $\mathcal{E}'$ does not satisfy the third item of the theorem. Contradiction since we supposed that $\mathcal{E}'$ satisfies all the three items.

□

Theorem 4.3.12. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF, let $\succeq_g$ be a relation from Definition 4.3.8 and let $\succeq^g_{max}$ be the set of maximal elements of $\mathcal{T}$ w.r.t. that relation. Then, $\mathcal{E} \in \succeq^g_{max}$ iff:

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$, and
- $(\forall a \in \mathcal{E})$ $sd(a, \mathcal{E})$ and
- $\mathcal{E}$ is a maximal set (w.r.t. set inclusion) which satisfies previous two items.

Proof. $\Rightarrow$ Let us suppose that $\mathcal{E}' \in \succeq^g_{max}$. We will prove that $\mathcal{E}'$ satisfies the three items of theorem.

- $\mathcal{E}'$ is conflict-free.
- Let $a' \in \mathcal{E}'$ and $a \in \mathcal{A}$. Since $\mathcal{E}' \in \succeq^g_{max}$, then $sd(a', \mathcal{E}', \{a\})$. Since a' was arbitrary, we have $sd(a', \mathcal{E}')$.
- Let us suppose that $(\exists \mathcal{E}'' \in \mathcal{A})$ s.t. $\mathcal{E}''$ is conflict-free and $\mathcal{E}' \subsetneq \mathcal{E}''$ and $\mathcal{E}''$ satisfies the first two items. In that case:
  - $\mathcal{E}''$ is conflict-free.
  - Since $(\forall x'' \in \mathcal{E}'')$ $sd(x'', \mathcal{E}'')$ then $(\forall \mathcal{E} \subseteq \mathcal{A})$ $\mathcal{E}'' \succeq_g \mathcal{E}$.

  From the two previous items, we see that $\mathcal{E}' \notin \succeq^g_{max}$, contradiction.

$\Leftarrow$ Let us suppose that the three conditions of the theorem are satisfied by $\mathcal{E}' \subseteq \mathcal{A}$ and let us prove that $\mathcal{E}' \in \succeq^g_{max}$.

- $\mathcal{E}'$ is conflict-free, trivial.
- Let $\mathcal{E} \subseteq \mathcal{A}$ be an arbitrary conflict-free set and let us prove that $(\forall x' \in \mathcal{E}')$ $sd(x', \mathcal{E}', \mathcal{E})$. Since we supposed that $(\forall x' \in \mathcal{E}')$ $sd(x', \mathcal{E}', \mathcal{A})$, which means that set $\mathcal{E}'$ strongly defends all its elements against attacks of all other elements, then $sd(x', \mathcal{E}', \mathcal{E})$. Thus, $\mathcal{E} \succeq_g \mathcal{E}'$.
- Let us suppose that $(\exists \mathcal{E}'' \subseteq \mathcal{A})$ s.t. $\mathcal{E}''$ is conflict-free and $\mathcal{E}' \subsetneq \mathcal{E}''$ and $\mathcal{E}'' \in \succeq_{max}$. In that case, it can be proven that $\mathcal{E}''$ satisfies two first items of this theorem:

- $\mathcal{E}''$ is conflict-free.
- Let $x'' \in \mathcal{E}''$ and let us prove that $sd(x'', \mathcal{E}'')$. Let $y \in \mathcal{A}$ and let us prove that $sd(x'', \mathcal{E}'', \{y\})$. This follows from the fact that $\{y\}$ is conflict-free and $\mathcal{E}'' \in \succeq^g_{max}$.

Since $\mathcal{E}''$ is conflict-free, $\mathcal{E}''$ satisfies the two first items of this theorem, and $\mathcal{E}' \subsetneq \mathcal{E}''$, then $\mathcal{E}'$ does not satisfy the third item of this theorem. Contradiction.

□

Theorem 4.3.13. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and let $\mathcal{R}' = \{(a, b) \mid a, b \in \mathcal{A},\ (a\mathcal{R}b \text{ and not}(b > a)) \text{ or } (b\mathcal{R}a \text{ and } a > b)\}$, and let $\succeq^g_{max}$ be the set of all maximal elements w.r.t. $\succeq_g$. Then: $\succeq^g_{max}$ contains only one set which is exactly the grounded extension of $(\mathcal{A}, \mathcal{R}')$.

Proof. For an argument a and a set $\mathcal{E}$, let $sd'(a, \mathcal{E})$ be defined as follows: $sd'(a, \mathcal{E})$ iff $\forall b \in \mathcal{A}$ if $b\mathcal{R}'a$ then $\exists c \in \mathcal{E} \setminus \{a\}$ s.t. $c\mathcal{R}'b$ and $sd'(c, \mathcal{E} \setminus \{a\})$. From Theorem 4.3.12 and from definition of $\mathcal{R}'$, we see that for $\mathcal{E} \subseteq \mathcal{A}$ we have $\mathcal{E} \in \succeq^g_{max}$ iff:

- $\mathcal{E} \in \mathcal{CF}(\mathcal{T})$, and
- $(\forall a \in \mathcal{E})\ sd'(a, \mathcal{E})$ and
- $\mathcal{E}$ is a maximal set (w.r.t. set inclusion) which satisfies previous two items.

In the rest of the proof, we will show that $\mathcal{E} \subseteq \mathcal{A}$ is a grounded extension of $(\mathcal{A}, \mathcal{R}')$ iff the three previous conditions are verified. Let $\texttt{GE}$ be the grounded extension of $(\mathcal{A}, \mathcal{R}')$. Using Proposition 50 and Proposition 51 (Baroni and Giacomin, 2007), we conclude that $\forall x \in \mathcal{A}$, $x \in \texttt{GE}$ iff $sd'(x, \texttt{GE})$. We will now prove that $\mathcal{E}'$ verifies the three conditions above iff $\mathcal{E}' = \texttt{GE}$.

⇒ Let $\mathcal{E}' = \texttt{GE}$.

- Since $\mathcal{E}'$ is the grounded extension then it is conflict-free.
- Since $\mathcal{E}'$ is the grounded extension, from Propositions 50 and 51 (Baroni and Giacomin, 2007), it strongly defends all its elements, i.e. $\forall a \in \mathcal{E}'$, $sd'(a, \mathcal{E}')$.

- Let us suppose that the third condition is not verified. This would mean that $\exists a' \in \mathcal{A} \setminus \mathcal{E}'$ s.t. $sd'(a, \mathcal{E})$. From Propositions 50 and 51 (Baroni and Giacomin, 2007) and the fact that $\mathcal{E}' = \texttt{GE}$, $\forall x \in \mathcal{A}$, $x \in \mathcal{E}'$ iff $sd(x, \mathcal{E}')$. Since $a' \notin \mathcal{E}'$ then $\neg sd'(a, \mathcal{E})$. Contradiction.

$\Leftarrow$ Let the three conditions be verified and let us prove that $\mathcal{E}' = \texttt{GE}$. From $(\forall a \in \mathcal{E}')$ $sd'(a, \mathcal{E}')$ and Proposition 51 (Baroni and Giacomin, 2007) we have that $\mathcal{E}' \subseteq \texttt{GE}$. Let us now prove that $\mathcal{E}' = \texttt{GE}$. Suppose not, i.e. suppose that $\mathcal{E}' \subsetneq \texttt{GE}$. We have proved in the first part of this theorem that GE verifies the three conditions stated above. Contradiction, since we have supposed that $\mathcal{E}'$ is a maximal set verifying the first two conditions, while GE verifies both of them. □

Proposition 4.4.2. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a basic PAF s.t. $\mathcal{R}$ is irreflexive and $\geq$ is a linear order.

- Stable, preferred and grounded extensions of $\mathcal{T}$ coincide.
- $\mathcal{T}$ has exactly one stable extension.
- If $|\mathcal{A}| = n$, then this extension is computed in $\mathcal{O}(n^2)$ time.

Proof. Let us consider the following algorithm.

```
input:
A: set of arguments
R: attack relation
>=: preference relation

output:
in: the only stable/preferred/grounded ext.
out: rejected arguments w.r.t. those semantics

/* Put all arguments in und. */
in = {};
out = {};
und = A;

/* While und is not empty,
sort arguments from und to in and out. */
while (not (und == {}) {
```

```
/* Select the best argument in und,
and move it to in. */
let a be the only argument in the set
{x in und | for all x' in und, x > x};
in = in union {a};
und = und - {a};

/* Since a is accepted, all arguments being
in conflict with it must be rejected. */
del = {x in und | x R a  or  a R x};
out = out union del;
und = und - del;
}
```

Let us prove that *in* is a stable extension of $\mathcal{T}$. It is clear that *in* is conflict-free. Let $x' \notin in$. From the previous algorithm, it is easy to see that there exists $x \in in$ s.t. $x > x'$ and ($x\mathcal{R}x'$ or $x'\mathcal{R}x$). In other words, $x\mathcal{R}'x'$. Thus, *in* is a stable extension of $\mathcal{T}$.

It has been proved by Dung (1995) that every stable extension is a preferred and a complete extension. Thus, *in* is a preferred and complete extension of $\mathcal{T}$.

Let us prove that *in* is the only complete extension. Suppose that $\mathcal{E} \subseteq \mathcal{A}$, with $\mathcal{E} \neq in$ is another complete extension. Since none of the arguments of *in* is attacked (w.r.t. $\mathcal{R}'$), it is clear that every complete extension must contain those arguments, i.e. $in \subsetneq \mathcal{E}$. But, since *in* is a stable extension, it is maximal conflict-free set, contradiction. So, we have shown that *in* is the only complete extension.

It has been shown by Dung (1995) that grounded extension is exactly the intersection of all complete extensions. Hence, *in* is the grounded extension of $\mathcal{T}$.

Let us now prove that *in* is the only stable and the only preferred extension. Suppose not, thus there exists another stable or preferred extension $\mathcal{E}$, such that $\mathcal{E} \neq in$. Since we supposed that $\mathcal{E}$ is stable or preferred, then $\mathcal{E}$ is for sure complete (Dung, 1995). But we have already shown that *in* was a unique complete extension, contradiction. Thus, *in* is the unique stable and preferred extension of $\mathcal{T}$.

The while loop is executed at most n times, where n is the number of arguments, and its execution contains at most n comparisons. Thus, algorithm's time complexity is $\mathcal{O}(n^2)$. □

Proposition 4.5.1. Let Σ be a propositional knowledge base and ($\mathtt{Arg}(\Sigma)$, Undercut) the argumentation framework built from Σ.

- For any consistent set $\mathcal{S} \subseteq \Sigma$, $\mathcal{S} = \mathtt{Base}(\mathtt{Arg}(\mathcal{S}))$.
- The function $\mathtt{Base} : \mathtt{Arg}(\Sigma) \rightarrow \Sigma$ is surjective.
- For any $\mathcal{E} \subseteq \mathtt{Arg}(\Sigma)$, $\mathcal{E} \subseteq \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$.
- The function $\mathtt{Arg} : \Sigma \rightarrow \mathtt{Arg}(\Sigma)$ is injective.

Proof.

- We show that $x \in \mathcal{S}$ iff $x \in \mathtt{Base}(\mathtt{Arg}(\mathcal{S}))$ where $\mathcal{S}$ is a consistent subset of Σ.
  $\Rightarrow$ Let $x \in \mathcal{S}$. Since $\mathcal{S}$ is consistent, then the set $\{x\}$ is consistent as well. Thus, $(\{x\}, x) \in \mathtt{Arg}(\mathcal{S})$. Consequently, $x \in \mathtt{Base}(\mathtt{Arg}(\mathcal{S}))$.
  $\Leftarrow$ Assume that $x \in \mathtt{Base}(\mathtt{Arg}(\mathcal{S}))$. Thus, $\exists a \in \mathtt{Arg}(\mathcal{S})$ s.t. $x \in \mathtt{Supp}(a)$. From the definition of argument, $\mathtt{Supp}(a) \subseteq \mathcal{S}$. Consequently, $x \in \mathcal{S}$.

- Let us show that the function $\mathtt{Base}$ is surjective. Let $\mathcal{S} \subseteq \Sigma$. From the first item of this property, the equality $\mathtt{Base}(\mathtt{Arg}(\mathcal{S})) = \mathcal{S}$ holds. It is clear that $\mathtt{Arg}(\mathcal{S}) \subseteq \mathtt{Arg}(\Sigma)$.
  The following counter-example shows that the function $\mathtt{Base}$ is not injective for any Σ: Let $\Sigma = \{x, x \rightarrow y\}$, $\mathcal{E} = \{(\{x\}, x), (\{x \rightarrow y\}, x \rightarrow y)\}$ and $\mathcal{E}' = \{(\{x\}, x), (\{x, x \rightarrow y\}, y)\}$. Since $\mathtt{Base}(\mathcal{E}) = \mathtt{Base}(\mathcal{E}') = \Sigma$, with $\mathcal{E} \neq \mathcal{E}'$ then $\mathtt{Base}$ is not injective.

- If $a \in \mathcal{E}$ where $\mathcal{E} \subseteq \mathtt{Arg}(\Sigma)$, then $\mathtt{Supp}(a) \subseteq \mathtt{Base}(\mathcal{E})$. Consequently, $a \in \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$.

- Let us prove that $\mathtt{Arg}$ is injective. Let $\mathcal{S}, \mathcal{S}' \subseteq \Sigma$ with $\mathcal{S} \neq \mathcal{S}'$. Then, it must be that $\mathcal{S} \setminus \mathcal{S}' \neq \emptyset$ or $\mathcal{S}' \setminus \mathcal{S} \neq \emptyset$ (or both). Without loss of generality, let $\mathcal{S} \setminus \mathcal{S}' \neq \emptyset$ and let $x \in \mathcal{S} \setminus \mathcal{S}'$. If $\{x\}$ is consistent, then, $(\{x\}, x) \in \mathtt{Arg}(\mathcal{S}) \setminus \mathtt{Arg}(\mathcal{S}')$. Thus, $\mathtt{Arg}(\mathcal{S}) \neq \mathtt{Arg}(\mathcal{S}')$.
  We will now present an example that shows that this function is not surjective. Let $\Sigma = \{x, x \rightarrow y\}$ and $\mathcal{E} = \{(\{x\}, x), (\{x \rightarrow y\}, x \rightarrow y)\}$. It is clear that there exists no $\mathcal{S} \subseteq \Sigma$ s.t. $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$, since such a set $\mathcal{S}$ would contain Σ and, consequently, $\mathtt{Arg}(\mathcal{S})$ would contain $(\{x, x \rightarrow y\}, y)$, an argument not belonging to $\mathcal{E}$.

□

Proposition 4.5.2. A set $\mathcal{S} \subseteq \Sigma$ is consistent *iff* $\mathtt{Arg}(\mathcal{S})$ is conflict-free.

Proof. Let $\mathcal{S} \subseteq \Sigma$.

- Assume that $\mathcal{S}$ is consistent and $\mathtt{Arg}(\mathcal{S})$ is not conflict-free. This means that there exist $a, a' \in \mathtt{Arg}(\mathcal{S})$ s.t. a undercuts a'. From the definition of undercut, it follows that $\mathtt{Supp}(a) \cup \mathtt{Supp}(a')$ is inconsistent. Besides, from the definition of argument, $\mathtt{Supp}(a) \subseteq \mathcal{S}$ and $\mathtt{Supp}(a') \subseteq \mathcal{S}$. Thus, $\mathtt{Supp}(a) \cup \mathtt{Supp}(a') \subseteq \mathcal{S}$. Then, $\mathcal{S}$ is inconsistent. Contradiction.

- Assume now that $\mathcal{S}$ is inconsistent. This means that there exists a finite set $\mathcal{S}' = \{h_1, \dots, h_k\}$ s.t.

  - $\mathcal{S}' \subseteq \mathcal{S}$
  - $\mathcal{S}' \vdash \bot$
  - $\mathcal{S}'$ is minimal (w.r.t. set inclusion) s.t. previous two items hold.

  Since $\mathcal{S}'$ is a minimal inconsistent set, then $\{h_1, \dots, h_{k-1}\}$ and $\{h_k\}$ are consistent. Thus, $(\{h_1, \dots, h_{k-1}\}, \neg h_k), (\{h_k\}, h_k) \in \mathtt{Arg}(\mathcal{S})$. Furthermore, those two arguments are conflicting (the former undercuts the latter). This means that $\mathtt{Arg}(\mathcal{S})$ is not conflict-free.

□

Proposition 4.5.3. Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \text{Undercut}, \geq)$ be a basic PAF and let $\geq$ be a total preorder (i.e. any pair of arguments is comparable). Then: for all stable extensions $\mathcal{E}$ and $\mathcal{E}'$ of $\mathcal{T}$, if $\mathcal{E} \neq \mathcal{E}'$, then $\neg(\mathcal{E} \succeq_d \mathcal{E}')$.

Proof. Let $\mathcal{A} = \mathtt{Arg}(\Sigma)$, $\mathcal{R} = \text{Undercut}$, $\mathcal{E}, \mathcal{E}'$ be two stable extensions of $(\mathcal{A}, \mathcal{R}, \geq)$, and $\mathcal{E} \succeq_d \mathcal{E}'$ with $\mathcal{E} \neq \mathcal{E}'$. It is clear that $\neg(\mathcal{E} \subseteq \mathcal{E}')$ and $\neg(\mathcal{E}' \subseteq \mathcal{E})$. Let $a'' \in \mathcal{E} \setminus \mathcal{E}'$ be an argument s.t. $\forall x \in \mathcal{E} \setminus \mathcal{E}'$, $a'' \geq x$. Since $\mathcal{E}'$ is a stable extension, then $\mathcal{E}'$ attacks (w.r.t. $\mathcal{R}'$) argument a''. Thus $\exists a' \in \mathcal{E}' \setminus \mathcal{E}$ s.t. $\neg(a'' > a')$. Since $\geq$ is total, then $a' \geq a''$. Thus, $\forall b \in \mathcal{E} \setminus \mathcal{E}'$, $a' \geq b$. Since $\mathcal{E} \succeq_d \mathcal{E}'$, then $\exists a \in \mathcal{E} \setminus \mathcal{E}'$, s.t. $a > a'$, contradiction. □

Theorem 4.5.1. Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a stratified knowledge base. For every preferred sub-theory $\mathcal{S}$ of Σ, it holds that:

- $\mathtt{Arg}(\mathcal{S})$ is a stable extension of ($\mathtt{Arg}(\Sigma)$, Undercut, $\geq_{wlp}$)
- $\mathcal{S} = \mathtt{Base}(\mathtt{Arg}(\mathcal{S}))$

Proof. Let $\mathcal{R} = Undercut$, and for $S \subseteq \Sigma$, for $i \in \{1, \ldots, n\}$, let us denote $S_i = S \cap \Sigma_i$. Let $\mathcal{S}$ be a preferred sub-theory of a knowledge base Σ. Thus, $\mathcal{S}$ is consistent. From Proposition 4.5.2, it follows that $\texttt{Arg}(\mathcal{S})$ is conflict-free. Assume that $\exists a \notin \texttt{Arg}(\mathcal{S})$. Since $a \notin \texttt{Arg}(\mathcal{S})$ and $\mathcal{S}$ is a maximal consistent subset of Σ (according to Proposition 4.5.4), then $\exists h \in \texttt{Supp}(a)$ s.t. $\mathcal{S} \cup \{h\} \vdash \bot$. Assume that $h \in \Sigma_j$. Thus, $\texttt{Level}(\texttt{Supp}(a)) \geq j$.

Since $\mathcal{S}$ is a preferred sub-theory of Σ, then $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j$ is a maximal (for set inclusion) consistent subset of $\Sigma_1 \cup \ldots \cup \Sigma_j$. Thus, $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j \cup \{h\} \vdash \bot$. This means that there exists an argument $(\mathcal{S}', \neg h) \in \texttt{Arg}(\mathcal{S})$ s.t. $\mathcal{S}' \subseteq \mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j$. Thus, $\texttt{Level}(\mathcal{S}') \leq j$. Consequently, $(\mathcal{S}', \neg h) \geq_{wlp} a$. Moreover, $(\mathcal{S}', \neg h)\mathcal{R}a$. Thus, $(\mathcal{S}', \neg h)\mathcal{R}'a$.
The second part of the theorem follows directly from Proposition 4.5.1.

□

Theorem 4.5.2. Let Σ be a stratified knowledge base. For every stable extension $\mathcal{E}$ of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{wlp})$, it holds that:

- $\texttt{Base}(\mathcal{E})$ is a preferred sub-theory of Σ
- $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$

Proof. Throughout the proof, for $S \subseteq \Sigma$ and $i \in \{1, \ldots, n\}$, we will use notation $\mathcal{S}_i = \mathcal{S} \cap \Sigma_i$. Also, $\texttt{PST}$ will denote the set of all preferred sub-theories of Σ.

- We will first show that if $\mathcal{S} \subseteq \Sigma$, $\mathcal{E} = \texttt{Arg}(\mathcal{S})$ and $\mathcal{E}$ is a stable extension then $\mathcal{S} \in \texttt{PST}$. We will suppose that $\mathcal{S} \notin \texttt{PST}$ and we will prove that $\mathcal{E}$ is not a stable extension. If $\mathcal{S}$ is not consistent, then Proposition 4.5.2 implies that $\mathcal{E}$ is not conflict-free. Let us study the case when $\mathcal{S}$ is consistent but it is not a preferred subtheory. Thus, there exists $i \in \{1, \ldots, n\}$ such that $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_i$ is not a maximal consistent set in $\Sigma_1, \ldots, \Sigma_i$. Let i be minimal s.t. $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_i$ is not a maximal consistent set in $\Sigma_1, \ldots, \Sigma_i$. This means that there exists $x \notin \mathcal{S}$ s.t. $x \in \Sigma_i$ and $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_i \cup \{x\}$ is consistent. Let $a' = (\{x\}, x)$. Since $\mathcal{E}$ is a stable extension, then $(\exists a \in \mathcal{E})$ s.t. $a\mathcal{R}'a'$. Since $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_i \cup \{x\}$ is consistent then no argument in $\mathcal{E}$ having level at most i cannot be in conflict with a'. Thus, we have that $\nexists a \in \mathcal{E}$ s.t. $a\mathcal{R}'a'$, which proves that $\mathcal{E}$ is not a stable extension.
- We will now prove that if $\mathcal{E} \subseteq \mathcal{A}$ is a stable extension of $(\mathcal{A}, \mathcal{R}, \geq)$ and $\mathcal{S} = \texttt{Base}(\mathcal{E})$ then $\mathcal{E} = \texttt{Arg}(\mathcal{S})$. Suppose the contrary. From Proposition 4.5.1, $\mathcal{E} \subseteq \texttt{Arg}(\texttt{Base}(\mathcal{E}))$, thus $\mathcal{E} \subsetneq \texttt{Arg}(\texttt{Base}(\mathcal{E}))$.

– Let us suppose that $\mathcal{S}$ is consistent. Since $\mathcal{S}$ is consistent, then Proposition 4.5.2 implies that $\texttt{Arg}(\mathcal{S})$ is conflict-free. Since we supposed that $\mathcal{E} \subsetneq \texttt{Arg}(\mathcal{S})$, then $\mathcal{E}$ is not maximal conflict-free, contradiction.

– Let us study the case when $\mathcal{S}$ is inconsistent. This means that there can be found a set $\mathcal{S}' = \{h'_1, \ldots, h'_k\}$ s.t.

  * $\mathcal{S}' \subseteq \mathcal{S}$
  * $\mathcal{S}' \vdash \bot$
  * $\mathcal{S}'$ is a minimal s.t. the previous two conditions are satisfied.

  Let us consider the set $\mathcal{E}'$ containing the following k arguments: $\mathcal{E}' = \{a'_1, \ldots, a'_k\}$, where $a'_i = (\mathcal{S}' \setminus h'_i, \neg h'_i)$. Since $(\forall h'_i \in \mathcal{S}')(\exists a \in \mathcal{E})$ s.t. $h'_i \in \texttt{Supp}(a)$ and since $\mathcal{E}$ is conflict-free then $(\nexists b \in \mathcal{E})$ s.t. $\texttt{Conc}(b) \in \{\neg h'_1, \ldots \neg h'_k\}$. Hence, $(\forall a'_i \in \mathcal{E}')$ we have that $a'_i \notin \mathcal{E}$. Formally, $\mathcal{E} \cap \mathcal{E}' = \emptyset$. This also means that, w.r.t. $\mathcal{R}$, no argument in $\mathcal{E}$ attacks any of arguments $a'_1, \ldots, a'_k$. Formally, $(\forall a' \in \mathcal{E}')(\nexists a \in \mathcal{E})$ s.t. $a\mathcal{R}a'$. Since $\mathcal{E}$ is a stable extension then arguments of $\mathcal{E}'$ must be attacked w.r.t. $\mathcal{R}'$. We have just seen that they are not attacked w.r.t. $\mathcal{R}$. This means that:

  $$(\forall i \in \{1, \ldots, k\})(\exists a_i \in \mathcal{E})(a'_i \mathcal{R} a_i) \wedge (a_i > a'_i).$$

  For undercuts to exist, it is necessary that:

  $$(\forall i \in \{1, \ldots, k\}) \;\; (h'_i \in \texttt{Supp}(a_i)) \;\wedge\; (a_i > a'_i).$$

  From $(\forall i \in \{1, \ldots, k\}) a_i > a'_i$ we have $(\forall i \in \{1, \ldots, k\})$
  $\texttt{Level}(\{h_i\}) \leq \texttt{Level}(\texttt{Supp}(a_i)) < \texttt{Level}(\texttt{Supp}(a'_i))$.
  This means that:

  $$(\forall i \in \{1, \ldots, k\}) \;\; \texttt{Level}(\{h'_i\}) < max_{j \neq i}\texttt{Level}(\{h'_j\}).$$

  Let $l_i = \texttt{Level}(h'_i)$, for all $i \in \{1, \ldots, k\}$ and let $l_m \in \mathcal{S}'$ be s.t. $l_m = max\{l_1, \ldots, l_k\}$. Then, from the previous facts, we have:

  $$\begin{array}{lcl} l_1 & < & l_m \\ & \ldots & \\ l_m & < & max(\{l_1, \ldots, l_k\} \setminus \{l_m\}) \\ & \ldots & \\ l_k & < & l_m \end{array}$$

The row m, i.e. $l_m < max(\{l_1, \ldots, l_k\} \setminus \{l_m\})$ is an obvious contradiction since we supposed that l_m is the maximal value in $\{l_1, \ldots, l_k\}$.

- Now, we have proved that:

  1. If $\mathcal{S} \subseteq \Sigma$, $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$ and $\mathcal{E}$ is a stable extension, then $\mathcal{S} \in \mathtt{PST}$,
  2. If $\mathcal{E}$ is a stable extension then $\mathcal{E} = \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$.

  Let $\mathcal{E}$ be a stable extension and let $\mathcal{S} = \mathtt{Base}(\mathcal{E})$. Then, from (2), $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. From (1), $\mathcal{S} \in \mathtt{PST}$.

□

Theorem 4.5.3. Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \text{Undercut}, \geq_{wlp})$ be a basic PAF built from a stratified knowledge base Σ. The stable extensions of $\mathcal{T}$ are exactly the $\mathtt{Arg}(\mathcal{S})$ where $\mathcal{S}$ ranges over the preferred sub-theories of Σ.

Proof. Let us use the notation $\mathtt{PST}$ for the set of all preferred sub-theories of Σ and $\mathtt{Ext}$ for the set of stable extensions of $\mathcal{T}$.

- Theorem 4.5.1 shows that $\mathtt{Arg}(\mathtt{PST}) \subseteq \mathtt{Ext}$.
- Proposition 4.5.1 implies that $\mathtt{Arg}$ is injective.
- Let $\mathcal{E} \in \mathtt{Ext}$ and let $\mathcal{S} = \mathtt{Base}(\mathcal{E})$. From Theorem 4.5.2, we have $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. Theorem 4.5.2 yields also the conclusion that $\mathcal{S} \in \mathtt{PST}$. Thus, $\mathtt{Arg} : \mathtt{PST} \to \mathtt{Ext}$ is surjective.

□

Proposition 4.5.5. Let $(\Sigma, \trianglerighteq)$ be a prioritized knowledge base, $\trianglerighteq$ be a total preorder and let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a corresponding stratified knowledge base, i.e. $\forall i, j \in \{1, \ldots, n\}$ $\forall x \in \Sigma_i$, $\forall y \in \Sigma_j$ we have $x \trianglerighteq y$ iff $i \leq j$. Then: $\forall \mathcal{S} \subseteq \Sigma$, $\mathcal{S}$ is a preferred sub-theory of $\Sigma_1 \cup \ldots \cup \Sigma_n$ iff $\mathcal{S}$ is a democratic sub-theory of $(\Sigma, \trianglerighteq)$.

Proof. $\Rightarrow$ Let $\mathcal{S}$ be a preferred sub-theory and let us suppose that $\mathcal{S}$ is not a democratic sub-theory. Thus, $\exists \mathcal{S}'$ s.t. $\mathcal{S}' \succeq_d \mathcal{S}$ and $\mathcal{S}' \neq \mathcal{S}$. Since $\mathcal{S}$ and $\mathcal{S}'$ are both maximal consistent sets, then $\mathcal{S} \setminus \mathcal{S}' \neq \emptyset$. Let $i \in \{1, \ldots, n\}$ be the minimal number s.t. $\mathcal{S}_i \setminus \mathcal{S}'_i \neq \emptyset$ and $x \in \mathcal{S}_i \setminus \mathcal{S}'_i$. Since $\mathcal{S}' \succeq_d \mathcal{S}$, then $\exists j < i$, $\exists y \in \mathcal{S}'_j \setminus \mathcal{S}_j$. This means that $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j \subsetneq \mathcal{S}'_1 \cup \ldots \cup \mathcal{S}'_j$. Consequently, $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j$ is not a maximal consistent set in $\Sigma_1 \cup \ldots \cup \Sigma_j$. Contradiction with the hypothesis that $\mathcal{S}$ is a preferred sub-theory.

$\Leftarrow$ Let $\mathcal{S}$ be a democratic sub-theory, and let us suppose that $\mathcal{S}$ is not a preferred sub-theory. Thus, there exists $j \in \{1, \ldots, n\}$ s.t. $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j$ is not a maximal consistent set in $\Sigma_1 \cup \ldots \cup \Sigma_j$. Let $x \in \Sigma_j \setminus \mathcal{S}$ be an argument s.t. $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j \cup \{x\}$ be a consistent set. Let $\mathcal{S}' = \mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j \cup \{x\}$. Then, $\mathcal{S}' \succeq_d \mathcal{S}$. Thus, $\mathcal{S}$ is not a democratic sub-theory.

□

Theorem 4.5.4. Let Σ be a knowledge base which is equipped with a partial preorder $\trianglerighteq$. For every democratic sub-theory $\mathcal{S}$ of Σ, it holds that $\mathtt{Arg}(S)$ is a stable extension of basic PAF $(\mathtt{Arg}(\Sigma), \text{Undercut}, \geq_{gwlp})$.

Proof. Let us denote by $\mathtt{DMS}(\Sigma)$ (or just $\mathtt{DMS}$) the set of all democratic sub-theories of Σ. We also write $x \triangleright x'$ iff $x \trianglerighteq x'$ and not $x' \trianglerighteq x$. Let $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. From Proposition 4.5.2, we see that $\mathcal{E}$ is conflict-free. We will prove that it attacks (w.r.t. $\mathcal{R}'$) any argument in its exterior. Let $a' \in \mathcal{A} \setminus \mathcal{E}$ be an arbitrary argument. Since $a' \notin \mathcal{E}$ then $\exists h' \in \mathtt{Supp}(a')$ s.t. $h' \notin \mathcal{S}$. From $\mathcal{S} \in \mathtt{DMS}(\Sigma)$ we have that $\mathcal{S}$ is a maximal consistent set. It is clear that $\mathcal{S} \cup \{h'\} \vdash \bot$. Let us identify all its minimal conflicting subsets. Formally, let $C_1, \ldots, C_k$ be all sets which satisfy the following three conditions:

1. $C_i \subseteq \mathcal{S}$
2. $C_i \cup \{h'\} \vdash \bot$
3. C_i is minimal (w.r.t. set inclusion) s.t. the two previous conditions are satisfied.

Those sets allow to construct the following k arguments: $a_1 = (C_1, \neg h'), \ldots,$ $a_k = (C_k, \neg h)$. It is obvious that each of them attacks a' w.r.t. $\mathcal{R}$. If at least one of them attacks a' w.r.t. $\mathcal{R}'$, then the proof is over. Suppose the contrary. This would mean that $\forall i \in \{1, \ldots, k\}$, $a' > a_i$. Thus, $(\forall i \in \{1, \ldots, k\})$ $(\exists h_i \in C_i)$ s.t. $h' \triangleright h_i$. In other words, for every argument a_i, there exists one formula $h_i \in \mathtt{Supp}(a_i)$, such that $h' \triangleright h_i$. Let $H = \{h_1, \ldots, h_k\}$.

Now, we can define a set $\mathcal{S}'$ as follows: $\mathcal{S}' = \mathcal{S} \cup \{h'\} \setminus H$. We will show that $\mathcal{S}'$ is consistent. Suppose the contrary. Since $\mathcal{S}$ is consistent, then any inconsistent subset of $\mathcal{S}'$ must contain h'. Let $K_1, \ldots, K_j$ be all sets which satisfy the following conditions:

1. $K_i \subseteq \mathcal{S}' \setminus \{h'\}$
2. $K_i \cup \{h'\} \vdash \bot$

3. K_i is a minimal set s.t. the previous two conditions hold.

Let $K = \{K_1, \ldots, K_j\}$ and $C = \{C_1, \ldots, C_k\}$. It is easy to see that $K \subseteq C$ (this follows immediately from the fact that $\mathcal{S}' \setminus \{h'\} \subseteq \mathcal{S}$). Furthermore, since $(\forall C_i \in C)$ $(\exists h \in H)$ s.t. $h \in C_i$ then $(\forall K_i \in K)$ $(\exists h \in H)$ s.t. $h \in K_i$. Since for all K_i, we have that $K_i \cap H = \emptyset$ then it must be that $j = 0$, i.e. $K = \emptyset$. In other words, there are no inconsistent subsets of S', which means that S' is consistent.

We can notice that $\mathcal{S}' \setminus \mathcal{S} = \{h'\}$ and $\mathcal{S} \setminus \mathcal{S}' = \{h_1, \ldots, h_k\}$. Since $\mathcal{S}'$ is consistent, we see that $\mathcal{S}' \succ \mathcal{S}$. Contradiction with $\mathcal{S} \in \mathtt{DMS}(\Sigma)$. □

Theorem 4.5.5. Let Σ be a knowledge base equipped with a partial preorder $\trianglerighteq$. For every stable extension $\mathcal{E}$ of $(\mathtt{Arg}(\Sigma),$ Undercut $, \geq_{gwlp})$, it holds that:

- $\mathtt{Base}(\mathcal{E})$ is a maximal (for set inclusion) consistent subset of Σ.
- $\mathcal{E} = \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$.

Proof. Let $\mathcal{S} = \mathtt{Base}(\mathcal{E})$.

- Let us suppose that $\mathcal{S}$ is consistent but that it is not a maximal consistent set. This means that $\exists h \in \Sigma \setminus \mathcal{S}$ s.t. $\mathcal{S} \cup \{h\}$ is consistent. From Proposition 4.5.2, $\mathcal{E}' = \mathtt{Arg}(\mathcal{S} \cup \{h\})$ is consistent. From Proposition 4.5.1, $\mathcal{E} \subseteq \mathcal{E}'$. The same result implies that $\mathcal{E} \neq \mathcal{E}'$. Thus, $\mathcal{E} \subsetneq \mathcal{E}'$, which means that $\mathcal{E}$ is not a maximal conflict-free set. Contradiction with the fact that $\mathcal{E}$ is a stable extension.

- Suppose now that $\mathcal{S}$ is inconsistent. This means that there can be found a set $S' = \{h'_1, \ldots, h'_k\}$ s.t.

  - $\mathcal{S}' \subseteq \mathcal{S}$
  - $\mathcal{S}' \vdash \bot$
  - $\mathcal{S}'$ is a minimal s.t. the previous two conditions are satisfied.

  Let us consider the set $\mathcal{E}'$ containing the following k arguments: $\mathcal{E}' = \{a'_1, \ldots, a'_k\}$, where $a'_i = (\mathcal{S}' \setminus h'_i, \neg h'_i)$. Since $(\forall h'_i \in \mathcal{S}')(\exists a \in \mathcal{E})$ s.t. $h'_i \in \mathtt{Supp}(a)$ and since $\mathcal{E}$ is conflict-free then $(\nexists b \in \mathcal{E})$ s.t. $\mathtt{Conc}(b) \in \{\neg h'_1, \ldots \neg h'_k\}$. Hence, $(\forall a'_i \in \mathcal{E}')$ we have that $a'_i \notin \mathcal{E}$. Formally, $\mathcal{E} \cap \mathcal{E}' = \emptyset$. This also means that, w.r.t. $\mathcal{R}$, no argument in $\mathcal{E}$ attacks any of arguments $a'_1, \ldots, a'_k$. Formally, $(\forall a' \in \mathcal{E}')(\nexists a \in \mathcal{E})$ s.t. $a \mathcal{R} a'$.

Since $\mathcal{E}$ is a stable extension then arguments of $\mathcal{E}'$ must be attacked w.r.t. $\mathcal{R}'$. We have just seen that they are not attacked w.r.t. $\mathcal{R}$. This means that:

$$(\forall i \in \{1,\ldots,k\})(\exists a_i \in \mathcal{E})(a'_i \mathcal{R} a_i) \wedge (a_i > a'_i).$$

For undercuts to exist, it is necessary that:

$$(\forall i \in \{1,\ldots,k\}) \ \ (h'_i \in \texttt{Supp}(a_i)) \ \wedge \ (a_i > a'_i).$$

For $i = 1$, we have: $\exists i_1 \in \{1,\ldots,k\}$ s.t. $h'_1 \rhd h'_{i_1}$. For $i = i_1$, we have that $\exists i_2 \in \{1,\ldots,k\}$ s.t. $h'_{i_1} \rhd h'_{i_2}$, thus, $h'_1 \rhd h'_{i_1} \rhd h'_{i_2}$. After k consecutive applications of the same rule, we obtain: $h'_1 \rhd h'_{i_1} \rhd \ldots \rhd h'_{i_k}$. It is clearly a contradiction since on one hand, all the formulae in the chain are different because of the strict preference between them, and, on the other hand, set $\{h'_1,\ldots,h'_k\}$ contains k formulae, thus at least two of them in a chain of $k+1$ formulae must coincide.

This ends the first part of the proof. Let us now prove that $\mathcal{E} = \texttt{Arg}(\mathcal{S})$. From Proposition 4.5.1, we have that $\mathcal{E} \subseteq \texttt{Arg}(\mathcal{S})$. Suppose that $\mathcal{E} \subsetneq \texttt{Arg}(\mathcal{S})$. In the first part of the proof, we have showed that $\mathcal{S}$ is a maximal consistent set. Thus, from Proposition 4.5.2, we have that $\texttt{Arg}(\mathcal{S})$ is conflict-free. This simply means that $\mathcal{E}$ is not a maximal conflict-free set, contradiction.

□

Theorem 4.5.6. Let $\mathcal{S}, \mathcal{S}' \subseteq \Sigma$ be maximal (for set inclusion) consistent subsets of Σ. It holds that $\mathcal{S} \succeq_d \mathcal{S}'$ iff $\texttt{Arg}(\mathcal{S}) \succeq_d \texttt{Arg}(\mathcal{S}')$.

Proof. $\Rightarrow$ Let $\mathcal{S} \succeq_d \mathcal{S}'$. Let $a' \in \mathcal{E}' \setminus \mathcal{E}$. Then $\exists h' \in \texttt{Supp}(a')$ s.t. $h' \in \mathcal{S}' \setminus \mathcal{S}$. Since $\mathcal{S} \succeq_d \mathcal{S}'$ then $\exists h \in \mathcal{S} \setminus \mathcal{S}'$ s.t. $h \rhd h'$. Let $a = (\{h\}, h)$. It is clear that $a \in \mathcal{S} \setminus \mathcal{S}'$ and $a > a'$. Thus, $\mathcal{E} \succeq_d \mathcal{E}'$.

$\Leftarrow$ Let $\mathcal{E} \succeq_d \mathcal{E}'$. Let $h' \in \mathcal{S}' \setminus \mathcal{S}$. Then $a' = (\{h'\}, h') \in \mathcal{E}' \setminus \mathcal{E}$. Thus, $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $a > a'$. Since $a \in \mathcal{E} \setminus \mathcal{E}'$, then $\exists h \in \texttt{Supp}(a)$ s.t. $h \in \mathcal{S} \setminus \mathcal{S}'$. It is clear that $h \rhd h'$. □

Theorem 4.5.7. Let Σ be equipped with a partial preorder $\trianglerighteq$.

- For every democratic sub-theory $\mathcal{S}$ of Σ, $\texttt{Arg}(\mathcal{S})$ is a stable extension of the rich PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}, \succeq_d$).
- For each stable extension $\mathcal{E}$ of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}, \succeq_d$), $\texttt{Base}(\mathcal{E})$ is a democratic sub-theory of Σ.

Proof. Let $\mathcal{R}$ = Undercut and let $\mathtt{DMS}(\Sigma)$ denote the set of all democratic sub-theories of Σ.

- From Theorem 4.5.4, we have that $\mathcal{E}$ is an extension of basic PAF $(\mathcal{A}, \mathcal{R}, \geq)$. We will prove that it is also an extension of rich PAF $(\mathcal{A}, \mathcal{R}, \geq, \succeq_d)$. Let us suppose the contrary, i.e. suppose that there exists $\mathcal{E}'$ s.t. $\mathcal{E}'$ is a stable extension and $\mathcal{E}' \succ_d \mathcal{E}$. Let $\mathcal{S}' = \mathtt{Base}(\mathcal{E}')$. From Theorem 4.5.5, $\mathcal{E}' = \mathtt{Arg}(\mathcal{S}')$. From the same theorem, we have that $\mathcal{S}'$ is maximal consistent set and from Theorem 4.5.6 that $\mathcal{S}' \succ_d \mathcal{S}$. Contradiction.

- Theorem 4.5.5 implies that $\mathcal{S}$ is a maximal conflict-free set and that $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. Suppose that $\mathcal{S} \notin \mathtt{DMS}(\Sigma)$. This means that $\exists \mathcal{S}' \subseteq \Sigma$ s.t. $\mathcal{S}' \in \mathtt{DMS}(\Sigma)$ and $\mathcal{S}' \succ_d \mathcal{S}$. From Theorem 4.5.4, $\mathcal{E}' = \mathtt{Arg}(\mathcal{S}')$ is a stable extension of a basic PAF. Theorem 4.5.6 implies that $\mathcal{E}' \succ_d \mathcal{E}$, contradiction.

□

Theorem 4.5.8. The stable extensions of $(\mathtt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}, \succeq_d)$ are exactly the $\mathtt{Arg}(\mathcal{S})$ where $\mathcal{S}$ ranges over the democratic subtheories of Σ.

Proof. Let us denote by $\mathtt{Ext}$ the set of all extensions of rich PAF $\mathcal{T} = (\mathtt{Arg}(\Sigma)$, Undercut, $\geq_{gwlp}, \succeq_d)$ and by $\mathtt{DMS}$ the set of the democratic subtheories of Σ. We will prove that $\mathtt{Arg} : \mathtt{DMS} \to \mathtt{Ext}$ is a bijection.

- Theorem 4.5.7 shows that $\mathtt{Arg}(\mathtt{DMS}) \subseteq \mathtt{Ext}$.

- Proposition 4.5.1 implies that $\mathtt{Arg}$ is injective.

- Let $\mathcal{E} \in \mathtt{Ext}$ and let $\mathcal{S} = \mathtt{Base}(\mathcal{E})$. From Theorem 4.5.5, we have $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. Theorem 4.5.7 yields the conclusion that $\mathcal{S} \in \mathtt{DMS}$. Thus, $\mathtt{Arg} : \mathtt{DMS} \to \mathtt{Ext}$ is surjective.

□

A.3 Proofs for results in Chapter 5

Proposition 5.3.1. Let $\mathcal{AF}$ be a decision framework. For all $a \in \mathcal{A}_b$, $\mathtt{Status}(a, \mathcal{AF}) = \mathtt{Status}(a, \mathcal{AF} \oplus e)$.

Proof. Let $a \in \mathcal{A}_b$. Since under grounded semantics, an argument can be either sceptically accepted or rejected, it is sufficient to show that $a \in \texttt{Sc}(\mathcal{AF}) \Rightarrow a \in \texttt{Sc}(\mathcal{AF} \oplus e)$ and $a \in \texttt{Rej}(\mathcal{AF}) \Rightarrow a \in \texttt{Rej}(\mathcal{AF} \oplus e)$.

- Assume that $a \in \texttt{Sc}(\mathcal{AF})$ and $a \in \texttt{Rej}(\mathcal{AF} \oplus e)$. This means that (1) $(\exists i \in \{1, 2, 3, \ldots\})$ $(\exists a_i \in \texttt{Sc}^i(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{A}_b)$.
  Let us now prove that:
  (2) if $(\exists i \in \{2, 3, \ldots\})$ $(\exists a_i \in \texttt{Sc}^i(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{A}_b)$ then $(\exists j \in \{1, 2, 3, \ldots\})$ $(j < i) \wedge (\exists a_j \in \texttt{Sc}^j(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{A}_b)$.

  Suppose that $(\exists i \in \{2, 3, \ldots\})$ $(\exists a_i \in \texttt{Sc}^i(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{A}_b)$. Since $a_i \in \texttt{Rej}(\mathcal{AF} \oplus e)$ then $(\exists x \in \mathcal{A} \cup \{e\})$ $(x, a_i) \in \texttt{Def}' \wedge (\nexists b \in \texttt{Sc}(\mathcal{AF} \oplus e))$ $(b, x) \in \texttt{Def}'$. Note that from $a_i \in \mathcal{A}_b$ and $(x, a_i) \in \texttt{Def}'$ we conclude that $x \in \mathcal{A}_b$. Since e is practical, then $x \neq e$. Thus, x has already existed before the agent has received the argument e. This implies $(\exists x \in \mathcal{A}_b)$ $(x, a_i) \in \texttt{Def}$. From $a_i \in \texttt{Sc}^i(\mathcal{AF})$ we conclude that some sceptically accepted argument defends argument a_i, i.e. $(\exists j \in \{1, 2, 3, \ldots\})$ $(j < i) \wedge (\exists a_j \in \texttt{Sc}^j(\mathcal{AF}) \cap \mathcal{A}_b)$. Since $(\nexists b \in \texttt{Sc}(\mathcal{AF} \oplus e))$ $(b, x) \in \texttt{Def}'$ it must be that $a_j \in \texttt{Rej}(\mathcal{AF} \oplus e)$. From (1) and (2) we get: $\exists a_1 \in \texttt{Sc}^1(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{A}_b$. Hence, a_1 is not defeated in $\mathcal{AF}$ and it is defeated in $\mathcal{AF} \oplus e$. So, $(e, a_1) \in \texttt{Def}'$. Contradiction, since e is practical and a is epistemic.

- Let $a \in \mathcal{A}_b$ be an epistemic argument such that $a \in \texttt{Rej}(\mathcal{AF})$. Let us suppose that $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$. This means that
  (1) $(\exists i \in \{1, 2, 3, \ldots\})$ $(\exists a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}) \cap \mathcal{A}_b)$.
  Let us now prove that:
  (2) if $(\exists i \in \{2, 3, \ldots\})$ $(\exists a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}) \cap \mathcal{A}_b)$ then $(\exists j \in \{1, 2, 3, \ldots\})$ $(j < i) \wedge (\exists a_j \in \texttt{Sc}^j(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}) \cap \mathcal{A}_b)$.

  Suppose that $(\exists i \in \{2, 3, \ldots\})$ $(\exists a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}) \cap \mathcal{A}_b)$. Since $a_i \in \texttt{Rej}(\mathcal{AF})$ then $(\exists x \in \mathcal{A})$ $(x, a_i) \in \texttt{Def} \wedge (\nexists b \in \texttt{Sc}(\mathcal{AF})$ $(b, x) \in \texttt{Def}$. Since $(x, a_i) \in \texttt{Def}$ and $a_i \in \mathcal{A}_b$ then $x \in \mathcal{A}_b$. But $a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e)$ implies that $(\exists j \in \{1, 2, 3, \ldots\})$ $(j < i)$ s.t. $(\exists a_j \in \texttt{Sc}^j(\mathcal{AF} \oplus e) \cap \mathcal{A}_b)$ s.t. $(a_j, x) \in \texttt{Def}'$. From $(a_j, x) \in \texttt{Def}'$ and $x \in \mathcal{A}_b$ we have that a_j is also epistemic (since practical arguments cannot attack epistemic ones). The fact that $a_j \in \mathcal{A}_b$ and e is practical implies that $a_j \neq e$. Thus, a_j existed before agent has received the new argument e. Since $(\nexists b \in \texttt{Sc}(\mathcal{AF}))$ $(b, x) \in \texttt{Def}$ then $a_j \in \texttt{Rej}(\mathcal{AF})$. Now we have proved (1) and (2). From (1) and (2) we have directly the following:

$(\exists a_1 \in \texttt{Sc}^1(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}) \cap \mathcal{A}_b)$. From $a_1 \in \texttt{Sc}^1(\mathcal{AF} \oplus e)$ we have $(\nexists y \in \mathcal{A} \cup \{e\})\ (y, a_1) \in \texttt{Def}'$ and from $a_1 \in \texttt{Rej}(\mathcal{AF})$ we have $(\exists y \in \mathcal{A})\ (y, a_1) \in \texttt{Def}$. Contradiction.

□

Proposition 5.3.2. Let $\mathcal{AF}$ be a decision framework. If $\exists a \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF})$ such that $(a, e) \in \texttt{Def}'_m$, then

- $e \in \texttt{Rej}(\mathcal{AF} \oplus e)$,
- $\texttt{GE}(\mathcal{AF}) = \texttt{GE}(\mathcal{AF} \oplus e)$
- for all $a \in \mathcal{A}_o$, $\texttt{Status}(a, \mathcal{AF}) = \texttt{Status}(a, \mathcal{AF} \oplus e)$.

Proof.

- Let $a \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF})$. From Proposition 5.3.1, $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$. Thus, $e \notin \texttt{GE}(\mathcal{AF} \oplus e)$ since $\texttt{GE}(\mathcal{AF} \oplus e)$ is conflict-free. Consequently, $e \in \texttt{Rej}(\mathcal{AF} \oplus e)$.
- $\Rightarrow$ We will now prove that $\texttt{Sc}(\mathcal{AF}) \subseteq \texttt{Sc}(\mathcal{AF} \oplus e)$. Suppose not. Then $(\exists b \in \mathcal{A})$ s.t. $b \in \texttt{Sc}(\mathcal{AF}) \wedge b \in \texttt{Rej}(\mathcal{AF} \oplus e)$. We will prove that:

  1. $(\exists i \in \{1, 2, 3, \ldots\})\ (\exists a_i \in \texttt{Sc}^i(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e))$
  2. if $(\exists i \in \{2, 3, \ldots\})\ (\exists a_i \in \texttt{Sc}^i(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e))$ then $(\exists j \in \{1, 2, 3, \ldots\})\ (j < i) \wedge (\exists a_j \in \texttt{Sc}^j(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e))$.

  Note that (1) is already proved. Let us now prove (2). Suppose that $(\exists i \in \{2, 3, \ldots\})\ (\exists a_i \in \texttt{Sc}^i(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e))$. Since $a_i \in \texttt{Rej}(\mathcal{AF} \oplus e)$ then $(\exists x \in \mathcal{A} \cup \{e\})\ (x, a_i) \in \texttt{Def}' \wedge (\nexists b \in \texttt{Sc}(\mathcal{AF} \oplus e))\ (b, x) \in \texttt{Def}'$. Suppose now that $e = x$. But $(\exists a \in \mathcal{A}_b \cup \texttt{Sc}(\mathcal{AF}))\ (a, e) \in \texttt{Def}$. Contradiction with $(\nexists b \in \texttt{Sc}(\mathcal{AF} \oplus e))\ (b, x) \in \texttt{Def}'$. Thus, $x \neq e$, and x was present in the framework $\mathcal{AF}$. Since $x \in \mathcal{A}$ and $(x, a_i) \in \texttt{Def}$, from $a_i \in \texttt{Sc}^i(\mathcal{AF})$ we conclude that some sceptically accepted argument defends argument a_i in $\mathcal{AF}$, i.e. $(\exists j \in \{1, 2, 3, \ldots\})\ (j < i) \wedge (\exists a_j \in \texttt{Sc}^j(\mathcal{AF}) \cap \mathcal{A}_b) \wedge (a_j, x) \in \texttt{Def}$. Since $(\nexists b \in \texttt{Sc}(\mathcal{AF} \oplus e))\ (b, x) \in \texttt{Def}$ it must be that $a_j \in \texttt{Rej}(\mathcal{AF} \oplus e)$. So, we proved (2). As the consequence of (1) and (2) together, it holds that: $\exists a_1 \in \texttt{Sc}^1(\mathcal{AF}) \cap \texttt{Rej}(\mathcal{AF} \oplus e)$. This means that $(\nexists b \in \mathcal{A})$ s.t. $(b, a_1) \in \texttt{Def}$ and $(\exists b \in \mathcal{A} \cup \{e\})\ (b, x) \in \texttt{Def}'$. So, $(e, a_1) \in \texttt{Def}'$. Note that e is the *only* argument that defeats a_1 in $\mathcal{AF} \oplus e$. But $(\exists a \in \texttt{Sc}(\mathcal{AF} \oplus e))\ (a, e) \in \texttt{Def}'$. Hence, a_1 is defended against all defeaters and, consequently, $a_1 \in \texttt{Sc}(\mathcal{AF} \oplus e)$. Contradiction.

$\Leftarrow$ We will now prove that $\texttt{Sc}(\mathcal{AF} \oplus e) \subseteq \texttt{Sc}(\mathcal{AF})$. Suppose not. Then $(\exists a_i \in \mathcal{A})\ a_i \in \texttt{Sc}(\mathcal{AF} \oplus e) \wedge a_i \in \texttt{Rej}(\mathcal{AF})$. We will prove that:

1. $(\exists i \in \{1, 2, 3, \ldots\})\ (\exists a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}))$
2. if $(\exists i \in \{2, 3, \ldots\})\ (\exists a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}))$ then $(\exists j \in \{1, 2, 3, \ldots\})\ (j < i) \wedge (\exists a_j \in \texttt{Sc}^j(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}))$.

Note that the (1) is already proved. Let us now prove (2). Suppose that $(\exists i \in \{2, 3, \ldots\})\ (\exists a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}))$. Since $a_i \in \texttt{Rej}(\mathcal{AF})$ then $(\exists x \in \mathcal{A})\ (x, a_i) \in \texttt{Def} \wedge (\nexists b \in \texttt{Sc}(\mathcal{AF})\ (b, x) \in \texttt{Def}$. Since $(x, a_i) \in \texttt{Def}$ and $a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e)$ then $(\exists j \in \{1, 2, 3, \ldots\})$ $(j < i) \wedge (\exists a_j \in \texttt{Sc}^j(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}))$. From $(\nexists b \in \texttt{Sc}(\mathcal{AF})$ $(b, x) \in \texttt{Def}$ we obtain that $a_j \in \texttt{Rej}(\mathcal{AF})$. Now we have proved (1) and (2). From (1) and (2) we have directly the following: $(\exists a_1 \in \texttt{Sc}^1(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}))$. From $a_1 \in \texttt{Sc}^1(\mathcal{AF} \oplus e)$ we have $(\nexists y \in \mathcal{A} \cup \{e\})\ (y, a_1) \in \texttt{Def}$ and from $a_1 \in \texttt{Rej}(\mathcal{AF})$ we have $(\exists y \in \mathcal{A})$ $(y, a_1) \in \texttt{Def}$. Contradiction.

- Since $\texttt{GE}(\mathcal{AF}) = \texttt{GE}(\mathcal{AF} \oplus e)$, then all arguments keep their original status.

□

Lemma A.3.1. Let $o \in \mathcal{O}$, $a_i \in \mathcal{H}(o)$, $a_i \in \texttt{Sc}^i(\mathcal{AF})$ and $x \in \mathcal{A}$ such that $(x, a_i) \in \texttt{Def}$.

1. If $x \in \mathcal{A}_b$ then $(\exists j \in \{1, 2, 3, \ldots\})\ (j < i) \wedge (\exists a_j \in \mathcal{A}_b \cap \texttt{Sc}^j(\mathcal{AF}))$ $(a_j, x) \in \texttt{Def}$,

2. If $x \in \mathcal{A}_o$ then $(\exists j \in \{1, 2, 3, \ldots\})\ (j < i) \wedge (\exists a_j \in (\mathcal{A}_b \cup \mathcal{H}(o)) \cap \texttt{Sc}^j(\mathcal{AF}))\ (a_j, x) \in \texttt{Def}$.

Proof. We first prove that if $a_i \in \mathcal{H}(o)$, $a_i \in \texttt{Sc}^i(\mathcal{AF})$, $x \in \mathcal{A}$ and $(x, a_i) \in \texttt{Def}$, then $(\exists j \in \{1, 2, 3, \ldots\})\ (j < i) \wedge (\exists a_j \in (\mathcal{A}_b \cup \mathcal{H}(o)) \cap \texttt{Sc}^j(\mathcal{AF}))$ $(a_j, x) \in \texttt{Def}$.

Assume that $(\nexists j \in \{1, 2, 3, \ldots\})\ (j < i) \wedge (\exists a_j \in \mathcal{H}(o) \cup \mathcal{A}_b)\ a_j \in \texttt{Sc}^j(\mathcal{AF}) \wedge (a_j, x) \in \texttt{Def}$. Since a_i is sceptically accepted and defeated, then it is defended, so $(\exists j \in \{1, 2, 3, \ldots\})\ (j < i) \wedge (\exists a_j \in \mathcal{A}_o \setminus \mathcal{H}(o))$ $a_j \in \texttt{Sc}^j(\mathcal{AF}) \wedge (a_j, x) \in \texttt{Def}$. Hence, $(\exists o' \in \mathcal{O})\ (o' \neq o)$ and $a_j \in \mathcal{H}(o')$. Since both a_i and a_j are in the grounded extension, there is no defeat between them. But, since $a_i \in \mathcal{H}(o)$ and $a_j \in \mathcal{H}(o')$, with $o' \neq o$, then

$(a_i, a_j) \in \mathcal{R}_o$ and $(a_j, a_i) \in \mathcal{R}_o$. Contradiction, since we must have either $a_i \mathtt{Def} a_j$ or $a_j \mathtt{Def} a_i$ (or both).

Suppose now that $x \in \mathcal{A}_b$. We have proved that $(\exists j \in \{1, 2, 3, \ldots\})$ $(j < i) \wedge$ $(\exists a_j \in (\mathcal{A}_b \cup \mathcal{H}(o)) \cap \mathtt{Sc}^j(\mathcal{AF}))$ $(a_j, x) \in \mathtt{Def}$. Suppose that $a_j \in \mathcal{H}(o)$. This means that a practical argument attacks an epistemic one. Contradiction. So, $a_j \in \mathcal{A}_b$. □

Proposition 5.3.3. Let $\mathcal{AF}$ be a decision framework.

- For all $a \in \mathcal{H}(\mathtt{Conc}(e))$, if $a \in \mathtt{Sc}(\mathcal{AF})$ then $a \in \mathtt{Sc}(\mathcal{AF} \oplus e)$.
- For all $a \in \mathcal{A}_o$, if $a \in \mathtt{Rej}(\mathcal{AF})$ and $a \in \mathtt{Sc}(\mathcal{AF} \oplus e)$, then $e \in \mathcal{H}(\mathtt{Conc}(a))$.

Proof. Let $o \in \mathcal{O}$ such that $e \in \mathcal{H}(o)$.

- Suppose that $\exists a \in \mathtt{Sc}(\mathcal{AF}) \cap \mathtt{Rej}(\mathcal{AF} \oplus e)$. We will prove that:

  1. $(\exists i \in \{1, 2, 3, \ldots\})$ $(\exists a_i \in (\mathtt{Sc}^i(\mathcal{AF}) \cap \mathtt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$
  2. if $(\exists i \in \{2, 3, \ldots\})$ $(\exists a_i \in (\mathtt{Sc}^i(\mathcal{AF}) \cap \mathtt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$ then $(\exists j \in \{1, 2, 3, \ldots\})$ $(j < i) \wedge$
     $(\exists a_j \in (\mathtt{Sc}^j(\mathcal{AF}) \cap \mathtt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$

  Note that we have already proved (1). Let us now prove (2). Suppose that $(\exists i \in \{2, 3, \ldots\})$ $(\exists a_i \in (\mathtt{Sc}^i(\mathcal{AF}) \cap \mathtt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$. Since argument a_i is rejected in the new framework, then $(\exists x \in \mathcal{A} \cup \{e\})$ $(x, a_i) \in \mathtt{Def}' \wedge (\nexists y \in \mathtt{Sc}(\mathcal{AF} \oplus e))$ $(y, x) \in \mathtt{Def}'$. Note that $x \neq e$, because $e \in \mathcal{H}(o)$ and arguments in favor of same option do not attack each other. Since $(a_i \in \mathtt{Sc}(\mathcal{AF}))$ and $(x, a_i) \in \mathtt{Def}$, then according to Lemma A.3.1, $(\exists j \in \{1, 2, 3, \ldots\})$ $(j < i) \wedge (\exists a_j \in \mathtt{Sc}^j(\mathcal{AF}))$ $(a_j \in \mathcal{H}(o) \cup \mathcal{A}_b) \wedge (a_j, x) \in \mathtt{Def}$. Note that $a_j \neq e$, because $a_j \in \mathtt{Sc}^j(\mathcal{AF})$ and $e \notin \mathtt{Sc}(\mathcal{AF})$. Since $(\nexists y \in \mathtt{Sc}(\mathcal{AF} \oplus e))$ $(y, x) \in \mathtt{Def}'$, then $a_j \in \mathtt{Rej}(\mathcal{AF} \oplus e)$. Argument a_j is practical, since $a_j \in \mathcal{A}_b$, according to Proposition 5.3.1, implies $a_j \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ which is in contradiction with the fact that $a_j \in \mathtt{Rej}(\mathcal{AF} \oplus e)$. So, $a_j \in \mathcal{H}(o)$. Now that we see that (1) and (2) are true, we may conclude that $(\exists a_1 \in (\mathtt{Sc}^1(\mathcal{AF}) \cap \mathtt{Rej}(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$. Since a_1 was not defeated in $\mathcal{AF}$ and it is defeated in $\mathcal{AF} \oplus e$, it holds that $(e, a_1) \in \mathtt{Def}'$. Contradiction, since $a_1 \in \mathcal{H}(o)$ and $e \in \mathcal{H}(o)$, and arguments in favor of same option do not defeat each other.

- Suppose the contrary. Then, $(\exists a \in \texttt{Rej}(\mathcal{AF}) \cap \texttt{Sc}(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$ and $e \notin \mathcal{H}(o)$. Since e is practical, it holds that $(\exists o' \in \mathcal{O})\ o' \neq o \wedge e \in \mathcal{H}(o')$. We will prove that:

  1. $(\exists i \in \{1,2,3,\ldots\})\ (\exists a_i \in (\texttt{Rej}(\mathcal{AF}) \cap \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$
  2. if $(\exists i \in \{2,3,\ldots\})\ (\exists a_i \in (\texttt{Rej}(\mathcal{AF}) \cap \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$ then $(\exists j \in \{1,2,3,\ldots\})\ (j < i) \wedge (\exists a_j \in \mathcal{H}(o) \cap \texttt{Sc}^j(\mathcal{AF} \oplus e) \cap (\texttt{Rej}(\mathcal{AF}))$

  Since $a \in \mathcal{H}(o)$, $a \in \texttt{Rej}(\mathcal{AF})$ and $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$, we see that (1) is true. So, let us prove (2). Suppose $(\exists i \in \{2,3,\ldots\})\ (\exists a_i \in (\texttt{Rej}(\mathcal{AF}) \cap \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$. Since a_i was rejected, $a_i \in \texttt{Rej}(\mathcal{AF})$, then $(\exists x \in \mathcal{A})\ (x, a_i) \in \texttt{Def} \wedge (\nexists y \in \texttt{Sc}(\mathcal{AF}))\ (y, x) \in \texttt{Def}$. Since $a_i \in \texttt{Sc}(\mathcal{AF} \oplus e)$ then, according to Lemma A.3.1, $(\exists j \in \{1,2,3,\ldots\})\ (j < i)$ s.t. $(\exists a_j \in (\texttt{Sc}^j(\mathcal{AF} \oplus e) \cap (\mathcal{H}(o) \cup \mathcal{A}_b))$ s.t. $(a_j, x) \in \texttt{Def}'$. We have $a_j \neq e$ because $a_j \in \mathcal{H}(o)$ and $e \notin \mathcal{H}(o)$. So, $a_j \in \mathcal{A}$. If $a_j \in \mathcal{A}_b$, then, according to Proposition 5.3.1, $a_j \in \texttt{Sc}(\mathcal{AF})$. Contradiction with the fact $(\nexists y \in \texttt{Sc}(\mathcal{AF}))\ (y, x) \in \texttt{Def}$. So, $a_j \in \mathcal{H}(o)$. On the other hand, since $a_i \in \texttt{Rej}(\mathcal{AF})$ then $(\nexists y \in \texttt{Sc}(\mathcal{AF}))\ (y, x) \in \texttt{Def}$. Hence, since $a_j \in \mathcal{A}$, then, it must be the case that $a_j \in \texttt{Rej}(\mathcal{AF})$. From (1) and (2) we have the following: $(\exists a_1 \in (\texttt{Rej}(\mathcal{AF}) \cap \texttt{Sc}^1(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$. So, a_1 is not defeated in $\mathcal{AF} \oplus e$ and a_1 is defeated in $\mathcal{AF}$. Contradiction.

□

Proposition 5.3.4. Let $\mathcal{AF}$ be a decision framework and $o \in \mathcal{O}_r(\mathcal{AF})$. It holds that $o \in \mathcal{O}_a(\mathcal{AF} \oplus e)$ iff $e \in \mathcal{H}(o)$ and $e \in \texttt{Sc}(\mathcal{AF} \oplus e)$.

Proof. $\Rightarrow$ Let $o \in \mathcal{O}_a(\mathcal{AF} \oplus e)$.

1. Let us prove that $e \in \mathcal{H}(o)$. Suppose not. Then $(\exists o' \in \mathcal{O})\ o \neq o' \wedge e \in \mathcal{H}(o')$. But, according to Proposition 5.3.3., all rejected arguments in favor of o will remain rejected, i.e. $\mathcal{H}(o) \subseteq \texttt{Rej}(\mathcal{AF} \oplus e)$. This means that $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$.

2. Let us now prove that $e \in \texttt{Sc}(\mathcal{AF} \oplus e)$. Suppose not. So, $e \in \texttt{Rej}(\mathcal{AF} \oplus e)$. Since $o \in \mathcal{O}_a(\mathcal{AF} \oplus e)$ then $(\exists a \in \mathcal{H}(o))\ a \in \texttt{Sc}(\mathcal{AF} \oplus e)$. Note that $a \neq e$ because $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$ and $e \in \texttt{Rej}(\mathcal{AF} \oplus e)$. This is equivalent to
   (a) $(\exists i \in \{1,2,3,\ldots\})\ (\exists a_i \in \mathcal{H}(o))\ (a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}))$.
   Let us prove that:
   (b) if $(\exists i \in \{2,3,\ldots\})\ (\exists a_i \in \mathcal{H}(o))\ (a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e) \cap \texttt{Rej}(\mathcal{AF}))$

then $(\exists j \in \{1,2,3,\ldots\})$ $(j < i) \wedge (\exists a_j \in \mathcal{H}(o))$ $(a_j \in \mathtt{Sc}^i(\mathcal{AF} \oplus e) \cap \mathtt{Rej}(\mathcal{AF}))$.

Suppose $(\exists i \in \{2,3,\ldots\})$ $(\exists a_i \in \mathcal{H}(o))$ $(a_i \in \mathtt{Sc}^i(\mathcal{AF} \oplus e) \cap \mathtt{Rej}(\mathcal{AF}))$. Since $a_i \in \mathtt{Rej}(\mathcal{AF})$ then $(\exists x \in \mathcal{A})$ $(x, a_i) \in \mathtt{Def} \wedge (\nexists y \in \mathtt{Sc}(\mathcal{AF}))$ $(y, x) \in \mathtt{Def}$. Since $a_i \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ then, according to Lemma A.3.1, $(\exists a_j \in \mathtt{Sc}^j(\mathcal{AF} \oplus e))$ s.t. $(a_j \in \mathcal{H}(o) \cup \mathcal{A}_b) \wedge (a_j, x) \in \mathtt{Def}'$. Here, we have $a_j \neq e$ because $a_j \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ and $e \notin \mathtt{Sc}(\mathcal{AF} \oplus e)$. So, a_j was already present before the agent has received the new argument e. Since $(\nexists y \in \mathtt{Sc}(\mathcal{AF}))$ $(y, x) \in \mathtt{Def}$ then $a_j \in \mathtt{Rej}(\mathcal{AF})$. Suppose that $a_j \in \mathcal{A}_b$. Then, according to Proposition 5.3.1, $a_j \in \mathtt{Rej}(\mathcal{AF} \oplus e)$, contradiction. So, $a_j \in \mathcal{H}(o)$. Now, when we have proved both (1) and (2), we conclude that $(\exists a_1 \in \mathcal{H}(o))$ $(a_1 \in \mathtt{Sc}^1(\mathcal{AF} \oplus e) \cap \mathtt{Rej}(\mathcal{AF}))$. Since a_1 is not defeated in $\mathcal{AF} \oplus e$, than it is not defeated in $\mathcal{AF}$. Contradiction with $a_1 \in \mathtt{Rej}(\mathcal{AF})$.

$\Leftarrow$ If $e \in \mathcal{H}(o)$ and $e \in \mathtt{Sc}(\mathcal{AF} \oplus e)$, then $\mathtt{Conc}(e)$ Then, the option o is acceptable according to Definition 5.2.3. ☐

Proposition 5.3.5. Let $\mathcal{AF} = (\mathcal{O}, \mathcal{A}, \mathtt{Def}, \mathcal{H})$ be a decision framework, and $\mathcal{AF} \oplus e$ its extension with argument e. It holds that $e \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ iff for all $a \in \mathcal{A}$, if $(a, e) \in \mathtt{Def}'$, then $\exists b \in \mathtt{Sc}(\mathcal{AF}) \cap (\mathcal{A}_b \cup \mathcal{H}(\mathtt{Conc}(e)))$ s.t. $(b, a) \in \mathtt{Def}$.

Proof. Let $o \in \mathcal{O}$ be an option such that $\mathtt{Conc}(e) = o$.
$\Rightarrow$ Since $e \in \mathtt{Sc}(\mathcal{AF} \oplus e)$, then $(\exists i \in \mathcal{N})$ s.t. $e \in \mathtt{Sc}^i(\mathcal{AF} \oplus e)$ and $e \notin \mathtt{Sc}^{i-1}(\mathcal{AF} \oplus e)$. Let us now suppose that this property does not hold, i.e. that:

$$(\exists x \in \mathcal{A})(x\mathtt{Def}'e \wedge (\forall a \in \mathcal{A})(a\mathtt{Def}x \Rightarrow a \notin \mathtt{Sc}(\mathcal{AF}) \cap (\mathcal{A}_b \cup \mathcal{H}(o)))) \quad \text{(A.1)}$$

Suppose $x \in \mathcal{A}_b$. Then, since $e \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ it holds that $(\exists \alpha \in \mathcal{A}_b \cap \mathtt{Sc}(\mathcal{AF} \oplus e))$ s.t. $(\alpha, x) \in \mathtt{Def}'$. From Proposition 5.3.1, $\alpha \in \mathtt{Sc}(\mathcal{AF})$, which ends the proof.

We will now study the case when $x \in \mathcal{A}_o$. Since $e \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ then from Lemma A.3.1 $(\exists y \in \mathtt{Sc}(\mathcal{AF} \oplus e))$ s.t. $(y, x) \in \mathtt{Def}' \wedge y \in \mathcal{A}_b \cup \mathcal{H}(o)$. Let us suppose that $(\forall x \in \mathcal{A})(x, e) \in \mathtt{Def}' \Rightarrow (\exists \alpha \in \mathtt{Sc}(\mathcal{AF} \oplus e) \cap \mathcal{A}_b)$ s.t. $(\alpha, x) \in \mathtt{Def}'$. Then, since for $\alpha \in \mathcal{A}_b$ it holds that $\alpha \in \mathtt{Sc}(\mathcal{AF})$ iff $\alpha \in \mathtt{Sc}(\mathcal{AF} \oplus e)$, the proof is over. Else, $(\exists x \in \mathcal{A})$ s.t. $(x, e) \in \mathtt{Def}' \wedge (\nexists \alpha \in \mathtt{Sc}(\mathcal{AF} \oplus e) \cap \mathcal{A}_b)$ s.t. $(\alpha, x) \in \mathtt{Def}'$. From the previous facts and from $(A.1)$, we have that for at least one such x, it holds that $(\forall l < i)(\forall a_l \in \mathtt{Sc}^l(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$

if $a_l \mathtt{Def}' x$ then $a_l \in \mathtt{Rej}(\mathcal{AF})$. Since at least one a_l verifies the previous condition, we have:

$$(\exists l < i)(\exists a_l \in \mathtt{Sc}^l(\mathcal{AF} \oplus e) \cap \mathcal{H}(o) \cap \mathtt{Rej}(\mathcal{AF})) \text{ s.t. } (a_l, x) \in \mathtt{Def}'. \quad \text{(A.2)}$$

It is clear that $a_l \neq e$, since $e \notin \mathtt{Sc}^i(\mathcal{AF} \oplus e)$.

We will prove that:

$$(\forall k \in \{1, \ldots, l\})(\text{ if } \mathtt{Sc}^k(\mathcal{AF} \oplus e) \cap \mathcal{H}(o) \cap \mathtt{Rej}(\mathcal{AF}) \neq \emptyset \text{ then}$$
$$(\exists j \in \{1, \ldots, k-1\}) \text{ s.t. } \mathtt{Sc}^j(\mathcal{AF} \oplus e) \cap \mathcal{H}(o) \cap \mathtt{Rej}(\mathcal{AF}) \neq \emptyset). \quad \text{(A.3)}$$

Let $(\exists a_k \in (\mathtt{Sc}^k(\mathcal{AF} \oplus e)) \cap \mathcal{H}(o))$ s.t. $a_k \in \mathtt{Rej}(\mathcal{AF})$. Note that $a_k \neq e$ since $e \notin \mathtt{Sc}^{i-1}(\mathcal{AF} \oplus e)$. Since $a_k \notin \mathtt{Sc}(\mathcal{AF})$ then $(\exists b \in \mathcal{A})(b, a_k) \in \mathtt{Def}$. It is impossible that for all such $b \in \mathcal{A}$ $(\exists \alpha \in \mathtt{Sc}(\mathcal{AF} \oplus e) \cap \mathcal{A}_b)(\alpha, b) \in \mathtt{Def}'$ since that would mean $\alpha \in \mathtt{Sc}(\mathcal{AF}) \wedge (\alpha, b) \in \mathtt{Def}$ so $a_k \in \mathtt{Sc}(\mathcal{AF})$, contradiction. Thus, from this fact and by using Lemma A.3.1, we obtain that $(\exists b \in \mathcal{A})(b, a_k) \in \mathtt{Def}$ and $\exists j < k$ $(\exists a_j \in \mathtt{Sc}^j(\mathcal{AF} \oplus e) \cap \mathcal{H}(o)))$ s.t. $(a_j, b) \in \mathtt{Def}$. Here again, $a_j \neq e$ since $e \notin \mathtt{Sc}^{i-1}(\mathcal{AF} \oplus e)$. If for every such an argument $b \in \mathcal{A}$ (s.t. $(b, a_k) \in \mathtt{Def} \wedge (\nexists \alpha \in \mathtt{Sc}(\mathcal{AF}))$ s.t. $(\alpha, b) \in \mathtt{Def}$) it holds that $(\exists a_j \in \mathtt{Sc}(\mathcal{AF}))$ s.t. $a_j \mathtt{Def} b$, then we have that $a_k \in \mathtt{Sc}(\mathcal{AF})$, contradiction. Thus, it must be that $(\exists j < k)(\exists a_j \in \mathtt{Sc}^j(\mathcal{AF} \oplus e) \cap \mathcal{H}(o))$ s.t. $a_j \in \mathtt{Rej}(\mathcal{AF})$, with $j < k$. This ends the proof for $(A.3)$. Together with $(A.2)$, this implies a contradiction, since this obviously creates an infinite, strictly decreasing sequence of natural numbers.

$\Leftarrow$ Let us suppose that e is defended from all attacks in $\mathcal{AF} \oplus e$ by arguments of $\mathtt{Sc}(\mathcal{AF}) \cap (\mathcal{H}(o) \cup \mathcal{A}_b)$. From Proposition 5.3.1 and Proposition 5.3.3 we have that

$$\mathtt{Sc}(\mathcal{AF}) \cap (\mathcal{H}(o) \cup \mathcal{A}_b) \subseteq \mathtt{Sc}(\mathcal{AF} \oplus e) \cap (\mathcal{H}(o) \cup \mathcal{A}_b).$$

This means that e is defended from all attacks in $\mathcal{AF} \oplus e$ by arguments of $\mathtt{Sc}(\mathcal{AF} \oplus e)$. Consequently, $e \in \mathtt{Sc}(\mathcal{AF} \oplus e)$. □

Lemma A.3.2. It holds that under grounded semantics $\mathtt{Dbe}(\mathcal{AF}) \subseteq \mathtt{Sc}(\mathcal{AF})$.

Proof. Let $\mathcal{AF} = (\mathcal{O}, \mathcal{A}, \mathtt{Def}, \mathcal{H})$ and $a \in \mathtt{Dbe}(\mathcal{AF})$. Let $Att(a) = \{x_i \in \mathcal{A} \mid (x_i, a) \in \mathtt{Def}\}$. Since the set $\mathcal{A}$ of arguments is finite, let us denote $Att(a) = \{x_1, \ldots, x_n\}$. From $a \in \mathtt{Dbe}(\mathcal{AF})$, we obtain $(\forall x_i \in \mathcal{A})$ $(\exists \alpha \in \mathtt{Sc}(\mathcal{AF}) \cap \mathcal{A}_b)$ such that $(\alpha, x_i) \in \mathtt{Def}$. Let $Defends(a) = \{\alpha_1, \ldots, \alpha_k\}$ be a set such that $Defends(a) \subseteq \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF})$ and $(\forall x_i \in Att(a))$ $(\exists \alpha_j \in Defends(a))$ $(\alpha_j, x_j) \in \texttt{Def}$. Since $Defends(a) \subseteq \texttt{Sc}(\mathcal{AF})$ then $(\forall \alpha_i \in Defends(a))$ $(\exists m_i \in \{1, 2, 3, \ldots\})$ s.t. $\alpha_i \in \texttt{Sc}^{m_i}(\mathcal{AF})$. Let $m = max\{m_1, \ldots, m_k\}$. It holds that $Defends(a) \subseteq \texttt{Sc}^m(\mathcal{AF})$. Then, according to the definition of grounded semantics, it holds that $a \in \texttt{Sc}^{m+1}(\mathcal{AF})$, since argument a is defended by arguments of $\texttt{Sc}^m(\mathcal{AF})$against all attacks. From $a \in \texttt{Sc}^{m+1}(\mathcal{AF})$, we have $a \in \texttt{Sc}(\mathcal{AF})$. □

Proposition 5.3.6. Let $\mathcal{AF}$ be a decision framework and $o \in \mathcal{O}_a(\mathcal{AF})$. It holds that $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$ iff

1. $e \notin \mathcal{H}(o)$, and
2. $\nexists a \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF})$ s.t. $(a, e) \in \texttt{Def}'_m$, and
3. $\forall a \in \texttt{Dbe}(\mathcal{AF}) \cap \mathcal{H}(o)$, $(e, a) \in \texttt{Def}'_o$.

Proof. $\Rightarrow$ Since $o \in \mathcal{O}_a(\mathcal{AF})$, then $(\exists a \in \mathcal{H}(o))$ $a \in \texttt{Sc}(\mathcal{AF})$. Let $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$.

1. Suppose $e \in \mathcal{H}(o)$. Then, according to Proposition 5.3.3, $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$. Consequently, $o \in \mathcal{O}_a(\mathcal{AF} \oplus e)$, contradiction.
2. Suppose that $(\exists x \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF}))$ $(x, e) \in \texttt{Def}$. According to Proposition 5.3.2, $\texttt{Sc}(\mathcal{AF} \oplus e) = \texttt{Sc}(\mathcal{AF})$ and $\texttt{Rej}(\mathcal{AF} \oplus e) = \texttt{Rej}(\mathcal{AF}) \cup \{e\}$. So, $a \in \texttt{Sc}(\mathcal{AF})$ implies $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$. Contradiction with the fact that $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$.
3. Suppose that $(\exists a \in \texttt{Dbe}(\mathcal{AF}) \cap \mathcal{H}(o))$ $(e, a) \notin \texttt{Def}'$. Since $a \in \texttt{Dbe}(\mathcal{AF})$, Lemma A.3.2 implies that $a \in \texttt{Sc}(\mathcal{AF})$. From $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$ we obtain $a \in \texttt{Rej}(\mathcal{AF} \oplus e)$. So, $(\exists x \in \mathcal{A})$ $(x, a) \in \texttt{Def}'$ s.t. $(\nexists b \in \texttt{Sc}(\mathcal{AF} \oplus e))$ $(b, x) \in \texttt{Def}'$. Note that $x \neq e$ because $(x, a) \in \texttt{Def}'$ and $(e, a) \notin \texttt{Def}'$. So, $x \in \mathcal{A}$. From $a \in \texttt{Dbe}(\mathcal{AF})$ we have $(\exists \alpha \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF})$ s.t. $(\alpha, x) \in \texttt{Def}$. From Proposition 5.3.1, we have $\alpha \in \texttt{Sc}(\mathcal{AF} \oplus e)$. Contradiction with $(\nexists b \in \texttt{Sc}(\mathcal{AF} \oplus e))$ $(b, x) \in \texttt{Def}'$.

$\Leftarrow$ Let $e \notin \mathcal{H}(o) \wedge (\nexists x \in \mathcal{A}_b \cap \texttt{Sc}(\mathcal{AF}))$ $(x, e) \in \texttt{Def}' \wedge (\forall a \in \texttt{Dbe}(\mathcal{AF}) \cap \mathcal{H}(o))$ $(e, a) \in \texttt{Def}'$. Suppose that $o \notin \mathcal{O}_r(\mathcal{AF} \oplus e)$. Thus, $o \in \mathcal{O}_a(\mathcal{AF} \oplus e)$. This means that $(\exists a \in \mathcal{H}(o))$ $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$. We will prove the following:

1. $(\exists i \in \{1, 2, 3, \ldots\})$ $(\exists a_i \in \mathcal{H}(o))$ $(a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e))$.
2. if $(\exists i \in \{2, 3, \ldots\})$ $(\exists a_i \in \mathcal{H}(o))$ s.t. $(a_i \in \texttt{Sc}^i(\mathcal{AF} \oplus e))$ then $(\exists j \in \{1, 2, 3, \ldots\})$ s.t. $(j < i) \wedge (\exists a_j \in \mathcal{H}(o))$ s.t. $(a_j \in \texttt{Sc}^j(\mathcal{AF} \oplus e))$.

Note that we have already proved (1), since $(\exists a \in \mathcal{H}(o))\ a \in \mathtt{Sc}(\mathcal{AF} \oplus e)$. Let us prove (2). Suppose that $(\exists i \in \{2, 3, \dots\})\ (\exists a_i \in \mathcal{H}(o))\ (a_i \in \mathtt{Sc}^i(\mathcal{AF} \oplus e))$. Let us explore two possibilities: $a \in \mathtt{Dbe}(\mathcal{AF})$ and $a \notin \mathtt{Dbe}(\mathcal{AF})$. Suppose that $a_i \in \mathtt{Dbe}(\mathcal{AF})$. Since $a_i \in \mathtt{Dbe}(\mathcal{AF}) \cap \mathcal{H}(o)$ then $(e, a_i) \in \mathtt{Def}'$. Since $a_i \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ and $(e, a) \in \mathtt{Def}'$ then, according to Lemma A.3.1, $(\exists j \in \{1, 2, 3, \dots\})\ j < i \wedge (\exists a_j \in \mathtt{Sc}^j(\mathcal{AF} \oplus e))\ (a_j \in \mathcal{A}_b \cup \mathcal{H}(o)) \wedge (a_j, e) \in \mathtt{Def}$. We will now show that $a_j \in \mathcal{H}(o)$. Suppose that $a_j \in \mathcal{A}_b$. According to Proposition 5.3.1, $a_j \in \mathtt{Sc}(\mathcal{AF})$. Contradiction with $(\nexists x \in \mathcal{A}_b \cap \mathtt{Sc}(\mathcal{AF}))\ (x, e) \in \mathtt{Def}$. Let us now explore the case when $a_i \notin \mathtt{Dbe}(\mathcal{AF})$. From Definition 5.3.1, we have $(\exists x \in \mathcal{A})\ (x, a_i) \in \mathtt{Def} \wedge (\nexists a_j \in \mathcal{A}_b \cap \mathtt{Sc}(\mathcal{AF} \oplus e))\ (a_j, x) \in \mathtt{Def}$. Since $a_i \in \mathtt{Sc}(\mathcal{AF} \oplus e)$ and $(x, a_i) \in \mathtt{Def}'$, Lemma A.3.1 implies that $(\exists j \in \{1, 2, 3, \dots\})$ s.t. $j < i \wedge (\exists a_j \in \mathtt{Sc}^j(\mathcal{AF} \oplus e))$ s.t. $(a_j \in \mathcal{A}_b \cup \mathcal{H}(o)) \wedge (a_j, e) \in \mathtt{Def}'$. Since $(\nexists a_j \in \mathcal{A}_b \cap \mathtt{Sc}(\mathcal{AF} \oplus e))$ s.t. $(a_j, x) \in \mathtt{Def}'$ then $a_j \in \mathcal{H}(o)$.

Now, we have proved (1) and (2). As the consequence, we have that: $(\exists a_1 \in \mathcal{H}(o))$ s.t. $(a_1 \in \mathtt{Sc}^1(\mathcal{AF} \oplus e))$. This means that a_1 is not defeated by any argument in $\mathcal{AF} \oplus e$. This implies that a_1 is not defeated by any argument in $\mathcal{AF}$, i.e. $a_1 \in \mathtt{Sc}^1(\mathcal{AF})$. Consequently, $a_1 \in \mathtt{Dbe}(\mathcal{AF})$. So, $(e, a_1) \in \mathtt{Def}'$. Contradiction with the fact that a_1 is not defeated in $\mathcal{AF} \oplus e$. □

Lemma A.3.3. Let $\mathcal{AF} = (\mathcal{A}, \mathcal{R}, \geq)$ be an argumentation framework, with $\mathcal{A} = \mathcal{A}_b \cup \mathcal{A}_o$, $\mathcal{R} = \mathcal{R}_b \cup \mathcal{R}_m \cup \mathcal{R}_o$ and $\geq = \geq_b \cup \geq_m \cup \geq_o$, and let $\mathcal{AF}_b = (\mathcal{A}_b, \mathcal{R}_b, \geq_b)$ be its epistemic part.

1. If $\mathcal{E}$ is a preferred extension of $\mathcal{AF}$, then $\mathcal{E} \cap \mathcal{A}_b$ is a preferred extension of $\mathcal{AF}_b$.

2. If $\mathcal{E}$ is a preferred extension of $\mathcal{AF}_b$, then $\exists \mathcal{E}' \subseteq \mathcal{A}_o$ s.t. $\mathcal{E} \cup \mathcal{E}'$ is a preferred extension of $\mathcal{AF}$.

Proof.

1. Let $\mathcal{E}$ be a preferred extension of $\mathcal{AF}$ and let $\mathcal{E}' = \mathcal{E} \cap \mathcal{A}_b$. It is trivial that $\mathcal{E}'$ is conflict-free. If $(\exists x \in \mathcal{A}_b)(\exists y \in \mathcal{E}')$ s.t. $(x, y) \in \mathtt{Def}$ and $(\nexists z \in \mathcal{E}')$ s.t. $(z, y) \in \mathtt{Def}$ then $\mathcal{E}$ is not admissible in $\mathcal{AF}$ because of the same attack in $\mathcal{AF}$. So, $\mathcal{E}'$ must be admissible in $\mathcal{AF}_b$. If $\mathcal{E}'$ is not preferred in $\mathcal{AF}_b$ then there exists $\mathcal{E}'' \subseteq \mathcal{A}_b$ s.t. $\mathcal{E}' \subsetneq \mathcal{E}''$ and $\mathcal{E}''$ is preferred in $\mathcal{AF}_b$. In this case, $\mathcal{E} \cup \mathcal{E}''$ admissible in $\mathcal{AF}$, contradiction.

2. Let $\mathcal{E}$ be a preferred extension of $\mathcal{AF}_b$. It is conflict-free and admissible in $\mathcal{AF}$. If it is not preferred, then exists $\mathcal{E}'' \subseteq \mathcal{A}$ such that $\mathcal{E} \subseteq \mathcal{E}''$

and $\mathcal{E}''$ is preferred extension of $\mathcal{AF}$. If $\mathcal{E}'' \cap \mathcal{A}_b = \mathcal{E}$, the proof is over. Else, from the first part of this property, we have that $(\mathcal{E}'' \cap \mathcal{A}_b)$ is a preferred extension of $\mathcal{AF}_b$. Contradiction with the fact that $\mathcal{E}$ is a preferred extension, since there exists a proper superset of $\mathcal{E}$ which is admissible, contradiction.

□

Proposition 5.3.7. Let $\mathcal{AF}$ be a decision framework. For all $a \in \mathcal{A}_b$, $\texttt{Status}(a, \mathcal{AF}) = \texttt{Status}(a, \mathcal{AF} \oplus e)$.

Proof. Let $a \in \mathcal{A}_b$.

1. Suppose that $a \in \texttt{Sc}(\mathcal{AF})$ and $a \notin \texttt{Sc}(\mathcal{AF} \oplus e)$. This means that exists an extension $\mathcal{E}$ in $\mathcal{AF} \oplus e$ s.t. $a \notin \mathcal{E}$. Let $\mathcal{E}' = \mathcal{E} \cap \mathcal{A}_b$. Note that the argumentation framework $\mathcal{AF}_b = (\mathcal{A}_b, \texttt{Def}_b)$ does not change when a new practical argument is received. From Lemma A.3.3, $\mathcal{E}'$ is a preferred extension of $\mathcal{AF}_b$. From the same lemma, there exists $\mathcal{E}'' \subseteq \mathcal{A}_o$ s.t. $\mathcal{E}' \cup \mathcal{E}''$ is a preferred extension of $\mathcal{AF}$. Thus, there exists a preferred extension $\mathcal{E}' \cup \mathcal{E}''$ such that $a \notin \mathcal{E}' \cup \mathcal{E}''$. Contradiction with the fact that $a \in \texttt{Sc}(\mathcal{AF})$.

2. Suppose that $a \in \texttt{Cr}(\mathcal{AF})$ and $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$. This means that there exists an extension $\mathcal{E}$ in $\mathcal{AF}$ such that $a \notin \mathcal{E}$. Let $\mathcal{E}' = \mathcal{E} \cap \mathcal{A}_b$. From Lemma A.3.3, $\mathcal{E}'$ is a preferred extension of $\mathcal{AF}_b$. From the same lemma, exists $\mathcal{E}'' \subseteq \mathcal{A}_o \cup \{e\}$ s.t. $\mathcal{E}' \cup \mathcal{E}''$ is a preferred extension of $\mathcal{AF} \oplus e$. Thus, there exists a preferred extension $\mathcal{E}' \cup \mathcal{E}''$ such that $a \notin \mathcal{E}' \cup \mathcal{E}''$. Contradiction with the fact that $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$.
   Assume now that $a \in \texttt{Cr}(\mathcal{AF})$ and $a \in \texttt{Rej}(\mathcal{AF} \oplus e)$. This means that exists an extension $\mathcal{E}$ in $\mathcal{AF}$ such that $a \in \mathcal{E}$. Let $\mathcal{E}' = \mathcal{E} \cap \mathcal{A}_b$. from Lemma A.3.3, $\mathcal{E}'$ is a preferred extension of $\mathcal{AF}_b$. From the same lemma, exists $\mathcal{E}'' \subseteq \mathcal{A}_o \cup \{e\}$ s.t. $\mathcal{E}' \cup \mathcal{E}''$ is a preferred extension of $\mathcal{AF} \oplus e$. Thus, there exists a preferred extension $\mathcal{E}' \cup \mathcal{E}''$ of such that $a \in \mathcal{E}' \cup \mathcal{E}''$. Contradiction with the fact that $a \in \texttt{Rej}(\mathcal{AF} \oplus e)$.

3. Suppose that $a \in \texttt{Rej}(\mathcal{AF})$ and $a \notin \texttt{Rej}(\mathcal{AF})$. This means that then exists an extension $\mathcal{E}$ in $\mathcal{AF} \oplus e$ such that $a \in \mathcal{E}$. Let $\mathcal{E}' = \mathcal{E} \cap \mathcal{A}_b$. From Lemma A.3.3, $\mathcal{E}'$ is a preferred extension of $\mathcal{AF}_b$. From the same lemma, exists $\mathcal{E}'' \subseteq \mathcal{A}_o$ s.t. $\mathcal{E}' \cup \mathcal{E}''$ is a preferred extension of $\mathcal{AF}$. Thus, there exists a preferred extension $\mathcal{E}' \cup \mathcal{E}''$ of $\mathcal{AF}$ such that $a \in \mathcal{E}' \cup \mathcal{E}''$. Contradiction with the fact that $a \in \texttt{Rej}(\mathcal{AF})$.

□

Proposition 5.3.8. Let $\mathcal{AF}$ be a decision framework. If $\exists a \in \mathcal{A}_b \cap \mathtt{Sc}(\mathcal{AF})$ such that $(a, e) \in \mathtt{Def}'_m$, then

- $e \in \mathtt{Rej}(\mathcal{AF} \oplus e)$,
- $\forall \mathcal{E} \subseteq \mathcal{A}$, $\mathcal{E}$ is a preferred extension of $\mathcal{AF}$ iff $\mathcal{E}$ is a preferred extension of $\mathcal{AF} \oplus e$,
- for all $a \in \mathcal{A}_o$, $\mathtt{Status}(\mathcal{AF}, a) = \mathtt{Status}(\mathcal{AF} \oplus e, a)$.

Proof.

1. By Lemma A.3.3, from $\alpha \in \mathtt{Sc}(\mathcal{AF})$, we have that $\alpha \in \mathtt{Sc}(\mathcal{AF} \oplus e)$. Since e is attacked by a sceptically accepted argument, it must be rejected since every extension contains α and every extension is conflict-free, thus no extension can contain argument e.

2. $\Rightarrow$ Let $\mathcal{E}$ be a preferred extension of $\mathcal{AF}$. It is obvious that it is conflict-free. It is admissible in $\mathcal{AF} \oplus e$ since it defends all its elements in $\mathcal{AF}$. So, it trivially defends the arguments in $\mathcal{AF} \oplus e$ from all attacks except from attacks of e. Since sceptically accepted arguments are in all extensions, $a \in \mathcal{E}$. So, a defends $\mathcal{E}$ from attacks of e in $\mathcal{AF} \oplus e$. Thus, $\mathcal{E}$ is admissible in $\mathcal{AF} \oplus e$. Suppose now that $\mathcal{E}$ is not a preferred extension of $\mathcal{AF} \oplus e$. Then, there exists $\mathcal{E}' \subseteq \mathcal{A} \cup \{e\}$ such that $\mathcal{E}'$ is preferred extension in $\mathcal{AF} \oplus e$ and $\mathcal{E} \subsetneq \mathcal{E}'$. Since e is rejected then $e \notin \mathcal{E}'$. But it is now easy to see that $\mathcal{E}'$ is then an admissible set in $\mathcal{AF}$, thus $\mathcal{E}$ is not a preferred extension of $\mathcal{AF}$.

   $\Leftarrow$ Let $\mathcal{E}$ be a preferred extension in $\mathcal{AF} \oplus e$. Since e is rejected then $e \notin \mathcal{E}$. It is clear that $\mathcal{E}$ is conflict-free. Since $\mathcal{E}$ is admissible in $\mathcal{AF} \oplus e$, i.e. it defends all its elements, then it is easy to conclude that it defends all its elements in $\mathcal{AF}$. We will now see that $\mathcal{E}$ is preferred in $\mathcal{AF}$. Let us suppose the contrary. Then, there exists $\mathcal{E}' \subseteq \mathcal{A}$ such that $\mathcal{E}'$ is preferred in $\mathcal{AF}$ and $\mathcal{E} \subsetneq \mathcal{E}'$. As shown above, this means that $\mathcal{E}'$ is admissible in $\mathcal{AF} \oplus e$.

3. Since extensions do not change, statuses of arguments do not change.

□

Lemma A.3.4. Let $(\mathcal{O}, \mathcal{A}, \mathtt{Def}, \mathcal{H})$ be a decision framework, and let $\mathcal{E}$ be one of its preferred extensions. Let $a \in \mathcal{E} \cap \mathcal{A}_o$ and $x \in \mathcal{A}$ s.t. $(x, a) \in \mathtt{Def}$. Then:

$$(\exists a_i \in \mathcal{E} \cap (\mathcal{A}_b \cup \mathcal{H}(\mathtt{Conc}(a)))) \text{ s.t. } (a_i, x) \in \mathtt{Def}$$

Proof. Let $o \in \mathcal{O}$ be such that $a \in \mathcal{H}(o)$ and let $(x, a) \in \texttt{Def}$. Since $a \in \mathcal{E}$ then $(\exists a_i \in \mathcal{E})(a_i, x) \in \texttt{Def}$. If $(\exists a_i \in \mathcal{E})(a_i, x) \in \texttt{Def} \wedge a_i \in \mathcal{E} \cap (\mathcal{A}_b \cup \mathcal{H}(o))$ the proof is over. Else, we have that $(\forall a_i \in \mathcal{E})(a_i, x) \in \texttt{Def} \Rightarrow a_i \in \mathcal{A}_o \backslash \mathcal{H}(o)$. Thus, $a \in \mathcal{E}$ and $a_i \in \mathcal{E}$ with $\texttt{Conc}(a) \neq \texttt{Conc}(a_i)$. This means that $a\texttt{Def}a_i$ or $a_i\texttt{Def}a$, contradiction. □

Proposition 5.3.9. Let $\mathcal{AF}$ be a decision framework. For all $a \in \mathcal{A}_o$ such that $\texttt{Conc}(a) = \texttt{Conc}(e)$, it holds that:

- If $a \in \texttt{Sc}(\mathcal{AF})$ then $a \in \texttt{Sc}(\mathcal{AF} \oplus e)$
- If $a \in \texttt{Cr}(\mathcal{AF})$ then $a \in \texttt{Sc}(\mathcal{AF} \oplus e) \cup \texttt{Cr}(\mathcal{AF} \oplus e)$

Proof. Let $o \in \mathcal{O}$, $a \in \mathcal{A}_o$ and let $a, e \in \mathcal{H}(o)$.

- Assume that $a \in \texttt{Sc}(\mathcal{AF})$ and $a \notin \texttt{Sc}(\mathcal{AF} \oplus e)$. This means that there exists a preferred extension of $\mathcal{AF} \oplus e$, $\mathcal{E}'$, such that $a \notin \mathcal{E}'$. It is easy to see that $\mathcal{E}' \backslash \{e\} \backslash \mathcal{H}(o)$ is admissible in $\mathcal{AF}$: it is trivial that it is conflict-free, and from Lemma A.3.4 we see that it defends all its elements since every practical argument can be defended either by an epistemic argument or by a practical argument having the same conclusion. Note also that, according to Lemma A.3.3, $\mathcal{E}' \cap \mathcal{A}_b$ is a preferred extension of the belief part $\mathcal{AF}_b = (\mathcal{A}_b, \mathcal{R}_b, \geq_b)$ of the framework $\mathcal{AF}$. So, there exists $\mathcal{E}'' \subseteq \mathcal{A}$ s.t. $(\mathcal{E}' \setminus \{e\} \setminus \mathcal{H}(o)) \subseteq \mathcal{E}''$ and $\mathcal{E}''$ is preferred extension of $\mathcal{AF}$. Note that, since $\mathcal{E}' \cap \mathcal{A}_b$ is a preferred extension of the belief part $\mathcal{AF}_b = (\mathcal{A}_b, \mathcal{R}_b, \geq_b)$, and $\mathcal{E}'' \cap \mathcal{A}_b$ (also according to Lemma A.3.3) is a preferred extension of $\mathcal{AF}_b$ and $\mathcal{E}' \cap \mathcal{A}_b \subseteq \mathcal{E}'' \cap \mathcal{A}_b$, then $\mathcal{E}' \cap \mathcal{A}_b = \mathcal{E}'' \cap \mathcal{A}_b$. Since $a \in \texttt{Sc}(\mathcal{AF})$ it must be that $a \in \mathcal{E}''$. Since practical arguments in favor of different options attack each other, then $\mathcal{E}'' \cap \mathcal{A}_o \subseteq \mathcal{H}(o)$. Thus, $\mathcal{E}' \cap \mathcal{A}_o \subseteq \mathcal{H}(o)$. Let us study the set $\mathcal{E}' \cup (\mathcal{E}'' \cap \mathcal{H}(o))$. Clearly, we have $(\mathcal{E}' \cup (\mathcal{E}'' \cap \mathcal{H}(o))) \cap \mathcal{A}_o \subseteq \mathcal{H}(o)$ Let us show that set $\mathcal{E}' \cup (\mathcal{E}'' \cap \mathcal{H}(o))$ is admissible in $\mathcal{AF} \oplus e$:
  - it is conflict-free as union of two conflict-free sets which do not attack each other since arguments in $\mathcal{H}(o)$ do not attack other arguments in $\mathcal{H}(o)$ and arguments in $\mathcal{H}(o) \cap \mathcal{E}''$ are not attacked by arguments in $\mathcal{E}' \cap \mathcal{A}_b$ (which is equal to $\mathcal{E}'' \cap \mathcal{A}_o$, and is a subset of a preferred extension).
  - it defends its elements since $\mathcal{E}'$ is admissible in $\mathcal{AF} \oplus e$ and $\mathcal{E}'' \cap (\mathcal{H}(o) \cup \mathcal{A}_b)$ is admissible in $\mathcal{AF} \oplus e$ and union of two admissible sets which do not attack each another is an admissible set.

Contradiction, since $\mathcal{E}'$ is a preferred extension in $\mathcal{AF} \oplus e$ and there exists its strict superset $\mathcal{E}' \cap (\mathcal{E}'' \cup \mathcal{H}(o))$ which is admissible in $\mathcal{AF} \oplus e$.

- Since $a \in \texttt{Cr}(\mathcal{AF})$ then $(\exists \mathcal{E} \subseteq \mathcal{A})$ s.t. $\mathcal{E}$ is a preferred extension in $\mathcal{AF}$ and $a \in \mathcal{E}$. As a consequence of Lemma A.3.4, $\mathcal{E}' = (\mathcal{H}(o) \cup \mathcal{A}_b) \cap \mathcal{E}$ is admissible in $\mathcal{AF} \oplus e$. Thus, $(\exists \mathcal{E}'' \subseteq \mathcal{A} \cup \{e\})$ s.t. $\mathcal{E}' \subseteq \mathcal{E}''$ and $\mathcal{E}''$ is a preferred extension of $\mathcal{AF} \oplus e$. This proves that $a \notin \texttt{Rej}(\mathcal{AF} \oplus e)$.

□

Proposition 5.3.10. Let $\mathcal{AF}$ be a decision framework, and $a \in \mathcal{A}_o$. If $a \in \texttt{Rej}(\mathcal{AF})$ and $a \in \texttt{Sc}(\mathcal{AF} \oplus e) \cup \texttt{Cr}(\mathcal{AF} \oplus e)$ then $\texttt{Conc}(a) = \texttt{Conc}(e)$.

Proof. Let $a \in \texttt{Rej}(\mathcal{AF})$ and $a \notin \texttt{Rej}(\mathcal{AF} \oplus e)$. Since $a \notin \texttt{Rej}(\mathcal{AF} \oplus e)$ then there exists $\mathcal{E} \subseteq \mathcal{A} \cup \{e\}$ s.t. $a \in \mathcal{E}$ and $\mathcal{E}$ is a preferred extension of $\mathcal{AF} \oplus e$. Let $\texttt{Conc}(e) = o$, with $o \in \mathcal{O}$. Set $\mathcal{E}' \setminus \mathcal{H}(o)$ is admissible in $\mathcal{AF}$: it is conflict-free (since it is conflict-free in $\mathcal{AF} \oplus e$) and from Lemma A.3.4, it defends all its elements. Since $a \in \texttt{Rej}(\mathcal{AF})$ then a cannot be in any admissible set of $\mathcal{AF}$ since for every admissible set there exists its superset which is a preferred extension, thus a would be in at least one preferred extension which could not be the case. Consequently, $a \notin \mathcal{E} \setminus \mathcal{H}(o)$. From $a \in \mathcal{E}$ and $a \notin \mathcal{E} \setminus \mathcal{H}(o')$ it follows that $a \in \mathcal{H}(o)$. □

Proposition 5.3.11. Let $\mathcal{AF}$ be a decision framework and $o \in \mathcal{O}_r(\mathcal{AF})$. Then $o \in \mathcal{O}_a(\mathcal{AF} \oplus e) \cup \mathcal{O}_n(\mathcal{AF} \oplus e)$ iff $e \in \mathcal{H}(o) \wedge e \notin \texttt{Rej}(\mathcal{AF} \oplus e)$.

Proof. $\Rightarrow$ Let us suppose that option $o \in \mathcal{O}$ was rejected before the argument e was received, i.e. $o \in \texttt{Rej}(\mathcal{AF})$ and that its status was improved, formally $o \in \mathcal{O}_a(\mathcal{AF} \oplus e) \cup \mathcal{O}_n(\mathcal{AF} \oplus e)$. This means that all the arguments in $\mathcal{H}(o)$ were rejected, and that in the framework $\mathcal{AF} \oplus e$ there exists at least one argument in favor of o which is not rejected. We see that $e \notin \mathcal{H}(o)$ is not possible since, that would mean that some of arguments in $\mathcal{H}(o)$ improved its status, and according to Proposition 5.3.10 that $e \in \mathcal{H}(o)$. So, we proved that $e \in \mathcal{H}(o)$. Let us now prove that $e \notin \texttt{Rej}(\mathcal{AF} \oplus e)$. Suppose the contrary, i.e. let $e \in \texttt{Rej}(\mathcal{AF} \oplus e)$. This means that $\exists \mathcal{E} \subseteq \mathcal{A}$ s.t. $\mathcal{E}$ is a preferred extension in $\mathcal{AF} \oplus e$ and that $(\exists a \in \mathcal{H}(o) \cap \mathcal{E})$. In other words, there exists a non-rejected argument in favor of o. From Lemma A.3.4 we see that set $\mathcal{E} \cap (\mathcal{A}_b \cup \mathcal{H}(o))$ is admissible in $\mathcal{AF} \oplus e$. It must also be admissible in $\mathcal{AF}$. This means that $a \notin \texttt{Rej}(\mathcal{AF})$ and, consequently $o \notin \texttt{Rej}(\mathcal{AF})$. Contradiction.
$\Leftarrow$ This part of proof is trivial, since it follows directly from Definition 5.2.3. □

Proposition 5.3.12. Let $\mathcal{AF} = (\mathcal{O}, \mathcal{A}_b \cup \mathcal{A}_o, \mathtt{Def}_b \cup \mathtt{Def}_o \cup \mathtt{Def}_m, \mathcal{H})$ be a decision framework. It holds that $e \notin \mathtt{Rej}(\mathcal{AF} \oplus e)$ iff $\exists \mathcal{E} \subseteq \mathcal{A}_b$ and $\exists \mathcal{E}' \subseteq \mathcal{H}(\mathtt{Conc}(e))$ such that:

1. $\mathcal{E} \cup \mathcal{E}'$ is conflict-free, and
2. $\mathcal{E}$ is a preferred extension of $(\mathcal{A}_b, \mathtt{Def}_b)$, and
3. $\forall a \in \mathcal{E}' \cup \{e\}$, if $\exists x \in \mathcal{A}$ s.t. $(x, a) \in \mathtt{Def}$, then $\exists a' \in \mathcal{E} \cup \mathcal{E}' \cup \{e\}$ s.t. $(a', x) \in \mathtt{Def}$.

Proof. Let $o \in \mathcal{O}$ such that $o = \mathtt{Conc}(e)$.
$\Rightarrow$ Let $e \notin \mathtt{Rej}(\mathcal{AF} \oplus e)$. In other words, $\exists \mathcal{E}' \subseteq \mathcal{A} \cup \{e\}$ s.t. $\mathcal{E}'$ is a preferred extension in $\mathcal{AF} \oplus e$ and $e \in \mathcal{E}'$. Let $\mathcal{E}_b = \mathcal{E}' \cap \mathcal{A}_b$ and $\mathcal{E}_o = \mathcal{E}' \cap \mathcal{H}(o)$.

1. It is obvious that $\mathcal{E}_b \cup \mathcal{E}_o$ is conflict-free.
2. Since $\mathcal{E}'$ is a preferred extension in $\mathcal{AF} \oplus e$, then from Lemma A.3.3 we have that $\mathcal{E}_b$ is a preferred extension of framework $\mathcal{AF}_b = (\mathcal{A}_b, \mathcal{R}_b, \geq_b)$.
3. Let $a \in \mathcal{E}_o \cup \{e\}$ and let $(x, a) \in \mathtt{Def}$. Since $\mathcal{E}'$ is a preferred extension in $\mathcal{AF} \oplus e$, then $\exists a' \in \mathcal{E}'$ s.t. $(a', x) \in \mathtt{Def}$.

$\Leftarrow$ Let us suppose that the three conditions are satisfied and let us prove that $e \notin \mathtt{Rej}(\mathcal{AF} \oplus e)$. We define $\mathcal{E}'$ as follows: $\mathcal{E}' = \mathcal{E}_b \cup \mathcal{E}_o \cup \{e\}$. Recall that $\mathcal{E}_b \cup \mathcal{E}_o$ is conflict-free. Since $\mathcal{E}_o \subseteq \mathcal{H}(o)$ then $\mathcal{E}_o \cup \{e\}$ is conflict-free. Argument e being practical, it cannot attack the arguments in $\mathcal{E}_b$. Suppose now that $\mathcal{E}_b$ attacks e, i.e. $(\exists \alpha \in \mathcal{E}_b)(\alpha, e) \in \mathtt{Def}$. In that case, from the third item, $(\exists \beta \in \mathcal{E}_b)(\beta, \alpha) \in \mathtt{Def}$, contradiction with the fact $\mathcal{E}_b$ is conflict-free. Thus, $\mathcal{E}'$ is conflict-free. Set $\mathcal{E}_b$ is a preferred extension in $(\mathcal{A}, \mathcal{R}_b, \geq_b)$. From Lemma A.3.3, it is a preferred extension in epistemic part $(\mathcal{A}', \mathcal{R}'_b, \geq'_b)$ of framework $\mathcal{AF} \oplus e$. Consequently, it defends its arguments. From the third item, $\mathcal{E}'$ defends arguments of $\mathcal{E}_o \cup \{e\}$. Thus, $\mathcal{E}'$ is an admissible extension of argumentation framework $\mathcal{AF} \oplus e$. Then, $\exists \mathcal{E} \subseteq \mathcal{A} \cup \{e\}$ s.t. $\mathcal{E}' \subseteq \mathcal{E}$, $e \in \mathcal{E}$ and $\mathcal{E}$ is a preferred extension of $\mathcal{AF} \oplus e$. So, $e \notin \mathtt{Rej}(\mathcal{AF} \oplus e)$. ☐

Proposition 5.3.13. Let $\mathcal{AF}$ be a decision framework and $o \in \mathcal{O}_a(\mathcal{AF}) \cup \mathcal{O}_n(\mathcal{AF})$. Then $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$ iff

1. $e \notin \mathcal{H}(o)$, and
2. there does not exist a preferred extension $\mathcal{E}$ of $\mathcal{AF}$ s.t. $\mathcal{E} \cap \mathcal{H}(o) \neq \emptyset$ and $\exists a \in \mathcal{E} \cap \mathcal{A}_b$ s.t. $(a, e) \in \mathtt{Def}'_m$, and

3. there does not exist a preferred extension $\mathcal{E}$ of $\mathcal{AF}$ s.t. there exists an admissible set $\mathcal{E}''$ of $\mathcal{AF}$ with $\mathcal{E}'' \cap \mathcal{A}_o \subseteq \mathcal{E} \cap \mathcal{H}(o)$ and $\mathcal{E}'' \cap \mathcal{A}_b = \mathcal{E} \cap \mathcal{A}_b$ and $\forall a \in \mathcal{E}'' \cap \mathcal{H}(o), (a, e) \in >'_o$ or $\exists a' \in \mathcal{E}'' \cap \mathcal{H}(o)$ s.t. $(e, a) \notin >'_o$.

Proof. $\Rightarrow$ Let $o \in \mathcal{O}_a(\mathcal{AF}) \cup \mathcal{O}_n(\mathcal{AF})$ and let us suppose that $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$. We prove that the three conditions stated in the proposition are satisfied.

1. From Proposition 5.3.9 we have that in the case when $e \in \mathcal{H}(o)$, all sceptically accepted arguments in $\mathcal{H}(o)$ will stay sceptically accepted and that all credulously accepted arguments in $\mathcal{H}(o)$ will either stay credulously accepted or become sceptically accepted. So, it must be that $e \notin \mathcal{H}(o)$.

2. Let us suppose that there exists a preferred extension of $\mathcal{AF}$, denoted $\mathcal{E}$, s.t. $a \in \mathcal{E} \cap \mathcal{H}(o)$ and $(\exists \alpha \in \mathcal{E} \cap \mathcal{A}_b)(\alpha, e) \in \mathcal{R}$. In that case, set $\mathcal{E}$ is admissible in $\mathcal{AF} \oplus e$, since it is conflict-free (trivial) and it defends all its elements: this come from the fact that $\mathcal{E}$ is admissible in $\mathcal{AF}$ and that it attacks e. So, there exists $\mathcal{E}' \subseteq \mathcal{A} \cup \{e\}$ which is a preferred extension in $\mathcal{AF} \oplus e$, such that $\mathcal{E} \subseteq \mathcal{E}'$. Hence, $a \notin \mathtt{Rej}(\mathcal{AF} \oplus e)$. Consequently, $o \notin \mathcal{O}_r(\mathcal{AF} \oplus e)$. Contradiction.

3. Let us suppose that the third condition of proposition is not satisfied, and let $\mathcal{E}'' \subseteq \mathcal{A}$ s.t. $\mathcal{E}'' \cap \mathcal{A}_o \subseteq \mathcal{E} \cap \mathcal{H}(o)$ and $\mathcal{E}'' \cap \mathcal{A}_b = \mathcal{E} \cap \mathcal{A}_b$ and $\mathcal{E}''$ is admissible in $\mathcal{AF}$ and $((\forall a \in \mathcal{E}'' \cap \mathcal{H}(o))(a, e) \in >_o$ or $(\exists a' \in \mathcal{E}'' \cap \mathcal{H}(o)$ s.t. $\neg(e, a) \in >_o))$. Since $\mathcal{E}''$ is admissible in $\mathcal{AF}$, then it is conflict-free and it defends all its arguments from all attacks in $\mathcal{AF}$. To check whether or not it is admissible in $\mathcal{AF} \oplus e$, it is sufficient to see that it defends itself also from attacks of e: in the case when $(\forall a \in \mathcal{E}'' \cap \mathcal{H}(o))(a, e) \in >_o$ then it is not defeated by e, in the case when $(\exists a' \in \mathcal{E}'' \cap \mathcal{H}(o))\neg(e, a) \in >_o$, we have $(a, e) \in \mathtt{Def}$ so in this case also we have that $\mathcal{E}''$ is admissible. This means that $o \notin \mathtt{Rej}(\mathcal{AF} \oplus e)$, contradiction.

$\Leftarrow$ Let us suppose that $o \in \mathcal{O}_a(\mathcal{AF}) \cup \mathcal{O}_n(\mathcal{AF})$ and that three conditions of the proposition are satisfied. We prove that $o \in \mathcal{O}_r(\mathcal{AF} \oplus e)$. Suppose the contrary. This would mean that $(\exists \mathcal{E} \subseteq \mathcal{A} \cup \{e\})$ s.t. $\mathcal{E}$ is a preferred extension of $\mathcal{AF} \oplus e$ and $\mathcal{E} \cap \mathcal{H}(o) \neq \emptyset$. Let $\mathcal{E}' = \mathcal{E} \cap (\mathcal{H}(o) \cup \mathcal{A}_b)$. From Proposition 5.3.9, $\mathcal{E}'$ is admissible in $\mathcal{AF} \oplus e$. Since $e \notin \mathcal{H}(o)$ then $(\forall a \in \mathcal{E}' \cap \mathcal{H}(o))(e, a) \in \mathcal{R} \wedge (a, e) \in \mathcal{R}$. Let us suppose that $(\exists \alpha \in \mathcal{E} \cap \mathcal{A}_b)(\alpha, e) \in \mathcal{R}$. Since $\mathcal{E}'$ is admissible in $\mathcal{AF} \oplus e$ then $\mathcal{E}'$ is admissible in $\mathcal{AF}$. This is in contradiction with the second condition of the proposition. Thus, $(\nexists \alpha \in \mathcal{E} \cap \mathcal{A}_b)(\alpha, e) \in \mathcal{R}$. Since $\mathcal{E}'$ is admissible in both $\mathcal{AF}$ and $\mathcal{AF} \oplus e$, and $(\nexists \alpha \in \mathcal{E} \cap \mathcal{A}_b)(\alpha, e) \in$

$\mathcal{R}$ then either e does not defeat any of arguments in $\mathcal{E}' \cap \mathcal{H}(o)$, formally $(\forall a \in \mathcal{E}'' \cap \mathcal{H}(o))(a, e) \in >_o$ or $\mathcal{E}' \cap \mathcal{H}(o)$ defeats e, formally $(\exists a' \in \mathcal{E}'' \cap \mathcal{H}(o)$ s.t. $\neg(e, a) \in >_o)$. This is in contradiction with the third condition of the proposition since from the fact that $\mathcal{E}'$ is admissible in $\mathcal{AF}$ it holds that there exists a preferred extension of $\mathcal{AF}$, denoted $\mathcal{E}''$, such that $\mathcal{E}' \subseteq \mathcal{E}''$. Note that it must be $\mathcal{E}' \cap \mathcal{A}_b = \mathcal{E}'' \cap \mathcal{A}_b$ since in the case when $\mathcal{E}' \cap \mathcal{A}_b \subsetneq \mathcal{E}'' \cap \mathcal{A}_b$, according to Lemma A.3.3, $\mathcal{E}'' \cap \mathcal{A}_b$ would have been a preferred extension of $\mathcal{AF}_b$, which is not possible since $\mathcal{E}' \cap \mathcal{A}_b$ is preferred extension of $\mathcal{AF}_b$. Thus, the hypothesis that $o \notin \mathcal{O}_r(\mathcal{AF} \oplus e)$ was false. □

A.4 Proofs for results in Chapter 6

Proposition 6.3.1. Let $\texttt{AF}^i = (\mathcal{O}^i, \mathcal{A}^i, \mathcal{R}^i, \geq^i, \mathcal{H}^i)$ be the theory of agent i. Let $e \in \texttt{Arg}_o(\mathcal{L})$ be such that $\texttt{Conc}^i(e) \notin \mathcal{O}^i$. If $\forall e' \in \mathcal{A}^i \cap \texttt{Arg}_o(\mathcal{L})$, $e >^i e'$ and $\mathcal{R}^i_m = \emptyset$, then $\texttt{Conc}^i(e)$ will be acceptable (under preferred, grounded as well as under stable semantics if stable extensions exist) after this offer and argument have been received.

Proof. The argument e is not attacked w.r.t. $\mathcal{R}'$. Let $\mathcal{E}$ be a preferred extension of the new framework. $\mathcal{E}$ does not attack e. e does not attack $\mathcal{E}$ since that would mean that $\mathcal{E}$ is not admissible. Thus, $\mathcal{E} \cup \{e\}$ is conflict-free, contradiction. Since e is not attacked, it must be in the grounded extension. Let $\mathcal{E}$ be a stable extension s.t. $e \notin \mathcal{E}$. But $\mathcal{E}$ does not attack e, contradiction. □

Proposition 6.3.2. Let $\texttt{AF}^i = (\mathcal{O}^i, \mathcal{A}^i, \mathcal{R}^i, \geq^i, \mathcal{H}^i)$ be the theory of agent i. Let $e \in \texttt{Arg}_o(\mathcal{L})$ be such that $\texttt{Conc}^i(e) \notin \mathcal{O}^i$. If $\exists a \in \mathcal{A}^i \cap \texttt{Arg}_b(\mathcal{L})$ such that a is sceptically accepted in $\texttt{AF}^i$ and $(a, e) \in \mathcal{R}(\mathcal{L})$, then $\texttt{Conc}^i(e)$ is rejected (under preferred, grounded and stable semantics) after the new offer and argument has been received.

Proof. From Proposition 5.3.1, a is sceptically accepted in the new framework. This means that it is in every extension. Thus, e is rejected since all extensions are conflict-free. □

Proposition 6.3.4. If o is an optimal solution for Agi, then there exists a dialogue $d = \langle m_1, \ldots, m_l \rangle$, such that o is an acceptable solution for Agi at the end of the dialogue d.

Proof. Let $\texttt{AF}^1 = (\mathcal{O}, \mathcal{A}^1_0, \mathcal{R}^1_0, \geq^1_0, \mathcal{H}^1_0)$ and $\texttt{AF}^2 = (\mathcal{O}, \mathcal{A}^2_0, \mathcal{R}^2_0, \geq^2_0, \mathcal{H}^2_0)$ be the initial agents' theories and $\mathcal{A}^u = \mathcal{A}^1_0 \cup \mathcal{A}^2_0$. From Definition 6.3.4, o is acceptable in $(\mathcal{O}, \mathcal{A}^u, \mathcal{R}(\mathcal{L})|_{\mathcal{A}^u}, \geq^i (\mathcal{L})|_{\mathcal{A}^u}, \mathcal{H}^i(\mathcal{L})|_{\mathcal{A}^u \cap \mathtt{Arg}_o(\mathcal{L})})$. This means that in a dialogue in which all arguments are exchanged, o is acceptable for Agi at the end of that dialogue. □

Proposition 6.3.2. Let $Ag1$ and $Ag2$ be agents and $\mathtt{AF}^1 = (\mathcal{O}, \mathcal{A}^1, \mathcal{R}^1, \geq^1, \mathcal{H}^1)$ and $\mathtt{AF}^2 = (\mathcal{O}, \mathcal{A}^2, \mathcal{R}^2, \geq^2, \mathcal{H}^2)$ their initial theories. Let $\mathcal{A} \subseteq \mathcal{A}^1 \cup \mathcal{A}^2$ be a set s.t. $\geq^1 |_{\mathcal{A}} = \geq^2 |_{\mathcal{A}}$ and let $\mathcal{A}$ be not attacked w.r.t. $\mathcal{R}'$ by arguments of $(\mathcal{A}^1 \cup \mathcal{A}^2) \setminus \mathcal{A}$. If $\mathcal{A}^1 \cap \mathcal{A} \supseteq \mathcal{A}^2 \cap \mathcal{A}$ and $\exists o \in \mathcal{O}$, $\exists a \in \mathcal{H}^1(o) \cap \mathcal{H}^2(o) \cap \mathcal{A}$ s.t. a is sceptically accepted in $\mathtt{AF}^1$, then there exists a dialogue $d = \langle m_1, \ldots, m_l \rangle$ s.t. o is a local solution at step $t \leq l$ of d.

Proof. Let $d = (m_1, \ldots, m_l)$ be a dialogue in which $Ag2$ does not send any arguments and $Ag1$ sends exactly all arguments from $\mathcal{A}^1 \cap \mathcal{A}$ to $Ag2$. Status of the argument a did not change for $Ag1$ from the beginning until the step l, and the status of this argument will be the same for $Ag1$ and $Ag2$ after this step. Since a is sceptically accepted for $Ag1$, than it is sceptically accepted for $Ag2$. Thus, offer o is now acceptable by both agents; consequently it is a local solution at step l. □

Bibliography

Amgoud, L. (2003), "A formal framework for handling conflicting desires." In *Proceedings of the 7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'03)*, 360–365, Springer.

Amgoud, L. and Ph. Besnard (2009), "Bridging the gap between abstract argumentation systems and logic." In *International Conference on Scalable Uncertainty Management (SUM'09)*, 12–27.

Amgoud, L. and Ph. Besnard (2010), "A formal analysis of logic-based argumentation systems." In *International Conference on Scalable Uncertainty Management (SUM'10)*, 42–55.

Amgoud, L., M. Caminada, C. Cayrol, MC. Lagasquie, and H. Prakken (2004), "Towards a consensual formal model: inference part, deliverable d2.2: Draft formal semantics for inference and decision-making." Technical report, ASPIC project.

Amgoud, L. and C. Cayrol (1998), "On the acceptability of arguments in preference-based argumentation." In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI'98)*, 1–7.

Amgoud, L. and C. Cayrol (2002a), "Inferring from inconsistency in preference-based argumentation frameworks." *Journal of Automated Reasoning*, 29 (2), 125–169.

Amgoud, L. and C. Cayrol (2002b), "A reasoning model based on the production of acceptable arguments." *Annals of Mathematics and Artificial Intelligence*, 34, 197–216.

Amgoud, L., C. Cayrol, and D. LeBerre (1996), "Comparing arguments using preference orderings for argument-based reasoning." In *Proceedings of the 8th International Conference on Tools with Artificial Intelligence (ICTAI'96)*, 400–403.

Amgoud, L., Y. Dimopoulos, and P. Moraitis (2007), "A unified and general framework for argumentation-based negotiation." In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'07)*, 963–970, ACM Press.

Amgoud, L., Y. Dimopoulos, and P. Moraitis (2008), "Making decisions through preference-based argumentation." In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR'08)*, 113–123.

Amgoud, L. and S. Kaci (2007), "An argumentation framework for merging conflicting knowledge bases." *International Journal of Approximate Reasoning*, 45, 321–340.

Amgoud, L., N. Maudet, and S. Parsons (2000a), "Modelling dialogues using argumentation." In *Proceedings of the 4th International Conference on Multi-Agent Systems (ICMAS'00)*, 31–38.

Amgoud, L. and S. Parsons (2002), "An argumentation framework for merging conflicting knowledge bases." In *Proceedings of the 8th European Conference on Logics in Artificial Intelligence (JELIA'02)*, 27–37.

Amgoud, L., S. Parsons, and N. Maudet (2000b), "Arguments, dialogue, and negotiation." In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI'00)*, 338–342, IOS Press.

Amgoud, L. and H. Prade (2004), "Reaching agreement through argumentation: A possibilistic approach." In *Proceedings of the 9th International Conference on the Principles of Knowledge Representation and Reasoning (KR'04)*, 175–182.

Amgoud, L. and H. Prade (2006), "Explaining qualitative decision under uncertainty by argumentation." In *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI'06)*, 219–224, AAAI Press.

Amgoud, L. and H. Prade (2009), "Using arguments for making and explaining decisions." *Artificial Intelligence Journal*, 173, 413–436.

Amgoud, L. and S. Vesic (2009a), "On revising argumentation-based decision systems." In *Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, (ECSQARU'09)*, 71–82, Springer.

Amgoud, L. and S. Vesic (2009b), "Repairing preference-based argumentation systems." In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'09)*, 665–670.

Amgoud, L. and S. Vesic (2010a), "Generalizing stable semantics by preferences." In *Proceedings of the 3rd International Conference on Computational Models of Argument (COMMA'10)*, 39–50.

Amgoud, L. and S. Vesic (2010b), "Handling inconsistency with preference-based argumentation." In *Proceedings of the 4th International Conference on Scalable uncertainty Management (SUM'10)*, 56–69.

Amgoud, L. and S. Vesic (2010c), "On the role of preferences in argumentation frameworks." In *Proceedings of the 22nd International Conference on Tools in AI, (ICTAI'10)*, 219–222.

Amgoud, L. and S. Vesic (2011a), "A formal analysis of the outcomes of argumentation-based negotiations." In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems, (AAMAS'11)*, In Press.

Amgoud, L. and S. Vesic (2011b), "A formal analysis of the role of argumentation in negotiation dialogues." *Journal of Logic and Computation*, In Press.

Amgoud, L. and S. Vesic (2011c), "On the equivalence of logic-based argumentation systems." In *Proceedings of the 5th International Conference on Scalable uncertainty Management (SUM'11)*, In Press.

Amgoud, L. and S. Vesic (2011d), "Revising option status in argument-based decision systems." *Journal of Logic and Computation*, In Press.

Amgoud, L. and S. Vesic (2011e), "Two roles of preferences in argumentation frameworks." In *Proceedings of the 11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'11)*, In Press.

Arrow, K. J. (1951), *Social choice and individual values.* J. Wiley, New York. 2nd edition, 1963.

Atkinson, K., T. Bench-Capon, and P. McBurney (2004), "Justifying practical reasoning." In *Proceedings of the Fourth Workshop on Computational Models of Natural Argument (CMNA'04)*, 87–90.

Barbera, S., W. Bossert, and P. K. Pattanaik (2001), "Ranking sets of objects." Technical report, Université de Montréal, Département de sciences économiques.

Baroni, P. and M. Giacomin (2007), "On principle-based evaluation of extension-based argumentation semantics." *Artificial Intelligence Journal*, 171, 675–700.

Baroni, P., M. Giacomin, and G. Guida (2005), "Scc-recursiveness: a general schema for argumentation semantics." *Artificial Intelligence Journal*, 168, 162–210.

Bench-Capon, T. J. M. (2003), "Persuasion in practical argument using value-based argumentation frameworks." *Journal of Logic and Computation*, 13(3), 429–448.

Benferhat, S., D. Dubois, and H. Prade (1993), "Argumentative inference in uncertain and inconsistent knowledge bases." In *Proceedings of the 9th Conference on Uncertainty in Artificial intelligence (UAI'93)*, 411–419.

Besnard, Ph. and A. Hunter (2001), "A logic-based theory of deductive arguments." *Artificial Intelligence Journal*, 128, 203–235.

Besnard, Ph. and A. Hunter (2008), *Elements of Argumentation*. MIT Press.

Black, E. and A. Hunter (2007), "A generative inquiry dialogue system." In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multi-Agents systems (AAMAS'07)*.

Bonet, B. and H. Geffner (1996a), "Arguing for decisions: A qualitative model of decision making." In *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence (UAI'96)*, 98–105.

Bonet, B. and H. Geffner (1996b), "Arguing for decisions: A qualitative model of decision making." In *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence (UAI'96)*, 98–105.

Brena, R., C. Chesevar, and J. Aguirre (2005), "Argumentation-supported information distribution in a multiagent system for knowledge management." In *2nd International Workshop on Argumentation in Multiagent Systems (ArgMAS 2005)*, 26–33.

Brewka, G. (1989), "Preferred subtheories: An extended logical framework for default reasoning." In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'89)*, 1043–1048.

Brewka, G., I. Niemela, and M. Truszczynski (2003), "Answer set optimization." In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'03)*, 867–872.

Brewka, G., I. Niemela, and M. Truszczynski (2008), "Preferences and non-monotonic reasoning." *AI Magazine*, 69–78.

Brewka, G., M. Truszczynski, and S. Woltran (2010), "Representing preferences among sets." In *Proceedings of the Twenty-Fourth Conference on Artificial Intelligence (AAAI'10)*, 273–278.

Caminada, M. (2006a), "On the issue of reinstatement in argumentation." In *Proceedings of the 10th European Conference on Logics in Artificial Intelligence (JELIA'06)*, 111–123, Springer.

Caminada, M. (2006b), "Semi-stable semantics." In *Proceedings of the 1st International Conference on Computational Models of Argument (COMMA'06)*, 121–130, IOS Press.

Caminada, M. and L. Amgoud (2007), "On the evaluation of argumentation formalisms." *Artificial Intelligence Journal*, 171 (5-6), 286–310.

Cayrol, C. (1995), "On the relation between argumentation and non-monotonic coherence-based entailment." In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95)*, 1443–1448.

Cayrol, C., F. Dupin de Saint Cyr, and M. Lagasquie-Schiex (2008), "Revision of an argumentation system." In *Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR'08)*, 124–134.

Cayrol, C., V. Royer, and C. Saurel (1993), "Management of preferences in assumption-based reasoning." *Lecture Notes in Computer Science*, 682, 13–22.

Chesnevar, C., M. Sabaté, and A. Maguitman (2006), "An argument-based decision support system for assessing natural language usage on the basis of the web corpus." *International Journal of Intelligent Systems*, ISSN 0884-8173.

Coste-Marquis, S., C. Devred, and P. Marquis (2005), "Prudent semantics for argumentation frameworks." In *Proceedings of the 17th International Conference on Tools with Artificial Intelligence (ICTAI'05)*, 568–572, IEEE.

Dash, R.K., N.R. Jennings, and D.C. Parkes (2003), "Computational-mechanism design: A call to arms." *IEEE intelligent systems*, 18, 40–47.

Dimopoulos, Y., P. Moraitis, and L. Amgoud (2009), "Extending argumentation to make good decisions." In *Proceedings of the 1st International Conference on Algorithmic Decision Theory (ADT'09)*, 225–236.

Dimopoulos, Y. and A. Torres (1996), "Graph theoretical structures in logic programs and default theories." *Theoretical Computer Science*, 170, 209–244.

Dubois, D. and H. Fargier (2006), "Qualitative decision making with bipolar information." In *Proceedings of the 10th International Conference on Principles of Knowledge Representation and Reasoning (KR'06)*, 175–186.

Dung, P. M. (1995), "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games." *Artificial Intelligence Journal*, 77, 321–357.

Dung, P.M., P. Mancarella, and F. Toni (2007), "Computing ideal skeptical argumentation." *Artificial Intelligence Journal*, 171, 642–674.

Dunne, P. (2007), "Computational properties of argument systems satisfying graph-theoretic constraints." *Artificial Intelligence Journal*, 171 (10-15), 701–729.

Dunne, P.E. and T.J.M. Bench-Capon (2002), "Coherence in finite argument systems." *Artificial Intelligence Journal*, 141, 287–203.

Elvang-Gøransson, M., J. Fox, and P. Krause (1993), "Acceptability of arguments as 'logical uncertainty'." In *Proceedings of the 2nd European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'93)*, 85–90.

Faratin, P. (2000), *Automated service negotiation between autonomous computational agents*. Ph.D. thesis, Queen Mary and Westfield College, University of London.

Fox, J. and S. Das (2000), *Safe and Sound. Artificial Intelligence in Hazardous Applications*. AAAI Press, The MIT Press.

Fox, J. and P. McBurney (2002), "Decision making by intelligent agents: logical argument, probabilistic inference and the maintenance of beliefs and acts." In *Proceedings 9th International Workshop on Non-Monotonic Reasoning (NMR'2002)*.

Fox, J. and S. Parsons (1997), "On using arguments for reasoning about actions and values." In *Proceedings of the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, Stanford.*

Garcia, A.J. and G.R. Simari (2004), "Defeasible logic programming: an argumentative approach." *Theory and Practice of Logic Programming*, 4, 95–138.

Gordon, T. and N. Karacapilidis (1997), "The zeno argumentation framework." In *Proceedings of the sixth international conference on Artificial intelligence and law*, 10 – 18, ACM Press.

Governatori, G., M. Maher, G. Antoniou, and D. Billington (2004), "Argumentation semantics for defeasible logic." *Journal of Logic and Computation*, 14, 675–702.

Hadidi, N., Y. Dimopoulos, and P. Moraitis (2010), "Argumentative alternating offers." In *Proceedings of the international conference on Autonomous Agents and Multi-Agent Systems (AAMAS'10)*, 441–448.

Hulstijn, J. and L. van der Torre (2004), "Combining goal generation and planning in an argumentation framework." In *Proceedings of the 10th Workshop on Non-Monotonic Reasoning (NMR'04).*

Jakobovits, H. and D. Vermeir (1999), "Robust semantics for argumentation frameworks." *Journal of Logic and Computation*, 9(2), 215–261.

Jennings, N.R., P. Faratin, A.R. Lomuscio, S. Parsons, M.J. Wooldridge, and C. Sierra (2001), "Automated negotiation: prospects, methods and challenges." *Group Decision and Negotiation*, 10, 199–215.

Kaci, S. (2010), "Refined preference-based argumentation framworks." In *Proceedings of the 3rd International Conference on Computational Models of Argument (COMMA'10)*, 299–310.

Kaci, S., L. van der Torre, and E. Weydert (2006), "Acyclic argumentation: Attack = conflict + preference." In *Proceedings of the European Conference on Artificial Intelligence (ECAI'06)*, 725–726.

Kakas, A. and P. Moraitis (2006), "Adaptive agent negotiation via argumentation." In *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multi-Agents systems (AAMAS'06)*, 384–391.

Kraus, S., K. Sycara, and A. Evenchik (1998), "Reaching agreements through argumentation: a logical model and implementation." *Artificial Intelligence Journal*, 104, 1–69.

Labreuche, Ch. (2006), "Argumentation of the decision made by several aggregation operators based on weights." In *Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'06)*, 683–690.

Modgil, S. (2009), "Reasoning about preferences in argumentation frameworks." *Artificial Intelligence Journal*, 173:9–10, 901–934.

Myerson, R.B. (1997), *Game theory: analysis of conflict*. Harvard University Press.

Oikarinen, E. and S. Woltran (2010), "Characterizing strong equivalence for argumentation frameworks." In *Proceedings of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR'10)*.

Parsons, S. and N. R. Jennings (1996), "Negotiation through argumentation—a preliminary report." In *Proceedings of the 2nd International Conference on Multi Agent Systems*, 267–274.

Parsons, S., M. Wooldridge, and L. Amgoud (2003), "Properties and complexity of some formal inter-agent dialogues." *Journal of Logic and Computation*, 13 (3), 347–376.

Pollock, J. (1992), "How to reason defeasibly." *Artificial Intelligence Journal*, 57, 1–42.

Prakken, H. (2006), "Formal systems for persuasion dialogue." *Knowledge Engineering Review*, 21, 163–188.

Prakken, H. (2011), "An abstract framework for argumentation with structured arguments." *Journal of Argument and Computation*, In Press.

Prakken, H. and G. Sartor (1997), "Argument-based extended logic programming with defeasible priorities." *Journal of Applied Non-Classical Logics*, 7, 25–75.

Rahwan, I. and L. Amgoud (2006), "An argumentation-based approach for practical reasoning." In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'06)*, 347–354, ACM Press.

Rahwan, I., Ph. Pasquier, L. Sonenberg, and F. Dignum (2007), "On the benefits of exploiting underlying goals in argument-based negotiation." In *National Conference on Artificial Intelligence (AAAI'07)*, 116–121.

Rahwan, I., S.D. Ramchurn, N.R. Jennings, P. Mcburney, S. Parsons, and L. Sonenberg (2004), "Argumentation-based negotiation." *The Knowledge Engineering Review*, 18, 343–375.

Reed, C. (1998), "Dialogue frames in agent communication." In *Proceedings of the 3rd International Conference on Multi Agent Systems (ICMAS'98)*, 246–253.

Rosenschein, J. and G. Zlotkin (1994), *Rules of Encounter: Designing Conventions for Automated Negotiation Among Computers.* MIT Press, Cambridge, Massachusetts.

Rotstein, N., M. Moguillansky, M. Falappa, A. Garcia, and G. Simari (2008), "Argument theory change: revision uponwarrant." In *Proceedings of the 2nd International Conference of Computational Models of Argument (COMMA'08)*, 336–347.

Roy, B. (1985), *Méthodologie multicritère d'aide à la décision.* Economica, Paris.

Savage, L. J. (1954), *The Foundations of Statistics.* Dover, New York. Reprinted by Dover, 1972.

Simari, G.R. and R.P. Loui (1992), "A mathematical treatment of defeasible reasoning and its implementation." *Artificial Intelligence Journal*, 53, 125–157.

Sycara, K. (1990), "Persuasive argumentation in negotiation." *Theory and Decision*, 28, 203–242.

Tan, S. W. and J. Pearl (1994), "Qualitative decision theory." In *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI'94)*, 928–933.

Tarski, A. (1956), *On Some Fundamental Concepts of Metamathematics.* Logic, Semantics, Metamathematic. Edited and translated by J. H. Woodger, Oxford University Press.

Tohmé, F. (1997), "Negotiation and defeasible reasons for choice." In *Proceedings of the Stanford Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning*, 95–102.

van Eemeren, F.H., R. Grootendorst, and F. Snoeck Henkemans (1996), *Fundamentals of Argumentation Theory: A Handbook of Historical Backgrounds and Contemporary Applications*. Lawrence Erlbaum Associates, Hillsdale NJ, USA.

Varian, H.R. (1995), "Economic mechanism design for computerized agents." In *Proceedings of the First USENIX Workshop on Electronic Commerce*, 13–21.

Verheij, B. (1996), "Two approaches to dialectical argumentation: admissible sets and argumentation stages." In *Proceedings of the Eighth Dutch Conference on Artificial Intelligence (NAIC 1996)*.

von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton University Press.

Walton, D. N. and E. C. W. Krabbe (1995), *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. SUNY Series in Logic and Language, State University of New York Press, Albany, NY, USA.

Wooldridge, M., P. E. Dunne, and S. Parsons (2006), "On the complexity of linking deductive and abstract argument systems." In *Proceedings of the Twenty First National Conference on Artificial Intelligence (AAAI'06)*.

Index