Argumentation – Justification, Localization and Propagation of Admissibility

Armin Hezart

MACQUARIE
University
SYDNEY·AUSTRALIA

Department of Computing
Faculty of Science and Engineering
Macquarie University, NSW 2109, Australia

Thesis supervised by

A/Prof. Abhaya Nayak
Prof. Mehmet A. Orgun

**This thesis is submitted in fullfilment of the requirement for the degree of
Doctor of Philosophy**

July 15, 2015

Declaration

I certify that the work in this thesis entitled *Argumentation – Justification, Localization and Propagation of Admissibility* has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree to any other university or institution other than Macquarie University. I also certify that the thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of the thesis itself have been appropriately acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signed: . . . .

Date: . . . .

Publications

This thesis has resulted in the following papers; my contribution to those papers is 80% in each paper.

1. The Localization of Admissibility and the Admissibility Backings (In Preparation)

2. The Role of one Argument in Admissibility of another and the Propagation of Admissibility Backings (In Preparation)

3. A. Hezart, A. Nayak and M. A. Orgun; Towards Context Dependent Defeasible Rules. In F. Sadri and K. Satoh (editors), Computational Logic in Multi-Agent Systems, 8th International Workshop, CLIMA VIII, Porto, Portugal, September 2007, Revised Selected and Invited Papers, Lecture Notes in Articial Intelligence, Vol.5056, Springer-Verlag, 2008, pages 193–213.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Assoc. Prof. Abhaya Nayak and my co-advisor Prof. Mehmet Orgun, for their excellent guidance, care, patience, and providing me with an excellent atmosphere for doing research. They patiently quided and encouraged me to think freely, and, approach the subjects from different points of view. I could not have imagined having a better advisor and co-advisor for my Ph.D study.

List of Figures

List of Tables

Abstract

Abstract argumentation frameworks are used to study various aspects of interaction between arguments. One most fundamental such interaction is that some arguments may attack some other arguments. In a sense, only those arguments finally matter that are successfully defended against the attackers. Such arguments are called admissible. This thesis incorporates three papers in this area of research, two of which are devoted to the admissibility of arguments.

The first paper introduces the notion of the admissibility backing of an argument – a minimal set of admissible arguments that can successfully defend (respectively attack) a given argument against its attackers (defenders). This paper shows how admissible backings can help us localize the admissibility of an argument. It does so by separating those arguments that are relevant to the admissibility of a given argument from those that are not. Independent corroboration for this approach is provided by showing how major results in Dung's approach to argumentation can be obtained using admissibility backings.

The second paper explores the propagation of admissibility backings in the following sense: under what condition, and in what way, an admissibility backing of a given argument will contribute to the backing of a different argument? It is shown that under certain conditions, the propagation is transitive. It is further shown how the propagation of admissibility backings can be used to partition an argumentation framework to independent sub-frameworks. This is indicative of an interesting approach to the splitting and merging of different argumentation frameworks, a theoretical investigation of which is left for future research.

The last paper explores a novel approach to marrying the argumentation frameworks to defeasible reasoning. The desired goal behind this approach is that an argument that is deemed justied in an argumentation framework should indeed satisfy our expectations as per defeasible reasoning. The efficacy of this approach is shown by providing a mapping from it to Dung's argumentation framework.

Contents

Chapter 1

Introduction

Arguments are an integral part of practical reasoning. Arguments are used both to represent how we reason, and tell us how we should reason. In this breath, arguments are used in many facets of our lives, from the mundane activities of the daily life, e.g., what laptop to buy, to the grand issues such as the domain and limits of civil liberties. For instance, in recent years, there have been a number of debates and subsequent reforms on issue of the same sex marriage. The presented arguments are sometimes simple and clear, and, sometimes long and complicated.

An argument is only prima facie justified.[1] We accept an argument until we weigh it against its counter arguments. Argumentation is then the process by which we weigh arguments against their counter arguments. The field of Argumentation theory is the study of various facets of arguments and argumentation. For instance, Rhetorics is the study of the ways arguments are formulated and delivered within natural languages to persuade an audience. Radical Argumentativism, Communication and Rhetoric, Formal Dialectics, Pragma-Dialectics, Informal Logic and the Formal Analysis of Fallacies are among other fields of studies in argumentation theory [vE95].

In Artificial Intelligence, argumentation theory is presented as one of the approaches to non-monotonic reasoning. Many authors have, though, argued that the applications of argumentation

---

[1] In legal reasoning, Prima facie justified means justified at first glance.

theory extend beyond the conventional readings of non-monotonic reasoning [PV01].

The monotonic reasoning is a form of reasoning where the drawn conclusions stay true regardless of the added premises represnting new findings. The classical deductive logics and the mathematical reasoning are the two examples of monotonic reasoning.

The nonmonotonic reasoning, on the other hand, allows an existing conclusion to be retracted on face of new information. Inductive reasoning is nonmonotonic. Since the early years of Artificial Intelligence, it is realized that any agent that needs to make decisions under uncertainty is bound to appeal to some form of nonmonotonic reasoning.

The adoption of the formal argumentation theory to model nonmonotonic reasoning is now more than three decades old. There are also many overlaps between the concepts in argumentation theory and the works of philosophers Toulmin [Tou58], Chisholm [Chi66], Pollock [Pol67] and Rescher [Res77]. All these works are presented under the subject *defeasible reasoning*, In an unforma lsetting, nonmonotonic reasoning can be viewed as the counterpart of the defeasible reasoning in philosophy. These works continue to be a valuable source of insight and inspiration in argumentation theory. For instance, Toulmin's model of arguments is used as a guiding tool for formal representations of legal arguments [Ver09, BPWA13].

There are two well regarded surveys on the argumentation theory [CML00, PV01]. The two surveys are complementary, and, thorough. There are though a few developments since then, mostly in relation to the applications and extensions of argumentation frameworks, notably the *argumentation schemes* [DPPS09, Sim11]. In this regard, there is not much that can be added by this author to the overview of argumentation theory. We though need a platform to create a point of view for this dissertation. The first half of this dissertation, the major half, uses Dung's abstract argumentation framework as its underlying framework. Dung's abstract framework as well as the semantics of argumentation theory is presented in the background chapter of this thesis. The second half of the thesis is in relation to a more model theoretic friendly inference

rules. A background theory for this half of the thesis requires an understanding of arguments and attack relations, as well as, the argumentation schemes.

The introduction section is set to present both a general understanding of argumentation theory while explicating relevant subjects to this thesis. The background section presents Dung's framework. The motivation section explains the general principles, goals and directions of this dissertation. The conclusion chapter discusses how the envisaged directions can be followed. We start with a brief account of the elements of argumentation theory.

1.1 An overview of argumentation theory

In reasoning by argumentation, an agent draws conclusions by constructing arguments. The arguments are constructed from the agent's knowledge base and the information at hand. The agent then decides which conclusions to believe by the process of argumentation. The arguments are generally constructed to serve the agent's intentions [Pol87].

The formal argumentation theory is comprised of a number of elements. The elements are,

– the arguments,

– the counter arguments,

– the evidence,

– the schemes and the rationales.

The schemes and rationals themselves comprise,

– what constitutes an evidence,

– what constitutes an argument and a counter argument,

– how a counter argument impacts the plausibility of an argument,

– given a pool of arguments, what arguments can be coined *plausible*,

and whether or not there are forms (degrees) of plausibility, e.g., *justified* vs.

provisionally acceptable, and, if there are then what they mean within the context of

qualitative reasoning,

– and finally, how to ensure that the conclusions of the justified arguments are

indeed what an agent expects to *believe* given all the information at hand,

i.e., whether or not the reasoning is *sound* with respect to the agents model of

the real world.

These elements are brought together as an argumentation framework or an argumentation system. Under this description, argumentation theory is a form of reasoning with a distinct structure where arguments are its primary building blocks.

The objective of research in the field of argumentation theory has therefore been the investigation (the identification, elaboration, characterization, formulation and implementation) of the various aspects of each and all of the above elements, often accompanied with a corresponding framework. In the context of defeasible reasoning the most influential work in argumentation theory is the work of John Pollock. Many of the adopted concepts, e.g., the *attack relation* or the types of *attack relations*, and, the theories, e.g. the *theory of warrant* and the method of *status assignment*, can be found within his works [Pol87, Pol92, Pol94, Pol96, Pol01, Pol10].

The most unifying work, though, is the introduction of *Abstract argumentation frameworks* by Phan Minh Dung in 1995 [Dun95b]. In his seminal paper [Dun95b], Dung presents an abstract argumentation framework where arguments are represented as abstract entities, free of their internal structure. The interaction among arguments is reduced to the argument, counter argument relation. The argument, counter argument relation, called the *attack relation*, is then presented as a binary relation between the two abstract frameworks. As part of major achievements of this paper, Dung claims that,

> "most of the major approaches to nonmonotonic reasoning and logic programming are special forms of this theory of argumentation."

This thesis uses Dung's abstract argumentation framework as its underlying framework. Ac-

cordingly, a detailed account of Dung's framework is provided in the background section of this chapter.

Subsequently, the argumentation-theoretic semantics is applied to a number of nonmonotonic reasoning systems, most notably in the area of logic programming [BTK93, KMD94, Dun95a, DS95]. These works are culminated into a seminal paper by Bondarenko and et al. [BDKT97]. In their paper [BDKT97], the authors present a general argumentation-theoretic assumption based framework for *default reasoning* that subsumes many of the approaches to nonmonotonic reasoning including the *Theorists Logic* and the *Default Logic* [McC80, Rei80, Moo84, Poo88].

The argumentation-theoretic assumption based framework is inspired by the theorists approach to default reasoning [Poo88]. The work uses the argumentation theoretic concepts, such as attack relation between sets of assumption, and, argumentation theoretic semantics, e.g. the *admissibility semantics*.

We now proceed by giving a short account of the elements of argumentation theory.

1.1.1 The elements of formal argumentation theory

1.1.1.1 Arguments

In defeasible reasoning, the large majority of realization of arguments follow a general scheme. An *argument* is simply a scheme for encoding the reasonings of the sort –

> for an agent, believing in the statements $P_1, \cdots, P_n$ is a reason for believing the statements $Q_1, \cdots, Q_m$.

Hence, we may view arguments as mappings, by an agent, from the sets of statements to the sets of statements. The statements $P_1 \cdots P_n$ is usually referred to as the *premises* of an argument and the statements $Q_1, \cdots, Q_m$ its *conclusions*.

Furthermore, the underlying meanings, *interpretation*, of statements are left to the agent. This

makes the reasoning by argumentation a symbolic form of reasoning. As a consequence, it is imperative to make sure that every argument contains all the information that pertains to the *justification* of the argument, independent of how the rest of the world pans out. We call this principle, the localization of the acceptability of an argument at the knowledge base (database) level.

Various authors proposed different formulations of arguments, each based on their target objectives. These formulations tend to be heavily intertwined with the choice of the underlying language and the schemes for arguments. Yet, almost all the formulations of arguments share certain characteristics. A common practice is to construct arguments in the manner we construct proofs in the classical logics. In these argumentation systems, the agent's knowledge base consists of a number of fixed *inference rules* of the sort $P_1, P_2, \cdots, P_n \rightarrow Q$. Accordingly, the reasoning by argumentation is regarded as a form of qualitative rule based reasoning where arguments provide the tentative proofs.

The inference rules may also be classified in terms of their defeasibility. Many of the proposed frameworks employ two types of inference rules, the strict rules (the deductive rules) and the defeasible rules. The argumentation systems that are based on logic programming with negation as failure have adopted these two forms of inference rules [Dun95b].

It is the nature of defeasibility of arguments that determines how the conflict between the arguments ought to be resolved. Naturally, any resolution between the conflicting arguments should meet expectations of the agents involved. In addition, the inference rules may also be assigned certain degrees of strengths [PS97, PS99, Pol01, Ben02].

Secondly, the arguments can be chained together such that conclusion of an argument may serve as the premise of another. In this manner, arguments are usually presented as tree structures of certain characteristics. Accordingly, it is common to represent arguments by the *inference graphs*, or, as simple triangles where the base and the top represent the premises and the conclu-

sion of the argument. For instance, in Simari & Loui framework [SL92], arguments form tree structures with some additional characteristics. The additional characteristic is being subjected to the *Occam's Razor maxim*. The principle of Occam's razor is to provide a minimal proof for the conclusion that also automatically prevents circular arguments.

Another distinct feature of the formulated arguments is that the pool of constructed arguments tends to be monotonic. That is, a constructed argument cannot be removed from the pool of arguments, upon the addition of new information. Simari & Loui's framework is an exception to this rule [SL92].

The source and the classifications of inference rules are themselves the subject of much research. In the second half of this dissertation, we propose a schema for representing arguments that extends beyond the conventional approaches.

1.1.1.2 Conflict between arguments

The issue of conflict between arguments is at the heart of argumentation theory. In general, two arguments are considered to be *conflicting* if given a context, acceptance of one argument prevents us from accepting the other. It is in this sense that arguments are regarded as tentative proofs where the addition of information may make them no longer justified.

In common terminology, an argument that makes another argument inapplicable is said to be the counter argument to that argument. In the literature, the counter argument-argument relation is referred to as the *defeat* or the *attack* relation.

The distillation of disagreements between arguments to a simple counter argument-argument relation is, however, not a straightforward affair [Pol94, Lou87, Lou89, PL92]. In fact, the forms of conflict between arguments and how to resolve them is one of the most contested subjects among the researchers, especially in the case of legal reasoning [PS99, Pra01a, PS04, Ver01a]. Every now and then, a new type of conflict between arguments is identified, e.g., to

attack the strength of an argument, or, to attack the attacking force of an argument [Mod06, MGS08].

Pollock identifies two kinds of defeat relations in general, the *rebutting* and the *undercutting* defeaters [Pol87]. A rebutting defeater could be any reason for denying the conclusion of an argument. The rebutting defeaters have the distinct characteristic that they must present a conclusion that is contrary to the conclusion of the argument in question. In the literature, this feature makes the rebutting defeat relation a symmetric defeat relation.

The undercutting defeaters, on the other hand, need not present a conclusion contradictory to the conclusion of the argument. The undercutting defeaters attack the underlying defeasible inference of the argument. For instance, when an argument attacks the reliability of the provided reasoning, the argument undercuts that argument. An example that is commonly presented for the undercutting defeat is [Pol87],

> the red appearance of an object is a sufficient reason to believe that it is red. But, appearing red under a red lighting no longer warrants to believe that the object is red.

An undercutting defeater, therefore, need not show that the conclusion of the argument under question is false. Hence, the undercutting defeat relation usually tends to be an asymmetric relation.

In his writings, Pollock argues that all forms of defeat relation between arguments can be presented as the rebutting and the undercutting defeat relations. On the other hand, he also extensively discusses another form of defeat among arguments that is exemplified by the Kyberg's Lottery paradox [Kyb61] (see appendix Lottery paradox).

Pollock presents the Lottery paradox in the context of *collective defeat*. The collective defeat says while in a group of arguments no single argument can be said to be individually defeated, the group is defeated as a whole. It is though difficult to reconcile the lottery paradox form of collective defeat to either category of the rebutting or undercutting defeat relations.

In all the presented argumentation systems, attacking a subargument of an argument always impacts the justification of the argument. The converse is however a hotly debated and unresolved subject. That is, whether or not there are circumstances where a rebutting attack mandates the retraction of any of the conclusions of subarguments of an argument, i.e., whether or not, a rebutting attack can propagate from an argument to some of its subarguments [PV01]. Some of such issues resulted in a certain axiomatic account of argumentation theory [CA05, CA07].

We can therefore see that the topic of an attack relation between arguments is a complicated, demanding and sometimes a controversial subject. This issue itself is even more complicated by the choice of an underlying language, and, how the arguments are constructed.

Following the literature, arguments are regarded as basic atoms in argumentation systems. This feature allows for arguments not to be directly tied to any particular underlying language. The choice of an underlying language, however, has a real impact on how the attack relation is identified and resolved.

For instance, in the argumentation systems where the underlying language is a classical language the attack relations are at the first instance symmetric. It is because, all the conflicts are in the form of conflicts between two contradictory sentences $P, \neg P$. We are then forced to have a separate mechanism, e.g. specificity, to break the tie between the two attacking arguments. An instance of such an argumentation framework is given in [BH01] where the underlying language and the consequence relation are fully classical.

The assumption based frameworks are another instance where the choice of an underlying language is directly tied to how we construct our database and how we define the attack relation between arguments. In the assumption based frameworks, although the underlying framework is a deductive framework, the inference rules do not automatically have a contrapositive counterpart. This feature allows them to define a Dung's style semantics based on the attack relation

between sets of assumptions [KT96, BDKT97].

The other forms of attack relations are arguments attacking an attack relation [Mod09], attack relation between sets of arguments [Boc02], the attack relation that reduces the strength of an argument [Ben02], and the attack relation that reduces the strength of an attack relation [MGS08].

Except in regard to the context dependent defeasible rules, the second half of this thesis, we only deal with the Dung's abstract argumentation framework. Most of the forms of attack relations, except in case of the Lottery paradox, fall within the purview of the binary attack relation of Dung's abstract framework. This happens in one of two ways. One way is to distill the attack relations into the form of the rebutting and undercutting attack relations. The rebutting and undercutting attack relations both can be represented in terms of binary attack relations. The other way is to translate a corresponding argumentation framework into Dung's abstract framework. Hence, by adopting Dung's framework, we implicitly showcase the span of applications of our results.

Next, we discuss argumentation schemes. The subject of argumentation schemes argues for a broader view of argumentation theory than its conventional presentation. The second part of this thesis follows this line of thinking.

1.1.1.3 Argument schemes

Argumentation schemes are initially introduced to model arguments and the ensuing argumentation in legal reasoning [Wal97], the kind of issues that involve expert opinion, witness testimony and circumstantial evidence. It is then presented as a general approach for capturing a wide range of practical reasoning.

The second half of this dissertation presents context sensitive defeasible rules. The relation between argumentation schemes and context sensitive defeasible rules is that practical arguments

generally constitute two parts, the primary reasons and the ancillary reasons [Wal03b, Wal03a].
In his defense of argumentation schemes, Walton discusses the need for this type of arguments.

> " An argument as used in a given case needs to be evaluated not just as a set of premises and a single conclusion, but as an inference drawn in a context on a balance of considerations. "

Argumentation schemes are also associated with the issues regarding burden of proof and the protocols for *argumentation games* [HPW04]. For instance to model the arguments of legal reasoning, each argument scheme is designed to capture certain types of statements. Each statement may involve a number of key elements such as expert opinion [Wal02, Pra01b, PRW03]. Each element may then open up new lines of inquiry and new arguments. The new arguments then follow their own schemes such as the credibility, the relevance, or the impartiality of an expert.

The field of argumentation schemes is fairly young, with many promises. The promises mostly emanate from the range and the procedure of inquiries into *what constitutes an admissible evidence*. In most of the approaches to nonmonotonic reasoning, the *information* is though always taken to be true. [2] However, in reality, of which multiagent systems are an instance, this is hardly true. On the other hand, what constitutes an admissible evidence is one of the pillars of legal reasoning. Consequently, the norms and procedures for a systematic inquiry into establishing the admissibility of evidence has been under constant scrutiny and revision. Obviously, there is much to be learned, on how to obtain and treat the information in a multiagent environment. Such studies will also be helpful in designing protocols for inter agent communication [PS98, PRW05, Wal07, Wal10]. These efforts include the conversational contexts, and the protocols for inter agents' negotiations e.g., negotiation by *persuasion*.

Many of works on the argumentation schemes make references to Toulmin's *general* model of arguments [Tou58]. Toulmin describes that an argument generally has five components,

---

[2] There are some exceptions here and there all related to the research on the nonmonotonic reasoning with non-prioritized information [FKIS02].

data, conclusion, warrant, backing and rebuttal. The data and conclusion correspond to the premises and conclusion of an argument. The warrant, backing and rebuttal are however more intricate topics, subject to various interpretations and schemes. The warrant is the underlying reasoning that connects data to conclusions. The rebuttal is all the situations where the warrant fails to apply. The backing is the higher order reasoning that acts as a warrant for the warrant itself. There are efforts to relate the Toulmin's model of arguments to the current models of argumentation [Ver05, Ver09].

1.1.1.4 Semantics

A large volume of research in argumentation theory is dedicated to the semantics of argumentation systems. This line of research has resulted in various semantics for various argumentation systems. The earliest form of semantics is the *theory of warrant*, presented by Pollock as part of his formal theory of defeasible reasoning. The theory of warrant was then adopted by some early frameworks including [SL92].

The theory of warrant is centered around two concepts, one is the concept of *reinstatement*, and, two, is the shift in the *burden of proof.* An argument that attacks some attackers of an argument, is said to *reinstate* that argument against those attackers. The rationale, here, is that an argument that is not justified cannot reject other arguments. Hence, an argument that successfully attacks an argument, nullifies all the attack force of that argument.

Between the two concepts, reinstatement is the most visible concept within argumentation theory, as it provides an abstraction at the level of what arguments can be labeled *justified* or *admissible*. The notion of burden of proof is different. The difference is that within the notion of burden of proof, there is a dynamic element that involves the search for counter arguments [Pra01a, PRW05]. Respectively, what constitutes an admissible argument is a more lucid and procedural concept. The notion of burden of proof is studied more within the domain of legal

reasoning, argumentation by query and the *argumentation games*.

The theory of warrant is what some view as a procedural semantics. The procedural definitions may also be treated as proof theories. To rectify the distinction between the two sides of theory of warrant, the semantics and the proof theory, new semantics are proposed. These formulations of semantics are the semantics by the *status assignment* or *labeling* [3] [Cam06a, Ver07], and, the *dialectical semantics* which are somewhat the equivalent counterparts of the *dialectical proof theories*.

The semantics by status assignment, labeling, the dialectical semantics and dialectical proof theories are all centered around the concept of reinstatement. An argument is labeled *in* only if it has no attacker or all its attackers are labeled *out*. Otherwise, the argument is labeled *out*. This formulation of semantics defines the acceptability status of an argument by the acceptability status of it attackers. Hence, they are sometimes referred to as the recursive definitions of semantics [PV01].

The theory of warrant and especially the dialectical proof theories all closely mirror the same process that we follow in debates, to seek and decide, the winning arguments. Some of the early works on formalizing the dialectic procedure as a proof theory include the works [SCG$^+$94, PS96a]. The other variations of dialectical proof theories include [CS07, GRS07, DKT06].

The dialectical proof theories are sometimes presented in the form of *argumentation games*. The argumentation games are initially introduced to account for the shift in the burden of proof [PS98]. They then are adopted for the application of standard dialectic proof theory.

Argumentation games consists of two groups of agents engaged in a turn based adversarial argumentation. The first group, the *proponents*, present the first argument. The challengers, the *opponents* then present a counter argument. Then in turn each group present their counter argument to an argument of the other group. It is then easy to see how argumentation games

---

[3] The semantics by status assignment and the semantics by labeling are more or less the same theories.

can be used for the dialectical proof theories. Argumentation games are however multi faceted and involve other elements than just the dialectic proof trees. They involve, a protocol for the game, a protocol for dialogues between agents and the inquiry protocols.

The most widely used semantics are though the fixed point semantics. The basic fixed point semantics are the *grounded, complete, preferred* and *stable* semantics. These semantics are first presented by Dung in conjunction with the introduction of abstract argumentation frameworks.

1.1.1.5 Abstract argumentation systems

The central idea behind the *abstract argumentation systems* is that the rationale by which we decide whether an argument is acceptable or not, is independent of any particular kind of an argument or the internal structure of arguments [Pol87, Dun95b, Vre97].

In this regard, arguments can be regarded as abstract entities represented by some set of symbols, where the only thing that determines the acceptability of arguments is the knowledge of an attack relation between the arguments. The attack relation (a, b) states that argument a attacks argument b.

The importance of abstract argumentation frameworks is well recognized in the literature.

> His (Dung's) article was a major breakthrough in three ways. It provided a general and intuitive semantics for the consequence notions of argumentation logics (and for non-monotonic logics in general); it made a precise comparison possible between different argumentation systems (by translating them into his abstract format); and it made a general study of formal properties of systems possible, which are inherited by instantiations of his general theory [Pra10].

Dung's abstract argumentation framework is the first and the simplest abstract argumentation framework. Dung's abstract argumentation framework AF is a tuple $AF = \langle AR, ATT \rangle$, comprised of a set of arguments AR, representing all (possible) arguments at hand, and, a set of preconceived fixed tuples (a, b), known as the attack relation ATT between arguments. The

centerpiece of Dung's framework is the acceptance relation between a set of arguments and an argument. If a set of S arguments can successfully fend off all the attackers of an argument a then we say S accepts a.

For an argument to be deemed acceptable it needs to be accepted by some set of arguments whose each of its arguments is acceptable. Otherwise, the argument is defended by some unwarranted argument, making the whole defense unwarranted. In abstract argumentation frameworks a set of arguments that can defend their members against any attacker is called an *admissible* set of arguments. Hence,

> we call an argument *admissible* if it is accepted by some admissible set, and,
> *dismissible* if it is attacked by some admissible set.

It is easy to see how the two notions of acceptance and reinstatement are like the two sides of the same coin. This fact is used to draw a two way translation between other proposed semantics, e.g., the theory of warrant, unique and multiple status assignment, and, Dung's presented fixed point semantics. For instance, the semantics by multiple status assignment roughly corresponds to the *credulous* semantics of preferred extensions.

Since their introduction, abstract argumentation frameworks have become the domain for modeling various types of argument interactions and the corresponding semantics. Hence, we can view the field of *abstract argumentation theory* as,

> a study of the properties of the interactions among arguments, based on the given
> set of relations between the abstract arguments.

A large portion of this dissertation can then be said to fall within the field of abstract argumentation theory. In this thesis, we take Dung's abstract argumentation framework as the background framework.

1.1.2 Computational complexity of argumentation frameworks

The computational complexity is the study of computational cost of algorithms in answering *Decision problems*. A decision problem is a problem whose answer is either yes or no (in other words, whether or not an object belongs to a set.) We may now regard a *problem* as a decision problem plus an algorithm for it. The computational cost is denoted by the computational cost of type of the *standard problem* to which a problem can be translated.

A non-deterministic algorithm is an algorithm which takes different routes probabilistically to the answer. Hence, each time it may reach its solution differently. Accordingly, a non-deterministic polynomial time problem, i.e., an NP-problem, is a problem for which there is a non-deterministic algorithm that can reach its solution in a polynomial time. On the other hand, the class of NP-complete problems are the class of problems that although it is easy (can be done in polynomial time) to verify a solution, there is no easy way to find a solution in the first place. The co-NP is the class of problems for which the search for a counterexample is an NP problem. Accordingly, the co-NP complete is the set of decision problems where the search for *no* instances is an NP-Complete problem.

In regard to abstract argumentation theory, we are interested in the questions that relate to the admissibility of arguments. The central questions are whether or not an argument is accepted or attacked by some admissible set. A full study of complexity properties of argumentation frameworks is done by Dunne et al. and Dimopoulos et al. [DT96, DNT99, DNT00, DNT02, DBC01, DBC02, DB04, DW09, DHM+11]. Most of these studies cover the worst case scenarios.

In coherent systems (see background theory section), there are algorithms, e.g., the dialectical proof theory methods, for finding whether or not an argument is credulously or skeptically accepted. The credulous acceptance is found to be an NP-complete problem and the skeptical acceptance to be a co-NPcomplete problem. The NPcomplete and co-NPcomplete problems are regarded to be computationally expensive problems.

In the first part of this thesis, we introduce admissibility backings of arguments. The admissibility backings of arguments are our way to get around the complexity issues of argumentation theory as suggested in [Vre06]. This is despite the fact that we cannot change the nature of complexity problems of argumentation frameworks. In the next section, we present the motivation behind this thesis.

1.2 The motivation

This dissertation consists of two parts. Although, each part deals with a different aspect of formal argumentation theory, they both share a central theme, namely *localization*. The first half of this dissertation investigates the localization of the admissibility of arguments, that is purported by the admissibility semantics in the assumption based argumentation frameworks [KT96, BDKT97]. The second half of this dissertation deals with the localization of a new form of inference rules that covers a large spectrum of practical reasoning.

This thesis is motivated by two underlying principles of reasoning by argumentation. The two principles are,[4]

- arguments localize *reasons*, and argumentation localizes *reasoning*;

- the reasoning by argumentation has two facets, the defeasible reasoning and the reasoning by inquiry.

Neither of these principles are explicitly stated as *principles*. To explain the reason for why these two are principles of reasoning by argumentation is beyond the scope of this dissertation, and, more likely than not requires years of research and experience. Nonetheless, the two

---

[4] The third principle is the place of intentional models of reasoning by argumentation. This principle is the central motivating factor behind the *context sensitive defeasible rules* that constitute the second part of this dissertation. However, there are many different opinions on the role of intentional models of reasoning by argumentation. Consequently, we had to present the motivating factor in terms of the non existence of intensional models of defeasible reasoning. There should however be a strong link between the intensional models of defeasible reasoning and argument schemes.

principles are pointed out by various researchers under various topics. Some of these topics are the localization of reasons [Pea88, Par98], the role of defeaters [Pol87], the role of assumptions [BDKT97], the roles and the lines of inquiry in reasoning by inquiry [Wal08, Wal03a].

1.2.1 Localization

By *localization* in general we mean,

Note 1.2.1. the structuring of the knowledge base such that given a topic, the boundaries between the relevant and the irrelevant information can be effectively and efficiently drawn.

The term *localization* itself is adopted from its use in the works of Judea Pearl [Pea88] in causal reasoning (for further discussion, see appendix 5.2.2). In there, the term *local* is used as *regardless of other things*. To localize the knowledge about an object X, we tie the knowledge of X to only the knowledge of limited number of objects, W, Y, Z, etc. The objects W, Y, Z are then regarded as the *most directly relevant* objects to X. The identification of objects W, Y, Z is guided by a general scheme. The scheme is that if we know W, Y, Z then we can know X and that any other information pertaining to other entities do not add to our knowledge of X. What is *knowable* about an object is still bounded by the domain of knowledge base.

The localization of reasons in argumentation theory is done by drawing concise boundaries over the applicability of inference rules. These boundaries are represented in the form of a connection between the antecedents, conclusion and the defeaters of the rule. In the same breath, we may say that the Toulmin's model of arguments, too, is set to localize the applicability of arguments. Accordingly, there is the implicit assumption that the reasoning by argumentation has an underlying working intentional model. The research on the intentional models of arguments is implicitly discussed under the topic of argumentation schemes. Hence, we claim that,

Note 1.2.2. the localization of defeasible inferences is one of the primary goals in the modeling of arguments.

1.2.2 Argumentation theory is a two pronged approach to reasoning

In the natural world, almost all forms of *reasoning* are done via some form of feedback loops, although, there are some solely feedforward systems, as well. Even the simple task of moving a hand to a target position is done by the constant feedback about the position and the correction of trajectory. The *feedback* information can be obtained passively or actively. The usual forms of perception by the sensory apparatus constitute the passive forms of feedback, e.g., the passive perception by vision or tactile senses. In practical reasoning, the active forms of feedback are generally done by inquiry. The argument for some system of inquiry should then be self evident.

Practical reasoning by nature is both defeasible and inaccurate. The most practical way to make it less defeasible and more accurate is by the *constant correction through better information*. The most practical way to obtain better information is by the *intelligent inquiry*. Hence, an indispensable aspect of practical reasoning is the reasoning by inquiry.

Note 1.2.3. The argumentation theory is (should), therefore, be a two pronged reasoning system. The first part constitutes the reasoning as presented by many argumentation frameworks. The second part should be the *reasoning by inquiry*.

The focus of research has, in large, been on the first phase of argumentation theory. However, in recent years, the researchers are, gradually, turning their attention to the second phase of the reasoning by argumentation which is how to conduct targeted inquiries.

In the literature, the inquiries fall within three categories. One form of inquiry is in regard to argumentation by the adversarial process. The argumentation by adversarial process is usually presented in the form of *argumentation games*. In argumentation games, the inquiry section is part of the protocol or is the algorithm. The protocol is responsible for telling the proponents or the opponents which arguments to present next. This form of inquiry is the most discussed form of inquiry in the literature [PS98, DKT06, GRS07, TDH12].

Another discussed form of inquiry is in relation to the suppositional arguments [Pol90, Bod02].

The suppositional arguments are also presented in the context of assumption based reasoning [TK95]. The third form of inquiry is in relation to the argument schemes [Wal03a]. This form of inquiry is exemplified by the inquiries about the reliability of (expert) witnesses [Ver01b].

A major goal of this dissertation is to assist the reasoning by inquiry, either by a more pragmatic and yet localized defeasible inference rules, or, structure an argumentation framework through the localization of the admissibility of its arguments. The following sections 1.2.3 and 1.2.4 summarize the problems we try to solve in this dissertation as well as the motivation behind them.

1.2.3 Ancillary reasons and context sensitive inference rules

It is argued that the models of arguments that are based on the conventional models of defeasible inference rules do not capture a wide range of practical reasoning. A number of authors then introduced new models of arguments, all grouped under the name argument schemes.

In the spirit of this new impetus, we introduce an alternative representation of defeasible rules that is *context sensitive*. The term *context* is borrowed from its references by Walton and Reed [RG01]. We present these defeasible inference rules in line with the two principles, notes (1.2.2) on page 18 and (1.2.3) on page 19. The main focus is though on the first principle of reducing the gap between the intentional and the extensional models of argumentation systems. In model-theoretic terms, a proper localization should capture, "what we get is what we expect to have". Every application of an infertence has a context. Hence, we made the context explicit in the rule.

We effectively assume that there exists a *local model* associated with an inference rule. The function of these local models is to capture the distinction between the relevant and the irrelevant context. Each local model presents the impact of context on the acceptability of an inference. The applicability is encoded in terms of a status value. The status value ranges over

the *fully acceptable* to *provisionally defeated* to *outright defeated*. In the process, we devise an argumentation system based on these inference rules.

Furthermore, the context is represented in terms of the primary and ancillary reasons. The inference rules are then presented in the usual form, "$\sigma : P \rightarrow Q$, assuming not-R". The antecedent P is still taken as the primary reason to believe in Q and R as the abnormality condition of the rule. For a rule to get triggered, the primary reasons need be present, as usual, and, the rule fails if the abnormality condition is triggered. The distinction is that, now, in addition to the primary reasons, P, there are also ancillary reasons in play. The ancillary reasons represent the circumstances that either strengthen or weaken a given rule. The effect of ancillary reasons is then encoded in a local model associated with the rule. In this manner, the inference rules are still black boxes that both localize and capture the intended underlying explanation associated with an inference [RG01, KR04, RW06].

1.2.4   A case for localizing admissibility of arguments

It is already established in the literature that the best way to investigate the issues regarding admissibility of arguments, is by means of the abstract argumentation frameworks. Accordingly, we present our study into the *localization of admissibility of argument* within Dung's abstract argumentation framework. For the definitions regarding the notion of admissibility, see either section 1.3 or section 1.1.1.5.

The presented semantics of argumentation systems are all, more or less, centered around (i.e., built upon or translated into) admissible sets of arguments.

> A central question in argumentation theory is whether or not an argument belongs to an admissible set.

This central question is followed by its *how-to* question.

> A central problem in argumentation theory is how to find that whether or not an argument belongs to some admissible set.

This problem, in general, is shown to be NP-complete [DBC02]. One thing we need to remember is that not every argument is relevant for the admissibility of an argument. Accordingly, one approach in addressing this problem is to localize the question to only those arguments that play a role in the admissibility of an argument.[5] Within this context now, the problem of localization is how to structure the knowledge such that the information is sorted by its relevance in regard to answering certain questions. In relation to the argumentation theory, the questions for which we seek answers are the ones in regard to the admissibility of arguments. The localization that we seek is therefore the localization of admissibility.

The central goal of the first part of this dissertation is then,

> the formulation and characterization of localization of the admissibility of arguments. This also includes its propagation along the attack sequences, followed by its applications.

The only arguments that are relevant in making an argument admissible (respectively dismissible) are those that belong to some minimal admissible set that either accepts or attacks an argument. It then becomes apparent that the localization of admissibility of arguments, can be formulated in terms of the minimal admissible sets that accept or attack an argument.

We call these minimal admissible sets the *admissibility backings of an argument*. The term backing is borrowed from the role of backings of an argument from Toulmin's model of arguments. In Toulmin's model the role of backings is to *support* an argument [Tou58]. We however distorted the term backings from supporting an argument to making the argument admissible or inadmissible. To ease the discussion, whenever there is no ambiguity we refer to the admissibility backings of arguments by the *backings of arguments*.

Our exploit of the *minimal admissible sets* hinges on one postulate and one assumption. The postulate is that,

---

[5] The term localization here is meant as the reduction of a global set to its relevant subsets.

the minimal admissible sets that accept or attack an argument capture the minimal requirement for the admissibility or dismissibility of that argument. Secondly, a minimal admissible set that accepts or attacks an argument remains such a minimal set irrespective of other arguments in the framework.

In Dung's abstract argumentation framework, this postulate is rather one of the main properties of *admissibility*. In [Dun95b], it is shown the set of all admissible sets form a complete partial order with respect to the set inclusion.

The assumption for the exploit is less straightforward. But, it allows for the implications of the postulate to extend to the twin issues of the dynamics of argumentation theory and the merging of argumentation frameworks. The assumption requires that,

a minimal admissible set that accepts or attacks an argument to remain such minimal admissible set in a *dynamic* argumentation framework.

Obviously, there is no such guarantee. Hence, to bring the assumption into a workable condition we need to make further considerations.

By the manner in which arguments are constructed, given a pool of information, there is a good likelihood that we can construct arguments indefinitely. That is, at any instance of our deliberations, we may have to consider an unbounded pool of arguments. However, though not conclusive, for all intensive purposes, we only need to consider a finite number of arguments. Firstly, every agent is resource bounded. So, it can only construct a limited number of arguments. Secondly, the real world is dynamic. The construction of an unlimited number of arguments is futile, as, they soon become irrelevant. Thirdly, it is known that as arguments get larger in length, there is more chance of their defeat [Lou87]. Hence, we may safely assume that we always work with a bounded pool of arguments.

We assume that we know all the arguments that are in play in a dynamic framework. This assumption makes the backings of arguments ever more workable. We use the backings of arguments to divide a framework into *independent* sub-argumentation frameworks. The full scope of how an argumentation framework can be divided into sub-frameworks of distinct character-

istics, is outside the scope of this thesis. However, to showcase the range of applications of the backings of arguments, in the last section of the first part of this dissertation, chapter three, we show how an argumentation framework can be partitioned into *independent* sub-frameworks. The independence relation between each sub-framework is technically governed by where the propagation of the backings halts.

How this lends itself to the dynamics of argumentation theory is that the arrival of new arguments should only impact the relevant sub-argumentation frameworks and leave the independent frameworks untouched. Moreover, if the new arguments are regarded as some sub-argumentation framework themselves, then, the problem of addition of new arguments can be framed in terms of the merging of two argumentation frameworks.

In short, to address the central problem of whether or not an argument is admissible that is inherently a NP-complete problem, given that we only deal with a finite number of arguments, we can reduce the search area to answer this question to a much smaller area.

From the above discussion, it is evident that we need to set up some basic structures for investigating many important issues in argumentation theory. Hence, we define the notion of *sub-argumentation framework relation*. In this thesis, we do not however explore the potentials of the sub-argumentation framework relation. It is, though, easy to see how the sub-argumentation framework relation lends itself to discussions on a range of important topics such as the equivalence relation between frameworks, the dependency relation between arguments, the admissibility preserving mapping between frameworks, and the preservation of admissibility of a set of arguments under certain conditions.

The utility of backings of arguments is not though limited to the efficient calculation of admissible sets or the dynamics and the merging of frameworks. Another utility of backings of arguments is in regard to the roles that arguments play in regard to the acceptability of another argument. The reason for why such knowledge is important is self evident. For instance, if we

wish to safeguard the admissibility of a certain argument, we need to know which arguments are instrumental for the defense of that argument and fortify them.

> The backings of arguments also provide a perfect ground for identifying which arguments have a hand in the admissibility of an argument, and, what are their significance.

A number of relationships between arguments are identified in the literature. The relationships are all based on the attack relation. Naturally, these relationships cast some net over whether or not an argument can have a role in the admissibility of another. The most primary relation is the attack relation itself. The attack relation identifies the role *attacker of an argument*. The next relations are the defense relation, the indirect attack relation and the indirect defense relation.

These classifications of argument relations do not however identify whether or not the labeled arguments do contribute to the admissibility of the target argument. For instance, the label *attacker* does not necessarily imply that an attacker of an argument plays a part in rejection of the argument. An attacker that itself is rejected, cannot reject an argument.

On the other hand, any member argument of some backing of an argument should play some role in the admissibility of that argument. Accordingly, we can say an argument is *active* for some argument, if it takes part in some backing of that argument. The followup question is then,

> whether or not the *active argument relation* is a transitive relation, and, if not, then, under what conditions it is transitive.

The answer to above question brings us to the issue of propagation of backings of arguments under the attack relation. What we mean by the propagation of backings, is the subset relation between the backings of arguments. For instance, for an argument a and its defender b, is every backing of a subset of some backing of b? The answer to this question also let us identify new relations between arguments.

The first half of our work on the backings of arguments deals with the formulation, elaboration

and characterization of the backings of arguments. The second half addresses the propagation of backings of arguments along the sequence of attack relations, and, its connection with the roles of arguments in regard to the admissibility or dismissibility of an argument. As already mentioned, the propagation of backings can then help in splitting an argumentation framework into sub-frameworks of distinct characteristics.

For any inquiry to be efficient and effective, we have to know what questions to ask. A structured knowledge base helps us in both the formulation of right questions and how to efficiently answer them. Following our discussion, we can see how the admissibility backings of arguments serve the second feature of the reasoning by argumentation, note (1.2.1) on page 18, which is reasoning by inquiry.

Before we start, we should make a few additional remarks that would help on how the definitions and results are constructed here. The notions of *acceptance* and *admissibility* are taken to be the most fundamental concepts in the formal argumentation theory. These notions therefore transcend their particular formulation in Dung's framework.

Hence, where it is feasible, we center the definitions directly around the admissible and conflict free sets. There are however instances where the definitions require references to some additional concepts. For instance, many of the presented relations embed some form of transitivity property. By a transitivity property we mean the manner by which a certain construct propagates (forward or backward) along the attack sequences. In this regard, the use of attack sequences provides the most feasible way to capture a number of properties within a definition.

Furthermore, in the chapters regarding the backings of arguments, we make an attempt to follow the following general format. Each discussion on a presented construct is comprised of three parts. The first part introduces the construct and its definition. The second part deals with the formulation of how the construct is realized. For instance, the conditions that would result in two arguments to become incompatible. We call these formulations, the characterization of the

construct (they can also be regarded as the base cases). Finally, the third part discusses the propagation of the construct under the attack relation or along the lines of attack sequences. In the next section, we discuss the background theory which is an account of Dung's abstract argumentation framework.

1.2.5   A summary of thesis achievements

In this section we provide a concise account of the achievements of this thesis. The major achievements of each chapter are generally identified by the headings of each section of these chapters. The problems that this thesis intends to address are explained in sections 1.2.4, 1.2.3 and the conclusion chapter.

Chapter 2 is entirely dedicated to the localization of admissibility (or dismissibility) of arguments within Dung's abstract argumentation framework. The sets of arguments that localize the admissibility of an argument are called the *admissibility backings* of that argument. The admissibility backings of arguments are usually referred to as the *backings of arguments*. The backings can be of two types, the *positive backings*, the backings that accept an argument, and, the *negative backings* the backings that attack an argument. The major achievement of this chapter are as follows.

1. The *Sub-argumentation frameworks*, and the *normal sub-argumentation frameworks* are presented to serve the formal setting under which all the finding of this thesis and subsequent future works can be succinctly represented. How the sub-argumentation frameworks can be utilized for further studies is discussed in section 5.2.1.3 of the concluding chapter.

2. A formal definition of the *grounded admissible sets* is given. Every argument that belongs to a grounded admissible set can be accepted beyond a reasonable doubt. Hence, it is important to know which backings of arguments are grounded admissible sets. The

grounded backings that are grounded admissible sets are referred to as the *grounded backings* of an argument. Grounded backings are defined in section 2.4.

3. A detailed presentation of the *admissibility backings* of argument is given in section 2.4.

4. In section 2.4, it is shown that the findings in [Dun95b] presented in section 2.4.2, can be recast in terms of the admissibility backings of arguments. Hence, the backings of arguments can be considered as a viable construct within the abstract argumentation frameworks.

5. In section 2.5, it is shown that there is a well defined relation between the backings of an argument and the backings of its attackers. The relation between the backings of an argument and those of its attackers are formulated in terms two algebraic operators $\sum$ and $\prod$. A full list of the properties of the operators $\sum$ and $\prod$ are given in appendix 5.2.2. The presented relation is in form of an equation. However, this equation cannot always be easily solved. Hence, we presented a *recursive algebraic function* by which the backings of an argument can be calculated. The function is called the *backing function*.

6. An important result of chapter 2 is using the backings of arguments to draw a precise boundary over all arguments that are relevant to the admissibility or rejection of an argument (section 2.6). This is an important result because it directly relates to the dependency relation between arguments. A distinct property of the admissibility backings is that each backing operates independent of other arguments in the framework. The finding presented here identifies the minimal sub-frameworks by which the admissibility status of an argument is preserved under the sup-argumentation framework relation.

7. A major finding of chapter 2 is to provide the formal characterization of classes of frameworks that are most suited to the motivation behind the backings of argument. The motivating principle in question is that the admissibility backings of an argument should carry *all the information* pertinent to the admissibility situation of that argument. The

identified class of frameworks are called the *normally stable and compact* argumentation frameworks.

Since, the *admissibility backings* localize the admissibility of arguments, they provide a suitable ground for investigating the functions that arguments may serve with respect to the admissibility of other arguments. This is the focus of chapter 3 which uses the backings of arguments to identify and characterize a number of roles that arguments play in regard to the admissibility of another.The presented results can also be used for partitioning a framework into sub-frameworks with distinct features. This would also serve as a prelude to the future work on the applications of the *admissibility backings* in the partitioning and the merging of argumentation frameworks. The major achievements of chapter 3 are as follows.

1. Four major argument relations are formally characterized. Their characterization also includes how these relations propagate along the sequences of attack relations, and, how each role is related to another such that they exhaustively cover the sequences of attack relations in an argumentation framework. Hence, they comprehensively address the propagation of admissibility backings along the sequences of attack relations. These presented relations can also be viewed as the roles one argument plays in regard to the admissibility of another. The identified roles are the *intercepted for, intercepting, critical for, incompatible with*, and, *redundant for* arguments.

2. Another major achievement of this thesis is use of the intercepts to partition a framework into *independent* sub-frameworks. These sub-frameworks are called the *disjointed by intercept sub-frameworks*. The independence between such sub-frameworks states that any change in one such sub-framework will not affect any of its non adjacent sub-frameworks. This statement is one of the results of section 3.7.

Chapter 4 presents a new type of defeasible inference rules that are suitable for the common instances of practical reasoning. The presented inference rules are called *context sensitive rules*.

They reflect the fact that in many instances of practical reasoning, we not only use the primary reasons to see whether or not the application of a specific inference rule is allowed, but, we also use the ancillary reasons to fine-tune the application of the inference rule. The major achievements of chapter 4 are as follows.

1. A defeasible argumentation framework is built based on the presented context sensitive defeasible rules. The presented framework can augment the systems based on the conventional default rules. The semantics that is provided for this framework is by translating the framework into Dung's abstract argumentation framework. The presented framework therefore can be read under the standard semantics in the literature.

2. In section 4.2, by means of two examples it is shown that the current representations of defeasible inference rules in the literature cannot account for the required reasoning, if they are formulated and applied as intended. The *required reasoning* is a reasoning that matches the *expected outcome*. In classical logics this property of the reasoning is referred to as the *soundness* of the reasoning.

3. In this chapter we introduce the notion of *missing arguments*. The missing arguments are the arguments that are either unknown or hidden to an agent, as shown by the examples in the motivation section 4.2. The importance of the missing arguments is that they provide both an explanation and an account of the deficiencies in the treatment of attack relation between conventional arguments. The deficiencies are highlighted with the two motivating examples. It is the missing arguments that facilitate the means to translate the presented framework into Dung's abstract argumentation framework.

This sums up the major achievements of this thesis. Next, we present Dung's abstract argumentation framework that serves as the background theory for this thesis. Dung's framework is the most discussed work in the literature, and hardly needs any introduction. We therefore present Dung's framework within the context of this thesis. This includes stating and explicating all the

relevant concepts and the findings of the paper [Dun95b] such as indirect attack relation, the attack cycles and the grounded extensions. The background theory also includes the notations and conventions that are consistently used within chapters 2 and 3 of this thesis. Chapter 4 uses its own notations and conventions that are defined within the motivation section of chapter 4. Thesis also includes an index of definitions and the defined symbols, e.g., the *positive backings* or $\langle a \rangle^+$.

1.3 Background theory, notations and conventions

In the first part of Introduction chapter, we provided a very short account of the basic idea behind the abstract argumentation theory and Dung's abstract argumentation framework. Dung's abstract framework is built around three central constructs. The three constructs are the *conflict free sets*, the *acceptance relation* and the *admissible sets*. Upon further studies, it is both suggested and shown that the three notions transcend any particular argumentation framework.

Dung's abstract argumentation framework AF is a simple tuple $AF = \langle AR, ATT \rangle$ where AR is a pool of abstract arguments and ATT is a binary attack relation between arguments. The binary attack relation (a, b) denotes that argument a attacks argument b.

Definition 1.3.1 (Dung95)**.** An argumentation framework AF is a pair $AF = \langle AR, ATT \rangle$, where AR is a set of arguments, and ATT is a binary attack-relation on AR, i.e., $ATT \subseteq AR \times AR$.

We first introduce the adopted notations and conventions.

1.3.1 Notations and conventions

1.3.1.1 Notations

The multi character names denote the special objects regarding an abstract argumentation framework, e.g., AR, ATT. The lower case English letters $a, b, c, a_i, \cdots$ are used to denote arguments, the upper case letters $A, B, C, A_i, \cdots$ the set of arguments, the upper case Calligraphy letters $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{A}_i, \cdots$ the set of sets of arguments, and, the bold capital letters $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{A}_i, \cdots$ the set of sets of sets of arguments. Hence, there is an ordering of kind $a \in AR$, $A \in 2^{AR}$, $\mathcal{A} \in 2^{2^{AR}}$, $\mathbf{A} \in 2^{2^{2^{AR}}}$.

The functions and the objects outside argumentation framework are represented by Greek letters. Moreover, throughout this thesis, any definition or result that is not original to this dissertation is marked by by its cited paper. For instance, *Dung95* says that the corresponding result comes from the paper [Dun95b].

1.3.1.2 Conventions

In this dissertation, we adopt the following terms and abbreviations. We say,

- a attacks b or a is a *direct attacker* of b and write $a \hookrightarrow b$ if $(a, b) \in ATT$;

- a attacks a set S of arguments and write $a \hookrightarrow S$ if a attacks some $b \in S$;

- S attacks a and write $S \hookrightarrow a$, if some $b \in S$ attacks a;

- S attacks a set T of arguments if S attacks some $b \in T$ and write $S \hookrightarrow T$;

- a is a *direct defender* of b if a attacks some attacker of b;

- a is *admissible*, if a is accepted by some admissible set (the *acceptance* relation is defined later), and, a is *dismissible*, if a is attacked by some admissible set.

We denote the set of attackers of an argument a by $\overline{a} = \{b \mid b \hookrightarrow a\}$, and, the set of all arguments that a attacks by $\overset{+}{a} = \{b \mid a \hookrightarrow b\}$. Correspondingly, we denote the set of attackers of a

set S of arguments by $\overline{S} = \{b \mid b \hookrightarrow S\}$, and, the set of all arguments that S attacks by $\overset{+}{S} = \{b \mid S \hookrightarrow b\}$.

In addition, we often use the term *in/direct* as an abbreviation for the expression *direct or indirect*, and, where there is no ambiguity we use shorter names and references. For instance, by *admissibility backings* or *backings* we mean admissibility backings of arguments. Moreover, in the thesis we often need a second framework AF^* along with AF. In such instances, without explicitly stating, by an argumentation framework AF^*, we mean $AF^* = \langle AR^*, ATT^* \rangle$.

It is a common practice to represent abstract argumentation frameworks in the form of graphs. The graphs are usually directed graphs where the nodes of a graph represent arguments and the edges represent attack relation. Hence, the attack relation $b \hookrightarrow a$ is represented by a directed edge from node b to node a.

1.3.2 Dung's abstract argumentation framework

The first construct regarding the acceptance of arguments is the *conflict freeness*. The notion of a conflict free set dictates that if we are to accept a set of arguments then the arguments should be collectively acceptable. Otherwise, we protest that the painted picture is not a rational picture. In Dung's framework, the criterion for a conflict free set of arguments is reduced to the conflict between the individual member arguments [Boc02].

Definition 1.3.2 (Dung95)**.** A set S of arguments is said to be *conflict free* if and only if there are no arguments a, b in S such that a attacks b.

Observation 1.3.3. *A set S of arguments is conflict free if and only if $S \cap \overline{S} = \varnothing$.*

The notion of conflict freeness, however, transcends its formulation in Dung's framework. The lottery paradox presents why the following definition of conflict freeness does not cover all cases of conflict freeness. In the lottery paradox, all agents present the same argument that they will not win the lottery. No argument is then in conflict with another argument, i.e., no argument

attacks another argument. On the other hand, the lottery being a fair lottery, somebody has to win the lottery. Hence, the set of all such arguments cannot be considered a conflict free set (somebody has to win the lottery). To address this issue we can define *conflict freeness* with respect to a set of arguments such that a set of arguments is conflict free if and only if it does not attack itself. This alternative definition is in fact presented as one of the main lemmas in Dung's Framework, see lemma 1.3.9 below. A more formal approach is adopted in [Boc02, KT99] where the notion of an attack relation includes the conflicts between sets of arguments.

The centerpiece of Dung's framework, as distinctly highlighted in [Dun95b], is the *acceptance* notion. We say, a set of arguments accepts an argument, if it defends that argument against any and all attackers.

Definition 1.3.4 (Dung95). An argument $a \in AR$ is acceptable with respect to a set S of arguments if and only if S attacks any argument $b \in AR$ that attacks a.

As a consequence, one can see that arguments that have no attackers are vacuously accepted by any set of arguments, including the empty set.

Observation 1.3.5. *A set T of arguments is accepted by S if and only if $\overline{T} \subseteq \overset{+}{S}$.*

The acceptance relation can also be formulated into an identifying marker for each particular framework, named the characteristic function.

Definition 1.3.6 (Dung95). The characteristic function, denoted by θ_{AF}, of an argumentation framework $AF = \langle\, AR\,,\, ATT\,\rangle$ is defined as follows.

$$\theta_{AF}\colon\ 2^{AR}\ \longrightarrow\ 2^{AR}$$

$$\theta_{AF}\,(S) = \{\, x \mid\ x \text{ is acceptable wrt } S\,\}.$$

It is not difficult to see that the characteristic function is monotonic with respect to set inclusion.

Lemma 1.3.7 (Dung95). *θ_{AF} is monotonic with respect to set inclusion.*

The next fundamental construct is the *admissible set*. A set of arguments is deemed to be admissible if it can defend its members against any attacker while remaining conflict free.

Definition 1.3.8 (Dung95). A conflict free set S of arguments is said to be *admissible* if and only if each argument in S is acceptable with respect to S.

Hence, we may say that, a set of arguments is admissible if it accepts itself. This result is articulated as one of the main lemmas of [Dun95b].

Lemma 1.3.9 (Dung95). *For a set S of arguments in an argumentation framework AF,*

1. *S is conflict free if and only if $S \not\rightarrow S$.*

2. *A conflict free set S is admissible if and only if $S \subseteq \theta_{AF}(S)$.*

The admissible sets are therefore the main yardstick by which to phrase and answer the inquiries regarding the acceptability of arguments. For instance, whether there are some admissible sets that accept (or attack) a certain argument. We may also assign *statuses* to arguments that characterize such general inquiries.

Definition 1.3.10. In an argumentation framework AF, we say that an argument is justified or has the status *justified* if it is attacked by no admissible set; the argument is *overruled* if it is accepted by no admissible set; and, an argument is *provisionally defeated* if it is both accepted by some and attacked by some admissible set. This is known as the *admissibility status* of an argument. Alternatively, we can represent the admissibility status of an argument by the status assignment function, $\epsilon_{AF} \colon AR \longrightarrow \{0, \frac{1}{2}, 1\}$. The values $0, \frac{1}{2}, 1$ respectively correspond to the status overruled, provisionally defeated and justified.

One argument allows us to draw a few conclusions about a subject matter. Two arguments draw a better picture than one argument, and three arguments give us more than two. Given the full set of arguments AR, we can select a number of subsets of AR to draw conclusions about a subject matter. The general question is then which subsets of AR to choose. To streamline such lines of inquiries, the admissible sets are categorized into some general classes. These

classes are often presented in the form of semantics of argumentation theory. In nonmonotonic reasoning, we tend to refer to such classes the *extensions* of the framework. More accurately, an extension is usually a complete subset of AR that adheres to some given criteria. The following are some of the extensions of AR that categorize the admissible sets of a given argumentation framework.

Definition 1.3.11 (Dung95)**.** Let S be an admissible set in an argumentation framework AF.

1. S is called a *complete extension* if and only if each argument that is acceptable with respect to S belongs to S.

2. S is said to be a *preferred extension* if and only if S is a maximal admissible set in AF (maximal with respect to set inclusion).

3. S is called the *grounded extension* of AF, denoted by **GR** if and only if **GR** is the least fixed point of θ_{AF}.

4. S is called a *stable extension* of AF if and only if S attacks any argument that does not belong to S.

It is easy to see that a stable extension is a maximal conflict free set that attacks every argument outside itself.

Observation 1.3.12. *For a set S of arguments in an argumentation framework AF, S is a stable extension if and only if $S = AR - \overset{+}{S}$.*

If we are interested in the fate of all the arguments, then the stable extensions are what we need. But, as it is usually the case, the scope of our interest is narrow and concern only a portion of all possible arguments. In this regard, the complete extensions provide a more local and suitable information. To address such intents, the finer grained categories of admissible sets are presented and categorized, e.g., the semistable semantics [Cam06a, Cam06b] or the SCC-recursive semantics [BGG05].

But, how fine grained a category of admissible sets should be so that we can answer whether or not an argument is admissible? The main theorem of [Dun95b] states that any admissible set that accepts or attacks the argument will do. The admissibility semantics in [BDKT97] follows this finding.

Theorem 1.3.13 (Dung95). *In an argumentation framework AF,*

1. *the set of all admissible sets form a complete partial order with respect to set inclusion.*

2. *For each admissible set S, there is some preferred extension E such that $S \subseteq E$.*

The theorem says that an admissible set remains admissible regardless of other arguments in AR. As highlighted in the introduction, this is the property around which the first half of this dissertation is built. The theorem also shows the reason for the admissible sets being regarded as the basis by which the semantics of argumentation frameworks are defined.

The second part of the theorem puts the result of the first part into perspective by drawing the subset relation between the various extensions. It is noted in [Dun95b] that since the empty set is always admissible, every argumentation framework has at least one preferred extension.

Example 1.3.14. The following example summarizes the background theory so far. The figure below shows the argumentation framework $AF_1 = \langle AR, ATT \rangle$ where $AR = \{a, b, c, d, e\}$, $ATT = \{(d, c), (c, b), (b, a), (a, b), (b, e)\}$.

We can see that the set $S = \{d, c, b\}$ accepts b, but, it is not a conflict free set. On the other hand, $T = \{c\}$ is a conflict free set and accepts e. However, T itself, is not an admissible set.

$$d \longrightarrow c \longrightarrow b \rightleftarrows a \qquad e$$

$$AF_1$$
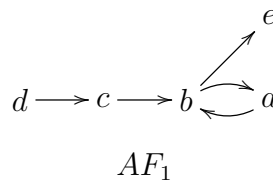
The admissible, complete, preferred and grounded extensions of AF_1 are:

Admissible sets: $\varnothing, \{d\}, \{d, a\}, \{a\}, \{d, b\}, \{a, e\}, \{d, a, e\}$.

Complete extensions: $\{d\}$, $\{d, b\}$, $\{d, a, e\}$.

Preferred extensions: $\{d, a, e\}$, $\{d, b\}$.

Stable extensions: $\{d, a, e\}$, $\{d, b\}$.

Grounded extensions: $\{d\}$.

Following the above extensions, we see that

- Every grounded extension is a complete extension but not vice versa.

- Every preferred extension is a complete extension but not vice versa.

- Every stable extension is a preferred extension but not vice versa.

Moreover, as a pretext for the *admissibility backings* of arguments, we see that not every complete extension is a minimal admissible set that accepts some argument. For instance, the set $W = \{a\}$ is a minimal admissible set that accepts a, but, W is not a complete extension. ∎

The coherent frameworks are of special interest to the admissibility backings. The most straightforward class of coherent frameworks is the *limited controversial* frameworks, see theorem 1.3.16 below. A limited controversial argumentation framework is simply a framework that does not contain any attack cycles of the odd length.

Definition 1.3.15 (Dung95)**.** An argumentation framework is said to be *coherent* if each preferred extension of AF is stable. We say that an argumentation framework is *relatively grounded* if its grounded extension coincides with the intersection of all preferred extensions. An argumentation framework is said to be *limited controversial* if there exists no infinite sequence of arguments $a_0, a_1, a_2, \cdots$ such that a_{i+1} is controversial with respect to a_i.

Theorem 1.3.16 (Dung95)**.** *Every limited controversial argumentation framework is coherent.*

Next, to highlight the role of arguments in relation to the admissibility of other arguments, we borrow a number of terms from [Dun95b], e.g., an attack cycle. The first of these terms are the

indirect attackers and the indirect defenders of an argument. The indirect attack and the indirect defense relations are originally defined by means of attack sequences. The attack sequences are also referred to as attack paths.

Definition 1.3.17 (Dung95). An argument b is said to be an *indirect attacker* of an argument a if there exists a finite sequence of arguments $a_0, a_1, \cdots, a_{2n+1}$ such that $a = a_0, b = a_{2n}$ and for each $0 \le i \le n$, a_{i+1} attacks a_i. Respectively, b is said to be an *indirect defender* of a if b attacks some indirect attacker of b. In addition, a is said to be a *controversial* argument for b if and only if a is both an in/direct attacker and an in/direct defender of a. (The term *in/direct* is read as direct or indirect).

Note that the in/direct attack, the in/direct defense, and the controversial relations, all are transitive under the attack relation. For instance, if a is controversial for b and $b \hookrightarrow c$ then a is controversial for c as well. Next, we define the sets of arguments that form an attack cycle. Although, we do not show it here, it is easy to see that in a limited controversial argumentation framework, no two arguments in a cycle can be controversial with respect to another.

Definition 1.3.18 (A Variation of Dung95). A set L of arguments is called a *cycle* or to form an *attack cycle* if and only if for any $a, b \in L$, there is some sequence of argument $a_0, a_1, \cdots, a_n$ where $a_{i+1} \hookrightarrow a_i$, $a_0 = a$, $a_n = b$.

Going back to example 1.3.14, the arguments a, b form an attack cycle (of even length). The argument d is an indirect defender of the argument b and an indirect attacker of the argument a. Many studies are carried out on Dung's framework. These studies mostly deal with the formation of the attack relation and and the ensuing semantics, e.g., the occurrence of odd length attack cycles and how to treat them. But, we conclude the background theory on Dung's framework here.

Chapter 2

The Localization of Admissibility and the Admissibility Backings

2.1   Introduction

Since their inception, abstract argumentation frameworks have become a potent mean for studying the various aspects of arguments interactions. Much of the research in this area is, though, focused on defining the new semantics. The proposed semantics are generally in form the maximal sets of arguments that pertain certain characteristics.

In this chapter, we, however, take the opposite route, and attempt to localize the admissibility of arguments, by means of minimal admissible sets that either accept or attack an argument. We call these admissible sets, the *admissibility backings* of arguments. This chapter is focused on the development, definition, and, investigation of the admissibility backings of arguments. Accordingly, we identify certain classes of coherent argumentation frameworks that suit best this endeavor.

In Dung's argumentation framework, the set of all admissible sets form a complete partial order with respect to the set inclusion. Hence, we can safely say that the admissibility backings of an argument capture the minimal requirements for the admissibility or the dismissibility of an argument.

41

This chapter is structured as follows. In section two, we introduce the sub-argumentation framework relation, but only as far as their relevance to this chapter. In section three, we present the *grounded admissible sets* and their relation to the *ground arguments*. In section four, we present the admissibility backings of arguments. In doing so, we first discuss the role of minimality clause in definition of admissibility backings. We then present the relation between the admissibility backings of an argument and of its attackers, and, the corresponding recursive function with respect to the attack relation. We conclude this section by discussing the independency of the admissibility backings. In section five, we present a class coherent argumentation frameworks that reflect the admissibility backings best. In the last two sections, we discuss the related research and conclusion.

2.2 Sub-argumentation frameworks

We build arguments based on the factual information at hand. When we receive new information we naturally build new arguments. The addition of new arguments may, however, impact our current findings about the standing arguments. The study of the changes in current findings with respect to an argumentation framework upon addition of the new arguments is usually referred to as the dynamics of argumentation theory. [1]

One way to represent the addition of a new set of arguments A, is by extending the current set AR to $AR \cup A$. Another way is to have some function α that selects the arguments at hand from a universal pool of arguments, where $\alpha(A) = AR \cup A$. We however opt to use the sub-argumentation frameworks. In this manner, the dynamics of argumentation theory are discussed, not in terms of the addition of new arguments, but presented with respect to the sub-argumentation framework relation.

---

[1] In the abstract argumentation frameworks, the addition of new arguments is expressive enough to represent the *subtraction* of the arguments. We can model the subtraction of an argument by adding the attacker of the argument to the mix.

Definition 2.2.1. An argumentation framework $AF' = \langle\, AR'\,,\ ATT'\,\rangle$ is a *sub-argumentation framework* of $AF = \langle\, AR\,,\ ATT\,\rangle$, written as $AF' \sqsubseteq AF$, if and only if $AR' \subseteq AR$, $ATT' \subseteq ATT$. AF' is said to be a *normal sub-argumentation framework* of AF, written as $AF' \sqsubseteq^N AF$, if and only if, $AF' \sqsubseteq AF$ and for all $a, b \in AR'$, if $(a, b) \in ATT$ then $(a, b) \in ATT'$.

Observation 2.2.2. *The sub-argumentation framework relation and the normal sub-argumentation framework relation are both partial order relations.*

The proof of this observation and the following results are given in appendix.

If an argument a does not play a role in relation to the acceptability of argument b, then we can say the acceptability of b is *independent* of a. The following observation 2.2.4 crudely highlights this independence relation with respect to the acceptability of arguments. The observation says that the addition of new arguments or the attack relations that are not connected to a current situation, do not affect the admissibility of the current arguments.

Definition 2.2.3. A sub-argumentation framework $AF' = \langle\, AR'\,,\ ATT'\,\rangle$ of an argumentation framework AF is said to be *closed* under the attack relation in AF if and only if AF' is a normal sub-argumentation framework of AF, and, for all arguments $a \in AR', \overline{a} \subseteq AR'$.

Observation 2.2.4. *A sub-argumentation framework AF' of AF is closed under the attack relation in AF if and only if, for every argument a in AF', if b is an in/direct attacker or in/direct defender of a in AF then b is in AF' as well.*

Next, we represent the main result of this section that casts a wide net over the independence relation between arguments. We present a more refined version of this result, in the admissibility backings section 2.4.

Theorem 2.2.5. *Let $AF' \sqsubseteq^N AF$ be closed under the attack relation in AF. Then, if S is an admissible set in AF', S is admissible in all AF'' where $AF' \sqsubseteq AF'' \sqsubseteq AF$.*

Working with sub-argumentation frameworks has two additional benefits. Dung's abstract argumentation framework is the simplest of the presented abstract argumentation frameworks.

It however serves as the basis for many other abstract argumentation frameworks, known as the *extended abstract argumentation frameworks*. In many of these frameworks, the attack relation, is effectively more than a binary relation between arguments, e.g. [Ben02, Mod06, MGS08].

For instance, in [Ben02], an attack from a to b is counted in the conventional sense only if b is not preferred to a where the preference ordering among arguments is set by external measures. Hence, to study the dynamics of argumentation theory, it is prudent to have the versatility of playing with the attack relations as well as the pool of arguments, all under the same setting. Working with sub-argumentation frameworks allows for such versatility.

The second advantage of working with the sub-argumentation frameworks is that it provides room for the classification of the argumentation frameworks, where an argumentation framework can be described in terms of its comprising classes of sub-frameworks. This formulation of argumentation frameworks can in turn be utilized in relation to other important research areas such as the process of *conjoining argumentation frameworks*, e.g., [CMDK$^+$07]. Under this approach the process of conjoining the argumentation frameworks can be investigated in terms of the conjoining of their sub-argumentation frameworks. All these topics are beyond the scope of this thesis, but to avoid the recast of our findings in our future works, we opted to set the setting for the dynamics of argumentation theory here.

2.3 Grounded admissible sets

The grounded extensions have a special place in both the argumentation theory and most of the approaches to nonmonotonic reasoning. If we draw a mapping between arguments and the statements that those arguments say about the world, then the statements that a grounded extension makes can be believed beyond a reasonable doubt. This is a distinct property of grounded extensions.

In the search for the backings of an argument, it is therefore important to identify which backings have this property of the grounded extensions. That is, which backings provide a beyond a reasonable doubt defense for an argument.

This feature of grounded extensions is tightly connected to the property of the *ground arguments*, those arguments that have no attackers. We however define the ground arguments based on the universality of their acceptance.

Definition 2.3.1. An argument a is said to be a *ground argument* if and only if $a \in \theta(\varnothing)$.

Observation 2.3.2. *(1) a is a ground argument if and only if $\overline{a} = \varnothing$, and, (2) For any set S of arguments, $\theta(\varnothing) \subseteq \theta(S)$.*

Given the distinct property of the grounded extensions, we would wish to identify other admissible sets with this property that if we are given only the admissible set, every argument in the admissible set can be accepted beyond a reasonable doubt. We call such admissible sets, the *grounded admissible* sets (see definiton 2.3.4 below). The obvious approach for testing whether an admissible set is grounded or not is to check if it is a subset of the grounded extension. However, as the following example illustrates this approach is not conclusive.

Example 2.3.3. The following argumentation framework, AF_2, has five nonempty admissible sets $S_1, \cdots, S_5$, with S_3 being the grounded extension.

$$S_1 = \{a\}, S_2 = \{a, c\}, S_3 = \{a, c, e\}, S_4 = \{c\}, S_5 = \{c, e\}.$$

$$a \longrightarrow b \rightleftharpoons c \longrightarrow d \longrightarrow e$$

$$AF_2$$

In this example, the admissible sets $S_1, \cdots, S_5$ are subsets of the grounded extension S_3, however, only S_1, S_2, S_3 are *grounded* in AF_2. For instance, if we take the admissible set $S_4 = \{c, e\}$, neither c nor e can be accepted beyond a reasonable doubt with respect to S_4. The

argument c does not conclusively defeat b. For the conclusive defeat of b we need argument a. But, $a \notin S_4$, hence S_4 cannot be considered a grounded admissible set. ∎

Example 2.3.3 suggests that the identification of the grounded admissible sets needs to be done via their ground arguments. Hence, we tie the grounded admissible sets to their ground arguments.

Definition 2.3.4. In an argumentation framework AF, the empty set $\varnothing$ is a *grounded admissible set* . Morevoer, if S is a grounded admissible set and S accepts a then $S \cup \{a\}$ is a grounded admissible set, too. We denote the set of all grounded admissible sets by $\mathcal{G}$.

Definition 2.3.5. An argument a is said to be *grounded* in an admissible set S if and only if $a \in S$ and there is some grounded admissible set $T \subseteq S - \{a\}$ such that T accepts a.

A special feature of the characteristic function in relation to the ground arguments is the cumulative property where $\theta^i(\varnothing) \subseteq \theta^{i+1}(\varnothing)$. This suggests that we can alternatively define the grounded admissible sets based on this cumulative structure that is entirely rooted in the ground arguments. This property of the grounded admissible sets is presented in theorem 2.3.6 below. To do this we need to identify the intended cumulative structure within a given set.

A comprehensive approach would be to define some *non-trivial characteristic* function $\widehat{\theta}(S)$ that isolates the ground arguments responsible for the admissibility of a set S. A function that accepts only the arguments that are non trivially defended by the set, e.g., $\widehat{\theta}(S) = \theta(S) - (\theta(\varnothing) - S)$. We however choose a simpler approach by curbing the characteristic function to some given set S : $\widehat{\theta}(T) = \theta(T) \cap S$.

Theorem 2.3.6. *A finite set S of arguments is a grounded admissible set if and only if there is a natural number m such that $S = \widehat{\theta}^m(\varnothing, S)$ where $\widehat{\theta}(T, S) = \theta(T) \cap S$.*

As a consequence of the above theorem, we have what we expected from the start that the grounded extension $\mathbf{GR}$ is the maximum grounded admissible set, i.e., $\mathbf{GR} = \bigcup_{S \in \mathcal{G}} S$, and that any subset of $\mathbf{GR}$ must be rooted in some grounded admissible set subset of $\mathbf{GR}$.

Lemma 2.3.7. *The grounded extension* **GR** *is the maximum (with respect to the set inclusion) grounded admissible set. If* $S \subseteq$ **GR** *is admissible then there is some minimal grounded admissible set* S' *such that* $S \subseteq S'$.

Due to the importance of the grounded extensions, there are many questions of interest. For instance, what are the maximal sub-frameworks and the minimal sup-frameworks of an argumentation framework in which a set of arguments is a grounded admissible set? However, these lines of inquiry are not directly related to this thesis. Hence, we end our discussion on the grounded admissible sets, and, move to define the admissibility backings of arguments.

2.4 Admissibility Backings

The British logician Stephen Toulmin, as part of his general characterization of arguments, presented a six-part model of an argument [Tou58]. He explains the *warrant* of an argument, as the component of the argument that licenses the inference of the claim form the data of the argument. The *backing* of an argument, in turn, is given as the support for the justification of the warrant, in cases where the warrant of argument is challenged. The backing of an argument is therefore generally regarded as the warrant for the warrant of an argument [Ver05].

In the context of abstract argumentation theory, given an envisaged pool of arguments, we present the *admissibility backings* of an argument as minimal admissible sets that either accept or reject an argument. We borrow the term *backing*, only due to the role of backings, that is to protect the justification (warrant) of an argument against possible counter arguments. This is however different from the usual use of backings of arguments in the literature.

Definition 2.4.1. Let $S \in 2^{AR}, a \in AR$ be given in some argumentation framework AF, and, $\mathcal{G}, \mathcal{P}, \mathcal{N}$ respectively denote the class of grounded admissible sets, the admissible sets that accept a, and, the admissible sets that attack a. Then, S is said to be a *positive backing* or a *backing for the acceptance* of a if and only if S is either a minimal element of $\mathcal{P}$ or a minimal

element of $\mathcal{P} \cap \mathcal{G}$ or both. Respectively, $S \in \mathcal{N}$ is a *negative backing* or a *backing for the rejection* of a if and only if S is a minimal element of $\mathcal{N}$ or is a minimal element of $\mathcal{N} \cap \mathcal{G}$. A positive or negative backing that is a grounded admissible set is called a *grounded backing* . The sets of all positive and negative backings of a are respectively denoted by $\langle a \rangle^{+}$ and $\langle a \rangle^{-}$.

In the previous section we pointed out the importance of the grounded admissible sets. However, not every minimal grounded admissible set that accepts an argument is necessarily a minimal admissible that accepts that argument, as was illustrated in example 2.3.3. Hence, in the definition of the backings of an argument, we give the grounded backings a special consideration regarding the minimality condition.

In example 2.3.3, although S_2 is the minimal grounded admissible set that accepts argument e, S_2 is not the minimal admissible set that accepts e. The only minimal admissible set that accepts e is S_4.

From this point on, unless it is stated otherwise, we refer to the admissibility backings only as the *backings, positive backings*, or the *negative backings*, depending on the context, and, by a *minimal set*, we mean a minimal set with respect to set inclusion.

2.4.1 The minimality clause of the admissibility backings

Some properties of the admissibility backings are due to the relation between the backings themselves, and, some are the direct result of their definition of being a minimal admissible set of a sort. We begin by characterizing minimal admissible sets that make up the admissibility backings.

An admissible set that accepts an argument needs not to be strictly comprised of the arguments that take part in the admissibility of the argument. For the backings however, due to the minimality clause of their definition, they need to be comprised of only the relevant arguments. The following lemma casts a wide boundary over the arguments that are relevant to the admissibility

of an argument.

Lemma 2.4.2. *Let DF_a, AT_a denote the set of in/direct defenders and the set of in/direct attackers of an argument a in AF. If $S \in \langle a \rangle^+$ then $S \subseteq DF_a$, and, if $S \in \langle a \rangle^-$ then $S \subseteq AT_a$.*

Let S be some conflict-free set that accepts a. For S to be a minimal of such sets, every argument in S must participate in the defense of a, i.e., there must be an onto mapping from $\bar{a}$ to S. We however need one additional condition. That is, every argument $c \in S$ must be necessary in S for the defense of a. Hence, we define the notion of a *critical defender*.

Definition 2.4.3. An argument c is said to be a *critical defender* of a in S if and only if c is the only argument in S that defends a against some $b \in \bar{a}$.

Lemma 2.4.4. $S \in \langle a \rangle^+$ *only if for every $c \in S$ there is some $d \in S \cup \{a\}$ such that c is a critical defender of d in S.*

Next, we like to see whether the backings of arguments can uniquely identify an argumentation framework. We know that since the characteristic function is built upon the notion of defense, it does not capture all the attack relations. Hence neither, the characteristic function, nor, the the set of all admissible sets of an argumentation framework can uniquely identify argumentation frameworks, as shown in the example 2.4.5. The backings, however, not only store the information about the acceptance of arguments, they also hold the information about their rejection. Despite this, due to the minimality clause in the definition of backings, there is still some information loss. As a result, the backings do not capture all the information on the attack relations, as shown in the second part of example 2.4.5. [2]

Example 2.4.5. The following argumentation frameworks AF_3, AF_4 share the same characteristic function $\theta_{AF_3} = \theta_{AF_4} = \{a, b\}$, as well as the set of all admissible sets $\mathcal{A}_{AF_3} = \mathcal{A}_{AF_4} = \{\varnothing, \{a\}, \{b\}, \{a, b\}\}$. On the other hand, the admissibility backings of their arguments are dis-

---

[2] A proper discussion of the unique identification of an argumentation framework involves the issue of the *equivalence relation* among argumentation frameworks, the types of such equivalence relations, and, whether or not they should be defined over the set of their sub-argumentation frameworks. Here, we however wanted to address the main question of whether the backings hold sufficient information to uniquely reconstruct an argumentation framework.

tinct where $\langle c \rangle^- = \{\{a\}\}$, $\langle d \rangle^- = \{\{b\}\}$ in AF_3, while $\langle c \rangle^- = \{\{a\}, \{b\}\}$, $\langle d \rangle^- = \{\{a\}, \{b\}\}$ in AF_4.

$$AF_3 \qquad\qquad AF_4$$
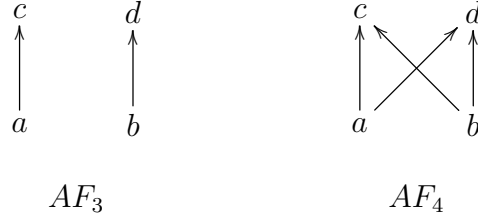
The following figure shows that although the two argumentation frameworks AF_5, AF_6 below are distinct, their respective arguments share the same backings. That is to say, there is some one-to-one and onto function ψ from the AR in AF_5 to the AR in AF_6, such that, for all $x \in AR$ in AF_5, we have $\langle x \rangle^+ = \langle \psi(x) \rangle^+$, and, $\langle x \rangle^- = \langle \psi(x) \rangle^-$. However, the set of attack relation, ATT, of AF_5 is different from the ATT of AF_6.

$$AF_5 \qquad\qquad AF_6$$
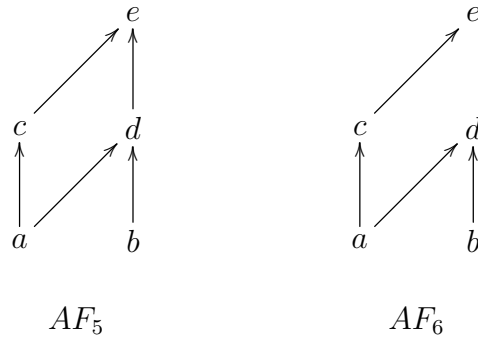
Hence, we remark that the admissibility backings do not always uniquely identify an argumentation framework.

2.4.2 Admissibility backings and the core results of Dung95

If the admissibility backings are to be considered as a viable construct within the abstract argumentation frameworks, it is necessary to show that the findings in [Dun95b] on the admissibility of arguments can be recast in terms of the admissibility backings. The results are presented as theorem 2.4.6, below. Theorem 2.4.6, however, does not reflect the intricacies of the minimality condition of the admissibility backings.

Theorem 2.4.6. *Let a be an argument in an argumentation framework AF.*

1. *If AF has a stable extension then $\langle a \rangle^+ \neq \varnothing$ or $\langle a \rangle^- \neq \varnothing$.*

2. *If a is a ground argument then $\langle a \rangle^+ = \{\varnothing\}$, $\langle a \rangle^- = \varnothing$.*

3. *If $S \in \langle a \rangle^-$ then $a \notin S$.*

4. *IF $S \in \langle a \rangle^+$ then for every $b \in \overline{a}$, there is some $T \in \langle b \rangle^-$ such that $T \subseteq S$.*

5. *The following statements are equivalent.*

    (a) *$\langle a \rangle^- = \varnothing$.*

    (b) *If $b \in \overline{a}$ then $\langle b \rangle^+ = \varnothing$.*

    (c) *The argument a belongs to all non-empty preferred extensions of AF.*

6. *The following statements are equivalent.*

    (a) *$\langle a \rangle^+ = \varnothing$.*

    (b) *Either there is some $b \in \overline{a}$, $\langle b \rangle^- = \varnothing$ or there is no conflict free set S in AF such that if $b \in \overline{a}$ then for some $T \in \langle b \rangle^-$, $T \subseteq S$.*

    (c) *The argument a belongs to no preferred extension.*

The results of theorem 2.4.6 are in large straightforward. For instance, 2.4.6.1 states that the existence of a stable extension ensures that every argument has some backings. The result 2.4.6.2 says that all the ground arguments have the same positive and negative backing. The result 2.4.6.3 states that, since backings are conflict free sets, we expect an argument not to take part in any of its negative backings. On the other hand, an argument that takes part in its own defense, may belong to some of its positive backings.

As a note in passing, an argument can simultaneously belong to some positive and some negative backing of another argument. In the following argumentation framework, AF_7, the positive and

negative backings of argument b are $\langle b \rangle^+ = \{\{a, b\}\}$, $\langle b \rangle^- = \{\{a, e\}\}$. Argument a is a controversial argument for b. It can be seen that a belongs to both the positive and the negative backings of b.

$$AF_7$$
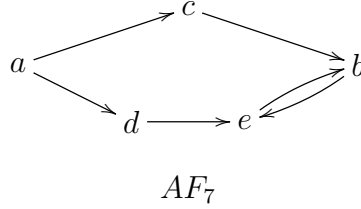
Due to the nature of the acceptance relation with respect to the attack relation, we expect a close relation between the backings of an argument and the backings of its attackers. This relation is partly presented by the results in theorems 2.4.6.5 and 2.4.6.6. The rest of this section is mostly dedicated to the formulation, derivation and the implications of this relation. To capture this relation we first define the operations $+, \circ$.

2.4.3 The operations $\dotplus$ and $\circ$

The claim in 2.4.6.4, states that if b is an attacker of a then since any positive backing S of a must attack b, it then must contain some negative backing T of b. To extend this result to all $b \in \bar{a}$, for every $b \in \bar{a}$ we must expect to find some $S_b \in \langle b \rangle^-$ such that $\bigcup_{b \in \bar{a}} S_b \subseteq S$. Moreover, since $S \in \langle a \rangle^+$ due to the minimality condition, we also have $S \subseteq \bigcup_{b \in \bar{a}} S_b$. This leads us to expect the positive backings for an argument to be somewhat a cross union of the negative backings of the attackers of that argument. That is if $\bar{a} = \{b, c\}$, $\langle b \rangle^- = \{S_1, S_2\}$, $\langle c \rangle^- = \{T_1, T_2\}$, and, $\dot{\circ}$ is the cross union operation, then $\langle a \rangle^+ = \langle b \rangle^- \dot{\circ} \langle c \rangle^- = \{S_1 \cup T_1, S_1 \cup T_2, S_2 \cup T_1, S_2 \cup T_2\}$.

Definition 2.4.7. Given an argumentation framework $AF = \langle AR, ATT \rangle$, the binary operations $\dotplus, \dot{\circ}$ over $2^{2^{AR}}$ are defined as follows.

- $A \in \mathcal{A} \dotplus \mathcal{B}$ if and only if A is a conflict free element of $\mathcal{A} \cup \mathcal{B}$.

- $A \in \mathcal{A} \dot{\circ} \mathcal{B}$ if and only if A is a conflict free element of $\{A \cup B \mid A \in \mathcal{A},\, B \in \mathcal{B}\}$.

The operation $\dot{\circ}$ as it stands, is still incomplete. We further need two additional adjustments. We need to enforce the conflict freeness, and, the minimality condition. The frameworks AF_8, AF_9, below highlight the need to amend the definiton of $\dot{\circ}$ as it is.

The argumentation framework AF_8 is a well discussed framework [BG04, Cam06b]. [3] Our argument of interest is the argument c. If we blindly apply the proposed operation $\dot{\circ}$ we will have $\langle c \rangle^+ = \langle b \rangle^- \dot{\circ} \langle a \rangle^- = \{\{a\}\} \dot{\circ} \{\{b\}\} = \{\{a, b\}\}$. But, set $\{a, b\}$ is not a conflict-free set. Hence, the operation $\dot{\circ}$ has to account for the safety of conflict free condition. The result of theorem 2.4.6.6b corresponds to the same issue.

$$AF_8 \qquad\qquad AF_9$$
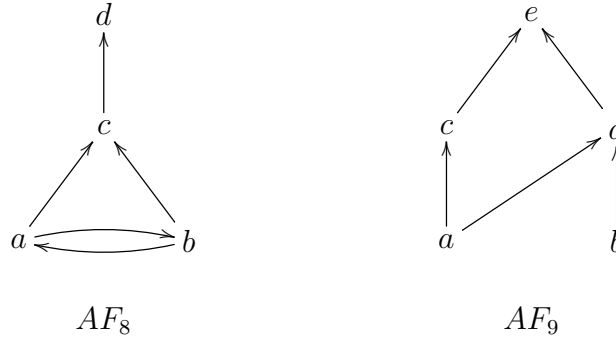
The minimality constraint also needs to be ensured at every application of $\dot{\circ}$. The reason is that $\min(\mathcal{A}) \cup \min(\mathcal{B})$ is not always equal to $\min(\mathcal{A} \cup \mathcal{B})$. The argumentation framework AF_9 illustrates why. Here, the argument of interest is e. In AF_9, the arguments c, d, share the negative backing $\{a\}$ where $\langle c \rangle^- = \{\{a\}\}$, $\langle d \rangle^- = \{\{a\}, \{b\}\}$. The application of $\dot{\circ}$ without the proper adjustments yields $\langle e \rangle^+ = \langle c \rangle^- \dot{\circ} \langle d \rangle^- = \{\{a\}\} \dot{\circ} \{\{a\}, \{b\}\} = \{\{a\}, \{a, b\}\}$. However, $\{a, b\}$ is not a minimal element of $\langle e \rangle^+$, and therefore is not an acceptable positive backing for e.

To ensure that the desired constraints, the minimality condition and being conflict free, are always met, we define the operations $+, \circ$ over $2^{2^{AR^\star}}$. The special symbol $\star$ denotes the admis-

---

[3] AF_8 is used to demonstrate that even in limited-controversial argumentation frameworks, the intersection of all preferred extensions does not necessarily coincide with the grounded extension. In this case, the argument d is not a grounded argument, while at the same time, it is accepted by all the preferred extensions. In general though, it seems not to be the case that we can say, we accept d beyond all reasonable doubts.

sible sets that are not grounded admissible sets.

Definition 2.4.8. For an argumentation framework AF, let $AF^\star = \langle AR^\star, ATT \rangle$, $AR^\star = AR \cup \{\star\}$, and operations $\dot{+}, \dot{\circ}$ be defined over $2^{2^{AR^\star}}$. The binary operations $+, \circ$ over $2^{2^{AR^\star}}$ are defined such that,

- $A \in \mathcal{A} + \mathcal{B}$ if and only if A is a minimal element of $\mathcal{A} \dot{+} \mathcal{B}$.

- $A \in \mathcal{A} \circ \mathcal{B}$ if and only if A is a minimal element of $\mathcal{A} \dot{\circ} \mathcal{B}$.

No backings of an argument both accepts and attacks an argument. Otherwise, it has to attack itself. The following theorem presents this finding in terms of operation $\circ$. The second part of the theorem uses the operation $+$ to present the result of theorem 2.4.6.1.

Theorem 2.4.9. *In a framework AF, for an argument a,*

1. *$\langle a \rangle^+ \dot{\circ} \langle a \rangle^- = \varnothing$.*

2. *If AF has a stable extension then $\langle a \rangle^+ \dot{+} \langle a \rangle^- \neq \varnothing$.*

Next, we define the pairs of operations $\dot{\sum}, \dot{\prod}$ and $\sum, \prod$ in the usual way. Let $\mathcal{A}_1, \mathcal{A}_2, \cdots$ be a sequence of the sets $\mathcal{A}_i \in 2^{2^{AR}}$ where $i \in \mathbb{N}$ (the set of natural numbers).

$$\dot{\sum_{i\in\mathbb{N}}} \mathcal{A}_i = ((\mathcal{A}_1 \dot{+} \mathcal{A}_2) \dot{+} \mathcal{A}_3) \dot{+} \cdots \qquad \sum_{i\in\mathbb{N}} \mathcal{A}_i = ((\mathcal{A}_1 + \mathcal{A}_2) + \mathcal{A}_3) + \cdots$$

$$\dot{\prod_{i\in\mathbb{N}}} \mathcal{A}_i = ((\mathcal{A}_1 \dot{\circ} \mathcal{A}_2) \dot{\circ} \mathcal{A}_3) \dot{\circ} \cdots \qquad \prod_{i\in\mathbb{N}} \mathcal{A}_i = ((\mathcal{A}_1 \circ \mathcal{A}_2) \circ \mathcal{A}_3) \circ \cdots$$

It can be seen that due to the commutativity and the associativity of $\dot{+}, \dot{\circ}$ and $+, \circ$, the operations $\dot{\sum}, \dot{\prod}$ and $\sum, \prod$ are both well defined. A full set of properties of $\dot{+}, \dot{\circ}$ and $+, \circ$ are given in appendix.

Moreover, following the reduction property $\mathcal{A} \circ \mathcal{B} = \mathcal{A} \circ (\mathcal{A} \circ \mathcal{B})$, it can be shown that for any two arbitrary indexing $\iota, \iota' \colon \mathbb{N} \longrightarrow \mathbf{A}$ of $\mathbf{A} \in 2^{2^{2^{AR}}}$ where $\mathbb{N}$ is the set of natural numbers,

the product of $\circ$ over $\mathbf{A}$ is the same.

$$\prod_{i \in \mathbf{N}} \mathcal{A}_{\iota(i)} = \prod_{i \in \mathbf{N}} \mathcal{A}_{\iota'(i)}.$$

Hence, we omit the indexing and instead write

$$\prod \mathbf{A} = \prod_{\mathcal{A} \in \mathbf{A}} \mathcal{A}.$$

2.5 The recursive property of admissibility backings

The relation between the backings of an argument and the backings of its attackers is what one would expect. If S is an admissible set that accepts $b \in \overline{a}$ then $S \cup \{b\}$ will be an admissible set that attacks a, and, if S is an admissible set that attacks all $b \in \overline{a}$ then S is an admissible set that accepts a.

To distinguish between the grounded and not grounded backings, we define a bijective mapping γ between $2^{2^{AR^\star}}$ and $2^{2^{AR}}$. The mapping γ maps every grounded backing into itself, and, every not-grounded backing S into $S \cup \{\star\}$.

Definition 2.5.1. In an argumentation framework $AF = \langle\, AR,\, ATT \,\rangle$, let $\mathbf{D} \subseteq 2^{2^{AR}}$ be the space of backings of arguments in AF and $AR^\star = AR \cup \{\star\}$. A bijective function γ from $\mathbf{D}$ to $\mathcal{D}^* \subseteq 2^{2^{AR^\star}}$ is a function where if $A \in \mathcal{A}$ and A is not a grounded backing then $A \cup \{\star\} \in \gamma(\mathcal{A})$.

The function γ assumes that we already know the backings of arguments including whether or not they exist for an argument. Its only purpose is then to allow the use of operations $+, \circ$ over the backings of arguments, as highlighted by the following observation.

Observation 2.5.2. *Let γ be a function in definition 2.5.1, then, $A \in \gamma^{-1}(\mathcal{A})$ if and only if* $A \in \{S - \{\star\} \mid S \in \mathcal{A}\}$.

Now we can present the main theorem of this chapter which states the relation between the backings of an argument and the backings of its attackers.

Theorem 2.5.3. *For an argument a and a set S of arguments,*

1. $\langle a \rangle^- = \gamma^{-1}(\mathcal{Z}^-)$ *where* $\mathcal{Z}^- = \displaystyle\sum_{b \in \bar{a}} (\{\{b\}\} \circ \gamma(\langle b \rangle^+))$.

2. $\langle a \rangle^+ = \gamma^{-1}(\mathcal{Z}^+)$ *where* $\mathcal{Z}^+ = \begin{cases} \displaystyle\prod_{b \in \bar{a}} \gamma(\langle b \rangle^-) & \text{if } \bar{a} \neq \varnothing, \\ \{\varnothing\} & \text{otherwise.} \end{cases}$

3. S *is not a grounded backing of* a *if* $S \cup \{\star\} \in \mathcal{Z}^+ \cup \mathcal{Z}^-$.

In formulating the backings of arguments, we inadvertently, also discussed the backings for a set of arguments. That is, if we define the negative backing for a set of arguments as the minimal admissible set that attacks every argument in the set, then, a positive backing for a is simply a negative backing for the set $B = \bar{a}$. Accordingly, we can define a positive backing for a set of arguments as the minimal admissible set that accepts all the arguments in the set. We however instead define operations of backings for a set of arguments with respect to the quantifiers *some* and *all*. The following theorem presents this result.

Theorem 2.5.4. *For a set A of arguments, let $\mathcal{H}, \mathcal{J}$ denote the class of admissible sets that accept all and accept some $a \in A$, and, $\mathcal{K}, \mathcal{L}$ denote the class of admissible sets that attack all and attack some $a \in A$.*

1. $S \in \gamma^{-1}(\displaystyle\prod_{a \in A} \gamma(\langle a \rangle^+))$ *if and only if S is a minimal element of $\mathcal{H}$ or S is a minimal element of $\mathcal{H} \cap \mathcal{G}$ that accepts every $a \in A$.*

2. $S \in \gamma^{-1}(\displaystyle\sum_{a \in A} \gamma(\langle a \rangle^+))$ *if and only if S is a minimal element of $\mathcal{J}$ or S is a minimal element of $\mathcal{J} \cap \mathcal{G}$ that accepts some $a \in A$.*

3. $S \in \gamma^{-1}(\displaystyle\prod_{a \in A} \gamma(\langle a \rangle^-))$ *if and only if S is a minimal element of $\mathcal{K}$ or S is a minimal element of $\mathcal{K} \cap \mathcal{G}$ that attacks every $a \in A$.*

4. $S \in \gamma^{-1}(\displaystyle\sum_{a \in A} \gamma(\langle a \rangle^-))$ *if and only if S is a minimal element of $\mathcal{L}$ or S is a minimal*

56

element of $\mathcal{L} \cap \mathcal{G}$ that attacks some $a \in A$.

Theorem 2.5.3 gives us two equations. The question is then whether or not we can derive the backings of these arguments by solving these equations. For instance, we want to find $\langle\, e\,\rangle^+$ in framework AF_{10} below. We backtrack through the attack relations, and, write the corresponding the equations. The operations $\oplus, \otimes$ stand for $+(\gamma)$ and $\circ(\gamma)$.

$$\langle\, a\,\rangle^+ = \langle\, b\,\rangle^+ = \langle\, d\,\rangle^+ = \{\varnothing\}$$

$$\langle\, x\,\rangle^- = \gamma^{-1}((\{\{a\}\} \otimes \langle\, a\,\rangle^+) + (\{\{b\}\} \oplus \langle\, b\,\rangle^+)) = \gamma^{-1}(\{\{a\}, \{b\}\}) = \{\{a\}, \{b\}\}$$

$$\langle\, c\,\rangle^+ = \langle\, x\,\rangle^- = \{\{a\}, \{b\}\}$$

$$\langle\, y\,\rangle^- = \gamma^{-1}((\{\{c\}\} \otimes \langle\, c\,\rangle^+) + (\{\{d\}\} \otimes \langle\, d\,\rangle^+)) = \gamma^{-1}(\{\{a, c\}, \{b, c\}\} + \{\{d\}\})$$

$$= \{\{a, c\}, \{b, c\}, \{d\}\}$$

$$\langle\, e\,\rangle^+ = \langle\, y\,\rangle^- = \{\{a, c\}, \{b, c\}, \{d\}\}.$$

The substitution of one result for another yields $\langle\, e\,\rangle^+ = \{\{a, c\}, \{b, c\}, \{d\}\}$.

$$AF_{10} \qquad\qquad\qquad AF_{11}$$
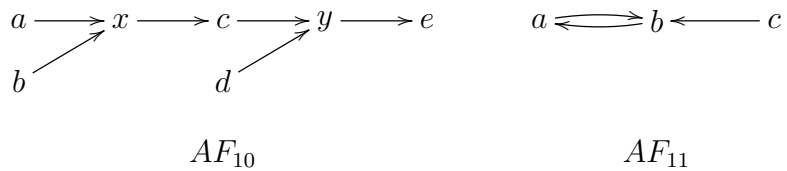
Although solving of the above equations is straightforward. This is not always the case, as difficulties may arise in the case of attack cycles. For instance, in AF_{11} above, if we write the

backing equations,

$$\langle\, a\,\rangle^{+} = \langle\, b\,\rangle^{-}$$

$$\langle\, b\,\rangle^{-} = \gamma^{-1}((\{\{a\}\} \otimes (\langle\, a\,\rangle^{+}) + (\{\{c\}\} \otimes \langle\, c\,\rangle^{+}))$$

$$\langle\, c\,\rangle^{+} = \{\varnothing\}$$

we will have,

$$\langle\, a\,\rangle^{+} = \gamma^{-1}((\{\{a\}\} \otimes \langle\, a\,\rangle^{+}) + \{\{c\}\}).$$

Hence, in case of the attack cycles it is common to face equations of the form,

$$\langle\, a\,\rangle^{+} = \gamma^{-1}((\mathcal{B} \otimes \langle\, a\,\rangle^{+}) + \mathcal{C}), \tag{2.1}$$

where solving them is not a straightforward matter. [4]

To overcome this problem, we use the result of theorem 2.5.3, and, define the recursive *backing function* $ß(d, T, j)$. The function $ß(d, T, j)$ is basically a different formulation of relation in theorem 2.5.3. It uses only the bare bone information which is the attack relations. The function $ß(d, T, j)$ itself is a mapping between the domains $\mathbb{D} = AR \times 2^{AR} \times \{0, 1\}$ and $2^{2^{AR^*}}$. The element d is the argument for which the backing is derived. T keeps a track of all the arguments visited. Hence, T ensures that the recursion path is finite, and, the function halts. The element j ensures that the crosswise relation between the backings of an argument and those of its attackers are followed correctly. The values $j = 0, j = 1$ respectively denote whether the function corresponds to the negative or the positive backings of argument d.

Definition 2.5.5. A backing function $ß(d, T, j)$ for a framework $AF = \langle\, AR,\, ATT\,\rangle$ is a

---

[4] The equation $\mathcal{A} = (\mathcal{A} \circ \mathcal{B}) + \mathcal{C}$ is in the same form as the equation (2.5.3). If in the right hand side of the equation, we substitute $\mathcal{A}$ by $\mathcal{B} \circ \mathcal{C}$, and the apply properties of $+, \circ$ given in appendix 5.2.2, we will get $(\mathcal{B} \circ (\mathcal{B}+\mathcal{C})) \circ \mathcal{C} = (\mathcal{B} \circ \mathcal{B}) + (\mathcal{B} \circ \mathcal{C}) + \mathcal{C} = \mathcal{B} + \mathcal{C}$. Hence, the equation (2.5.3) has a solution $\langle\, a\,\rangle^{+} = \{\{a\}\} + \{\{c\}\} = \{\{a, c\}\}$. However, we still need to show that this is the only solution.

We did not fully investigate how to find the solutions to multiple equations of the form of equation (2.5.3). We instead decided to sidestep the issue by defining the *recursive backing function*.

function from domain $\mathbb{D} = AR \times 2^{AR} \times \{0, 1\}$ to $2^{2^{AR^*}}$ where $AR^* = AR \cup \{\star\}$.

$$\ss(d, T, j) = \begin{cases} \begin{cases} \displaystyle\sum_{b \in \overline{d}} (\ss(b, T \cup \{d\}, 1) \circ \{\{b\}\}) & \text{if } \overline{d} \neq \varnothing, \ d \notin T, \ j = 0 \\[2em] \displaystyle\prod_{b \in \overline{d}} \ss(b, T \cup \{d\}, 0) & \text{if } \overline{d} \neq \varnothing, \ d \notin T, \ j = 1 \end{cases} \\[2em] \{\{\star\}\} & \text{if } \overline{d} \neq \varnothing, \ d \in T \\[2em] \begin{cases} \varnothing & \text{if } \overline{d} = \varnothing, \ j = 0 \\[2em] \{\varnothing\} & \text{if } \overline{d} = \varnothing, \ j = 1 \end{cases} \end{cases}$$

In definition of $\ss(d, T, j)$, the first pair of cases directly refer to the relation in theorem 2.5.3, whereas the bottom three cases are in regard to the special conditions. The last pair of cases are in regard to the ground arguments d where $\overline{d} = \varnothing$. The other three cases regard to the not-ground arguments. The value of $\ss(d, T, j)$ for ground arguments is a constant where $\ss(d, T, 0)$ for $\langle d \rangle^-$ and $\ss(d, T, 1)$ for $\langle d \rangle^+$ are respectively $\varnothing, \{\varnothing\}$. The second case identifies whether the recursion is traversed through an attack cycle. This is checked by the condition $d \notin T$ where T holds the visited arguments. If so, then the value of $\ss(d, T, 1)$ includes the special symbol $\star$ marking that the identified set is not a grounded set.

The set T is a free variable in $\ss(d, T, j)$. Its role is to trace the recursion steps. To derive the backings of an argument a, we therefore need to initialize the arguments of the function $\ss(d, T, j)$ to values $d = a, T = \varnothing$. The operations $\sum, \prod$ guarantee that the value of $\ss(a, \varnothing, j)$ is a set of sets where each member set is a minimal conflict free set that accepts or attacks a. This makes the result of function $\ss(a, \varnothing, j)$, the backings of argument a, as is shown by theorem 2.5.6 below.

Theorem 2.5.6. *For a finite argumentation framework AF, a backing function $\beta(d, S, j)$ always*

halts. Moreover, for an argument a,

1. $\langle a \rangle^+ = \gamma^{-1}(\beta(a, \varnothing, 1))$, *and,* $\langle a \rangle^- = \gamma^{-1}(\beta(a, \varnothing, 0))$.

2. S *is not a grounded backing of* a *if* $S \cup \{\star\} \in \beta(a, \varnothing, 1) \cup \beta(a, \varnothing, 0)$.

The following example illustrates both the working of function $\beta(a, \varnothing, j)$ and the approach taken for the proof of theorem 2.5.6.

Example 2.5.7. The following diagram represents argumentation framework $AF_{12} = \langle AR, ATT \rangle$. The goal is to test the findings of theorem 2.5.6 for argument a.

$$AF_{12}$$
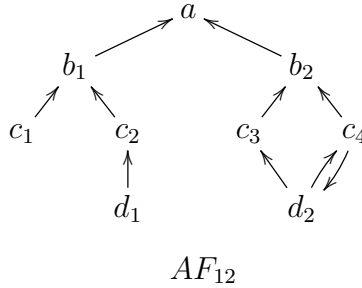
In the following, let $\vec{x} \in AR \times 2^{AR} \times \{0, 1\}$. For any $\vec{x} = (x, T, j)$, then the values of x, T are specific to the particular $\vec{x}$, while the value of j is free. We denote the values of $\beta(\vec{x})$ for $j = 1$ and $j = 0$, each by $\beta^1(\vec{x})$ and $\beta^0(\vec{x})$.

Let $\vec{a} = (a, \varnothing, j)$, we compute $\beta^1(\vec{a})$, $\beta^0(\vec{a})$ by directly following the definition 2.5.5. In doing so, we also draw the route of recursion in terms of $\vec{x}$, the arguments of function $\beta(\vec{x})$.

The recursion map of $\beta(\vec{a})$
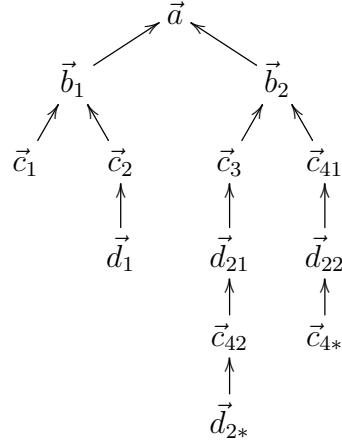
$$\vec{a} = (a, \varnothing, j)$$
$$\vec{b_1} = (b_1, \{a\}, j)$$
$$\vec{b_2} = (b_2, \{a\}, j)$$
$$\vec{c_1} = (c_1, \{a, b_1\}, j)$$
$$\vec{c_2} = (c_2, \{a, b_1\}, j)$$
$$\vec{d_1} = (d_1, \{a, b_1, c_2\}, j)$$
$$\vec{c_3} = (c_3, \{a, b_2\}, j)$$
$$\vec{c_{41}} = (c_4, \{a, b_2\}, j)$$
$$\vec{d_{21}} = (d_2, \{a, b_2, c_3\}, j)$$
$$\vec{d_{22}} = (d_2, \{a, b_2, c_4\}, j)$$
$$\vec{c_{42}} = (c_4, \{a, b_2, c_3, d_2\}, j)$$
$$\vec{c_{4*}} = (c_4, \{a, b_2, c_4, d_2\}, j)$$
$$\vec{d_{2*}} = (d_2, \{a, b_2, c_3, d_2, c_4\}, j)$$

The corresponding values $\beta(\vec{x})$ for the nodes $\vec{x}$ of the recursion map are as follows.

$$\beta^1(\vec{d_1}) = \beta^1(\vec{c_1}) = \{\varnothing\}.$$

$$\beta^0(\vec{d_1}) = \beta^0(\vec{c_1}) = \varnothing.$$

$$\beta^1(\vec{c_2}) = \beta^0(\vec{d_1}) = \varnothing.$$

$$\beta^0(\vec{c_2}) = \{\{d_1\}\} \circ \beta^1(\vec{d_1}) = \{\{d_1\}\} \circ \{\varnothing\} = \{\{d_1\}\}.$$

$$\beta^1(\vec{b_1}) = \beta^0(\vec{c_1}) \circ \beta^0(\vec{c_2}) = \varnothing \circ \beta^0(\vec{c_2}) = \varnothing.$$

$$\beta^0(\vec{b_1}) = (\{\{c_1\}\} \circ \beta^1(\vec{c_1})) + (\{\{c_2\}\} \circ \beta^1(\vec{c_2})) = (\{\{c_1\}\} \circ \{\varnothing\}) + (\{\{c_2\}\} \circ \varnothing)$$

$$= \{\{c_1\}\} + \varnothing = \{\{c_1\}\}.$$

$$\beta^1(\vec{c_{4*}}) = \beta^0(\vec{c_{4*}}) = \beta^1(\vec{d_{2*}}) = \beta^0(\vec{d_{2*}}) = \{\{\star\}\}.$$

$$\beta^1(\vec{c_{42}}) = \beta^0(\vec{d_{2*}}) = \{\{\star\}\}.$$

$$\beta^0(\vec{c_{42}}) = \{\{d_2\}\} \circ \beta^1(\vec{d_{2*}}) = \{\{d_2\}\} \circ \{\{\star\}\} = \{\{d_2, \star\}\}.$$

$$\beta^1(\vec{d_{22}}) = \beta^(\vec{d_{4*}}) = \{\{\star\}\}.$$

$$\beta^0(\vec{d_{22}}) = \{\{c_4\}\} \circ \beta^1(\vec{c_{4*}}) = \{\{c_4\}\} \circ \{\{\star\}\} = \{\{c_4, \star\}\}.$$

61

$$\text{ß}^1(\vec{d_{21}}) = \text{ß}^0(\vec{c_{42}}) = \{\{d_2, \star\}\}.$$

$$\text{ß}^0(\vec{d_{21}}) = \{\{c_4\}\} \circ \text{ß}^1(\vec{c_{42}}) = \{\{c_4\}\} \circ \{\{\star\}\} = \{\{c_4, \star\}\}.$$

$$\text{ß}^1(\vec{c_3}) = \text{ß}^0(\vec{d_{21}}) = \{\{c_4, \star\}\}.$$

$$\text{ß}^0(\vec{c_3}) = \{\{d_2\}\} \circ \text{ß}^1(\vec{d_{21}}) = \{\{d_2\}\} \circ \{\{d_2, \star\}\} = \{\{d_2, \star\}\}.$$

$$\text{ß}^1(\vec{c_{41}}) = \text{ß}^0(\vec{d_{22}}) = \{\{c_4, \star\}\}.$$

$$\text{ß}^0(\vec{c_{41}}) = \{\{d_2\}\} \circ \text{ß}^1(\vec{d_{22}}) = \{\{d_2\}\} \circ \{\{\star\}\} = \{\{d_2, \star\}\}.$$

$$\text{ß}^1(\vec{b_2}) = \text{ß}^0(\vec{c_3}) \circ \text{ß}^0(\vec{c_{41}}) = \{\{d_2, \star\}\} \circ \{\{d_2, \star\}\} = \{\{\}\} = \{\{d_2, \star\}\}.$$

$$\text{ß}^0(\vec{b_2}) = (\{\{c_3\}\} \circ \text{ß}^1(\vec{c_3})) + (\{\{c_4\}\} \circ \text{ß}^1(\vec{c_{41}})) = (\{\{c_3\}\} \circ \{\{c_4, \star\}\}) +$$

$$(\{\{c_4\}\} \circ \{\{c_4, \star\}\}) = \{\{c_3, c_4, \star\}\} + \{\{c_4, \star\}\} = \{\{c_4, \star\}\}.$$

Hence, we can draw $\text{ß}^1(\vec{a})$ and $\text{ß}^0(\vec{a})$

$$\text{ß}^1(\vec{a}) = \text{ß}^0(\vec{b_1}) \circ \text{ß}^0(\vec{b_2}) = \{\{c_1\}\} \circ \{\{c_4, \star\}\} = \{\{c_1, c_4, \star\}\}.$$

$$\text{ß}^0(\vec{a}) = (\{\{b_1\}\} \circ \text{ß}^1(\vec{b_1})) + (\{\{b_2\}\} \circ \text{ß}^1(\vec{b_2})) = (\{\{b_1\}\} \circ \varnothing) + (\{\{b_2\}\}$$

$$\circ \{\{d_2, \star\}\}) = \varnothing + \{\{b_2, d_2, \star\}\} = \{\{b_2, d_2, \star\}\}.$$

Accordingly, $\langle a \rangle^+$ and $\langle a \rangle^-$ are

$$\langle a \rangle^+ = \{S - \{\star\} \mid S \in \text{ß}^1(a)\} = \{\{c_1, c_4\}\},$$

$$\langle a \rangle^- = \{S - \{\star\} \mid S \in \text{ß}^0(a)\} = \{\{b_2, d_2\}\}.$$

The recursion in $\text{ß}(\vec{a})$ follows the attack relation. It is therefore not surprising to see that the recursion map of $\text{ß}(\vec{a})$ looks parallel to the graphical representation of AF_{10} itself. The difference between an argumentation framework and its corresponding map is that the map contains no cycle, and, it looks like a directed tree except may be for some leaf nodes. The reason for

this feature of the map lies in the tracking set T in (d, T, j). The tracking set T keep a record of the elements in the path of the recursion. On the other hand, the backing function ensures that the path does not contain any cycles. As a consequence, every set T is unique, except may be for some leaf nodes. ∎

2.6 The independency of admissibility backings

In theorem 2.2.5, we cast a wide net over the arguments that play no role in the acceptance or rejection of an argument. In this section, we make this net precise, narrowing it down to the member arguments of the backings of an argument. We start by recasting the finding in theorem 2.2.5 in terms of the backing functions.

Observation 2.6.1. *For an argument a and a set S of arguments, if $S \cap P_a = \varnothing$, where P_a is the set of all in/direct attackers and in/direct defenders of a, then $\beta(a, S, j) = \beta(a, \varnothing, j)$.*

The above observation states that in computing $\beta(a, \varnothing, j)$, if we start with any irrelevant set S and compute $\beta(a, S, j)$, we can then substitute the value of $\beta(a, S, j)$ for $\beta(a, \varnothing, j)$. Conversely, if we derive $\beta(a, \varnothing, j)$, then, we can have the value of $\beta(a, \varnothing, j)$ for any $\beta(a, S, j)$ where S is an irrelevant set with respect to the attacking or defending of the argument a. This is a useful property, as in deriving the $\beta(a, \varnothing, j)$, we can use the found values for $\beta(b, S, j)$ of b the in/direct attackers/defenders of a, for the original $\beta(b, \varnothing, j)$, and, vice versa.

For instance, in example 2.5.7, we can take $\beta^1(\vec{c}_{41}) = \{\{c_4, \star\}\}$ for $\langle c_4 \rangle^+ = \beta(c_4, \varnothing, 1) = \{\{c_4, \star\}\}$, and, $\beta^0(\vec{c}_{41}) = \{\{d_2, \star\}\}$ for $\langle c_4 \rangle^- = \beta(c_4, \varnothing, 0) = \{\{d_2, \star\}\}$.

We can improve on the results of theorem 2.2.5 and observation 2.6.1, by drawing a more precise boundary for the set S in theorem 2.6.1. It should be no surprise that this boundary can be identified by the members of backings of an argument. If an argument b, neither belongs to, nor attacks a backing of an argument a, then b plays no role in relation to the admissibility of a. The following theorem presents this finding.

Theorem 2.6.2. *For an argument a and sets S, T of arguments,*

1. *$\beta(a, S, j) = \beta(a, \varnothing, j)$ if and only if for all $T \in \langle a \rangle^+ \cup \langle a \rangle^-$, both $S \cap T = \varnothing$ and $S \not\rightsquigarrow T$.*

2. *If $\beta(a, S, j) = \beta(a, \varnothing, j)$ and $\beta(a, T, j) = \beta(a, \varnothing, j)$ then $\beta(a, S \cup T, j) = \beta(a, \varnothing, j)$.*

We can further improve on the result of theorem 2.6.2. We can show that every backing of a, accepts or rejects a independently of the other backings of a.

For every backing S of an argument a, there is a minimal sub- argumentation framework $AF_S \sqsubseteq AF$ that captures S. For AF_S to capture S, it should contain both a and $\overline{S}$, as well as, the attack relation between the members of S and $\overline{S}$, i.e, $AF_S^N = \langle AR_S, ATT_S \rangle \sqsubseteq AF$ where $AR_S = S \cup \overline{S}$, $ATT_S = (AR_S \times AR_S) \cap ATT$.

From the definition of backings, we observe that every backing of a, accepts or rejects a independently of the other backings of a. One way to present this, is to say that the status of S as a backing for a, is preserved, in any normal sub-argumentation framework of AF that contains AF_S. That is, for any $S \in \langle a \rangle^+$ and any $AF' \sqsupset AF_S$, S is a minimal admissible set in AF' that accepts a, irrespective of whether or not other backings of a are present in AF'. Moreover, AF_S is the minimal of such normal sub-argumentation frameworks for S.

Conversely, let $\mathcal{AF}$ be the class of argumentation frameworks AF_T such that some admissible set T accepts or rejects a, and that, T remains admissible in all AF_i for which $AF_T \sqsubseteq AF_i$. Then, all minimal elements AF_* of such sub-argumentation frameworks $\mathcal{AF}$ each coincides with with some backing S of a.

This correspondence between the backings of arguments and the minimal sub-argumentation frameworks in which the acceptance or rejection of a is preserved throughout the super frameworks is presented as the following theorem.

Theorem 2.6.3. *For an argument a in an argumentation framework AF,*

1. *For every backing S of a in AF there is some minimal $AF_S \sqsubseteq AF$ under $\sqsubseteq$, such that for all $AF_i \sqsupseteq AF_S$, S is a backing for a in AF_i. Moreover, AF_S contains no controversial arguments, and, has one and only one preferred extension $E_S = S$ that accepts (respectively rejects) a.*

2. *Conversely, if AF_S is a minimal sub framework of AF that for all $AF_i \sqsupseteq AF_S$, S is an admissible set that accepts or rejects a in AF_i, then S is a backing for a in AF.*

3. *For a backing S of a, its corresponding AF_S^* is given as, $AF_S^* = \langle AR_S^*, ATT_S^* \rangle \sqsubseteq^N AF$ where $AR_S^* = S \cup \overline{S}$. AF_S^* is the minimal normal sub-argumentation framework that contains all minimal sub-argumentation frameworks AF_S of result in (2.6.3.1).*

There are two remarks to be made. The first is that the minimal normal sub-framework AF_S^* may contain more than one backing of a. For instance, in framework AF_7, the minimal normal sub-framework for argument b is AF_7 itself. AF_7 however contains two backings of b, one positive and one negative backing of b.

The second remark is that the minimal sub-frameworks in theorem 2.6.3 do not fully identify the class $\mathcal{AF}_a^m$ of minimal sub-frameworks AF^m for which if a is accepted by some admissible set then a is accepted by some admissible set in all normal sub-frameworks AF' where $AF^m \sqsubseteq AF' \sqsubseteq^N AF$. The following example illustrates this remark.

Example 2.6.4. In the following framework AF_{13a}, argument a_1 has a positive backing $B = \{d, b_1\}$ where $\langle a \rangle^+ = \{\{d, b_1\}\}$. The minimal sub-framework that includes a_1 and B remains a backing of a_1 in all its sup-frameworks is AF_{13b}. This is the result of theorem 2.6.3. However, AF_{13b} is not the only minimal sub-framework for which in all its normal sup-frameworks a_1 has positive backing. AF_{13c} is such sub-framework. For AF_{13c}, in all normal sub-frameworks AF' where $AF_{13c} \sqsubseteq AF' \sqsubseteq^N AF$, argument a_1 has a positive backing.

$$AF_{13a} \qquad AF_{13b} \qquad AF_{13c}$$
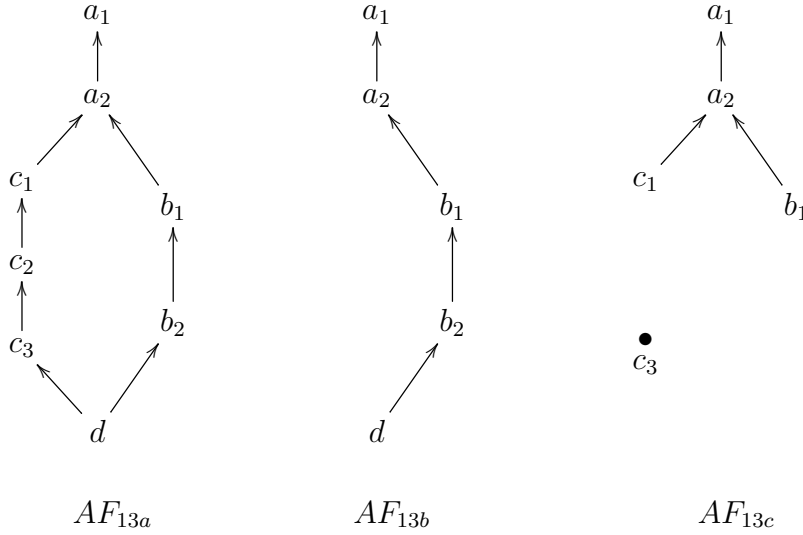
This concludes the results for how backings independently decide the admissibility of arguments. In the next section, we present the class of argumentation frameworks that is the proper fit for the admissibility backings.

2.7 The proper class of argumentation frameworks for admissibility backings

The admissibility backings, in the manner they are defined, are simply a special class of admissible sets that accept or reject an argument. In defining the admissibility backings, we do not demand any special constraints on the argumentation framework, and, the presented results stand for all argumentation frameworks including those with no stable extensions.

The motivation behind the admissibility backings, however, extends beyond their definition. The motivation is to provide the *full information* regarding the admissibility of arguments. In this section, we discuss how the notion of *full information*, is compromised in frameworks which are not coherent. In response, we then identify a class of coherent frameworks that stay true to this motivation.

In the background section, it is shown that a framework that is not coherent must have some

attack cycle of odd length. The problems that the attack cycles of odd length and the incoherent frameworks present, as well as, the treatment of those problems, are well discussed in the literature. The simplest form of attack cycles of odd length are self attacking arguments. In relation to the ramifications of attack cycles of odd length, Dung's framework does not distinguish between the self attacking arguments and the other forms of attack cycles of odd length. We view that the self attacking arguments belong to the special class of *absurd arguments*. Hence, we separate the self attacking arguments from the rest of the class of attack cycles of odd length. We say an argumentation framework is *rational* if it does not contain any self attacking argument.

2.7.1 The rational argumentation frameworks

One of the oldest paradoxes in logic is the liar's paradox [PV01]. Accepting the self attacking arguments, poses the same logical paradox as is presented in the liar's paradox. In regard to the acceptance of an argument, there are generally two mutually exclusive choices, either to accept the argument or to reject the argument.

In Dung's framework, if we choose to accept a self attacking argument then we are forced to reject all the arguments that it attacks, including the self attacking argument itself. This is clearly a logical impossibility. Hence, between the choice of accepting or rejecting a self attacking argument, we are always forced to reject the argument, independent of any particular framework. In other words, a self attacking argument can never belong to any admissible set. In this sense, we may claim that,

> in the abstract argumentation frameworks the closest notion to the *absurdity* is the self attacking arguments.

If we take a basic principle of rationality to be avoiding the *absurdity*, then, for an abstract argumentation framework to be considered rational, we must either exclude the self attacking arguments, or, any self attacking argument must be universally attacked by some metalogical

constant argument, e.g., some universal *empty-argument*. We adopt the former approach.

Definition 2.7.1. We say an argumentation framework AF is *rational* if and only if AF does not include any self attacking argument.

Observation 2.7.2. *If AF is rational then all $AF' \sqsubseteq AF$ are rational too.*

There is another major problem with the self attacking arguments. A self attacking argument that does not have a admissible attackers, is neither accepted or attacked by any admissible set. Let us call this the property $\mathfrak{U}$.

> An argument is said to have property $\mathfrak{U}$ if no admissible set either accepts or attacks that argument. Hence, an argument a has property $\mathfrak{U}$ if and only if $\langle a \rangle^+ = \langle a \rangle^- = \varnothing$.

The self attacking arguments that do not have an admissible attacker possess property $\mathfrak{U}$. The problem that self attacking arguments with property $\mathfrak{U}$ present, is that they initiate the propagation of the property $\mathfrak{U}$ onward (under the attack relation). That is, an argument which is effectively defended (or attacked) by a self attacking argument, will possess the property $\mathfrak{U}$ as well. Hence, we are faced with a situation where an argument that itself is not self attacking, is treated, with respect to all intensive purposes, like a self attacking argument. Obviously, this does not make sense.

In Dung's framework this problem can be tied to the nonexistence of stable extensions. For instance, the argumentation framework AF_{14}, below, does not have a stable extension. AF_{14} though has a grounded extension, namely $\{q\}$, which is also a complete and preferred extension. In AF_{14}, none of the arguments a, d, e is accepted or attacked by an admissible set, and, therefore none of these arguments have a positive or a negative backing.

$$AF_{14} \qquad\qquad AF_{15}$$
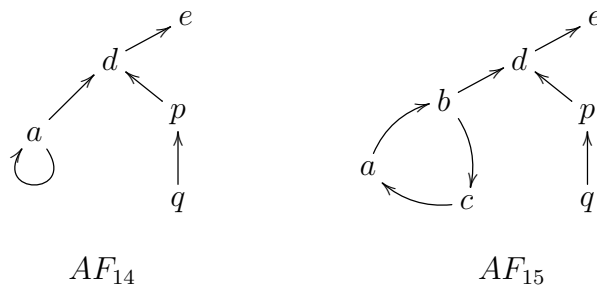
Under Dung's admissibility semantics, the arguments that indirectly attack themselves are also prone to have the property $\mathfrak{U}$. For instance, we can replace argument a in AF_{14} with an attack cycle of odd length a, b, c, and get framework AF_{15}. In AF_{15}, arguments a, d, e still have no positive or a negative backings. Hence, forcing an argumentation framework to be rational, does not address our concerns regarding the arguments with no admissibility backings.

Now, the problem with arguments with no admissibility backings is that we are not clear how to read them and treat them. The issue is that,[5]

> how an argument that is not in conflict with any admissible set, not to be accepted by some admissible set?
> Conversely, how an argument that is not accepted by any admissible, not to be in conflict with some admissible set ?

The treatment of this subject is beyond the scope of this thesis. We instead sidestep the issue and focus on the motivation behind the backings of arguments. The motivation is to localize the information regarding the admissibility of an argument, in form of the arguments responsible for the acceptance or rejection of that argument. Clearly, the statement $\langle a \rangle^+ = \langle a \rangle^- = \varnothing$, does not provide such information.

To account for this shortcoming, we have two options. For one, we can extend the definition of admissibility backings, and, allow for a new category of backings, namely the *in-inadmissible sets*, the set of arguments that are neither not conflict free, nor attacked by any admissible set. The admissibility backings presented here, however, do not allow for the *not conflict free* set of arguments. Hence, we go with the second option. We instead, focus our efforts on identifying the class of argumentation frameworks where all arguments have some backings, i.e., $\langle a \rangle^+ \neq \varnothing$ or $\langle a \rangle^- \neq \varnothing$. We call these classes of frameworks, the strongly, and, the normally stable argumentation frameworks.

---

[5] We can present this issue in terms of plausible models of the world. Let us assume that each admissible set correspond to a (minimal) plausible model of the world that we can construct from the evidence at hand. Hence, we are faced with a situation in which an argument is not in conflict with any plausible model of the world, and yet, it cannot be included in any.

2.7.2 The strongly stable and the normally stable argumentation frameworks

A general approach to secure that every argument has some backing, is to ensure that the framework has some stable extension, the theorem 2.4.6.1. The class of frameworks with some stable extension, does not however, generalize the class of frameworks for which every argument has some positive or some negative backings. For instance, in the framework AF_{16}, below, every argument is attacked or accepted by some admissible set, but, AF_{16}, itself, does not have a stable extension.

$$d_1 \nwarrow \quad \quad \quad \nearrow c_1$$

$$d_3 \leftarrow b \rightleftarrows a \rightarrow c_3$$

$$d_2 \nearrow \quad \quad \quad \nwarrow c_2$$

$$AF_{16}$$
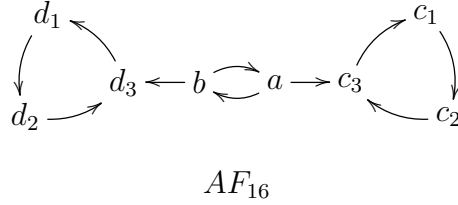
In addition, it is not just enough for a framework to have some stable extension, but, all the sub-argumentation frameworks of the framework should have some stable extension as well. The reason for this provision is directly related to the discussion regarding the need for sub-argumentation frameworks. For instance, an argumentation framework that has some stable extension, may not have any stable extension after the addition of some new arguments. To this end, we first define the *strongly stable* argumentation frameworks.

Definition 2.7.3. We say an argumentation framework AF is *stable* if and only if AF has some stable extension. We say AF is *strongly stable* if and only if all AF', $AF' \sqsubseteq AF$ are stable.

Observation 2.7.4. *Every sub-argumentation framework of a strongly stable argumentation framework is strongly stable too.*

We observe that the class of strongly stable argumentation frameworks is closed under the sub-argumentation framework relation. It then follows that the class of strongly stable argumentation frameworks form a partial order under the sub-argumentation framework relation. In

addition, it turns out that the class of strongly stable argumentation frameworks equates with the class of limited-controversial argumentation frameworks, as shown by theorem 2.7.5, below.

Theorem 2.7.5. *An argumentation framework AF is strongly stable if and only if AF is limited-controversial.*

The limited-controversial argumentation frameworks are coherent. The advantage of working with coherent frameworks is that, not only, every framework has some stable extension, but, every preferred extension is a stable extension as well. The disadvantage of confining ourselves to the limited-controversial frameworks is that they are too restrictive. An argumentation framework is limited-controversial if and only if it contains no attack cycle of odd length. The issue is that not all attack cycles of odd length are problematic or avoidable or the result of some rare occurrences.
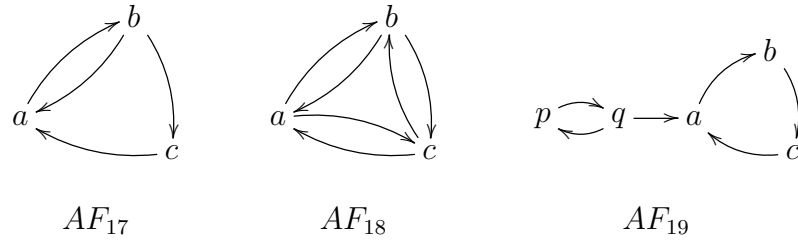
$$AF_{17} \qquad AF_{18} \qquad AF_{19}$$

For instance, all the three frameworks AF_{17}, AF_{18}, AF_{19}, below, have some stable extension. It is difficult to justify that the configuration in AF_{18} is some rare occurrence, or, the configuration in AF_{17} is easily avoidable. In short, in order to cover a greater range of realistic, yet well behaved, arguments interaction, we need to go beyond the strongly stable frameworks. We do this by relaxing the requirement for, *every sub-framework of a framework needs to be stable* to *every normal sub-framework of a framework to be stable*. When we consider admissibility of arguments in the real sense, we need to take account the full arguments interactions, given by the full set of attack relation. Consequently, we are only required to consider the stability of normal sub-argumentation frameworks.

71

Definition 2.7.6. We say an argumentation framework AF is *normally stable* if and only if all AF', $AF' \sqsubseteq^N AF$ are stable.

We observe that the class of normally stable frameworks is a superset of the class of strongly stable frameworks. The class of normally stable frameworks is also closed under the *normal* sub-argumentation framework relation, the observation 2.7.7, where they form a partial order under the normal sub-argumentation framework relation.

Observation 2.7.7. *In an argumentation framework AF ,*

1. *Every normally stable argumentation framework is rational.*

2. *Every strongly stable argumentation framework is normally stable too.*

3. *Every normal sub-argumentation framework of a normally stable argumentation framework is normally stable too.*

Following our example above, while none of the frameworks $AF_{17}, AF_{18}, AF_{19}$, is a strongly stable framework, the two intended frameworks AF_{17}, AF_{18} are normally stable frameworks. Hence, we use the frameworks AF_{17}, AF_{18} to characterize the normally stable frameworks. The distinctive feature of the frameworks AF_{17}, AF_{18}, is that, every attack cycle of odd length L, contains an attack cycle $L' \subset L$, of even length. It is easy to see that this property is preserved through all the normal sub-frameworks of a framework. The theorem 2.7.8, below, stipulates this characterization of the normally stable frameworks.

Theorem 2.7.8. *An argumentation framework AF is normally stable if and only if every attack cycle, L, of odd length, contains an attack cycle, L', of even length.*

One way to test the merits of normally stable frameworks, is by adding the new arguments and attack relations to the mix while preserving the integrity of the current framework, i.e., keeping the new framework, a normal framework. We can see that, in case of the frameworks AF_{17}, AF_{18}, if we introduce any new argument, the resulting framework will still be a stable

framework. On the other hand, in case of AF_{19}, if we add the new argument z where $z \hookrightarrow q$, the resulting framework is no longer a stable framework. This outcome can easily be explained by theorem 2.7.8. The addition of a new argument z that attacks any current argument, effectively, equates with removing the attacked argument from the framework. That is to say that we are dealing with a normal sub-framework of the current framework. By theorem 2.7.8, any normal sub-framework is still a stable framework. However, in the case of AF_{19}, the resulting sub-framework contains an attack cycle of odd length, and therefore, is not stable.

There is also a second matter of concern regarding framework AF_{19}. It is that not all preferred extensions in AF_{19} are stable extensions. As a result, we have a situation where $\langle b \rangle^+ \neq \varnothing$ and $\langle b \rangle^- = \varnothing$. Normally, when we are given, only the information $\langle b \rangle^+ \neq \varnothing$, $\langle b \rangle^- = \varnothing$, we expect that since no admissible set has any conflict with b, b should to accepted by all admissible sets. However, this is not the case. The information that b is not accepted by all admissible sets, is not readily carried by the information in $\langle b \rangle^+, \langle b \rangle^-$. If we wish for the admissibility backings to contain all the information that is deducible at the first glance, a framework needs to be coherent. As it happens, every normally stable argumentation framework is a coherent framework, as stated by the following theorem.

Theorem 2.7.9. *Every normally stable argumentation framework is coherent.*

Hence, we have come full a circle and defined a class of coherent frameworks that every normal sub-framework of it is coherent as well, namely, normally stable frameworks.

In the motivation chapter, section 1.2.4, we stated our reasons for why its is prudent to work with finite argumentation frameworks. Many of results in the literature, too, are presented in relation to the finite argumentation frameworks. In this regard, we define the *compact* argumentation frameworks. The compact frameworks maintain the important characteristic of finite frameworks while relaxing the constraint on finiteness of the framework.

2.7.3 Compact argumentation frameworks

In [Dun95b], an argumentation framework is said to be *finitary* if for every argument a, $\bar{a}$ is a finite set. However, this property does not ensure that every argument is accepted or rejected by some finite admissible set. For instance, the argumentation framework $AF = \langle AR, ATT \rangle$, where $AR = \{a_1, a_2, \cdots\}$, $ATT = \{(a_{i+1}, a_i) \mid i : \text{Natural number}\}$, is a coherent finitary argumentation framework, with two stables extensions. AF does not, however, have a finite admissible set. To make sure that every argument, whether accepted or rejected, is accepted or rejected by some finite admissible set, we define the *compact* argumentation frameworks. This definition of a compact argumentation framework is originally presented in [Boc02]. We present a slight variation of it here.

Definition 2.7.10. An argumentation framework $AF = \langle AR, ATT \rangle$ is said to be *compact* if and only if for all arguments $a \in AR$, the set of in/direct attackers and defenders of a, denoted by DF_a, AT_a are both finite, i.e., $DF_a \cup AT_a$ is a finite set.

Observation 2.7.11. *If AF is compact then all $AF' \sqsubseteq AF$ are compact too.*

It can be seen that following the theorem 2.2.5, any result regarding a finite argumentation equally holds for its corresponding infinite but compact argumentation framework.

In working with the admissibility backings of arguments, we effectively work with the normally stable frameworks that are compact.

Definition 2.7.12. We call an argumentation framework AF to be a $\mathcal{NSC}\text{-}AF$ framework if and only if it is both normally stable and compact.

This concludes our presentation of admissibility backings. There are a few studies that are directly related to our work here, and, many that are indirectly related. In the next section, we briefly discuss some of the related works.

2.8 Related research

The most direct reference to the admissibility backings is the *minimal admissible defense sets*, by Vreeswijk in [Vre06]. The motivation behind [Vre06], is to present an efficient algorithm for calculating the admissible sets. The minimal admissible defense sets are then the nucleus around which the algorithm is built. To test the computational efficiency of the algorithm, the algorithm is run against a number of average and worst case scenarios. The computational results are found to be promising.

The paper [Vre06] does not formally define minimal admissible defense sets. This can be attributed to the motivation behind [Vre06] where the role of minimal admissible defense sets is only to serve the algorithm. There is, however, a close relation between the algorithm in [Vre06] and the backing functions. The output of the algorithm is similar to the output of the backing function. The difference lies with the intention behind the algorithm that is to efficiently produce the admissible sets. Hence, there is no formal verification, or otherwise, of the minimality condition, that the presented admissible sets are indeed the minimal admissible defense sets.

In relation to the related research on sub-argumentation frameworks, there are many implicit, and, some explicit references. However, none of the works formally defines the sub-argumentation framework relation. The explicit references are mostly in regard to the merging, splitting and the dynamics of argumentation systems [CMDK$^+$07, Bau11]. The implicit references are mostly found in the proofs and definitions, in cases where only a subset of the whole set of arguments is considered [Dun95b, BG04, Cam06a]. The implicit references are in large in regard to the normal sub-argumentation frameworks.

Finally, there are a number works on how to treat the attack cycles of odd length, notably [BG04, Cam06b]. Within the studies on the abstract argumentation frameworks, the approaches on how to deal with the attack cycles of odd length, in general, fall within two mutually exclusive camps,

those that attempt to provide a semantics for the attack cycles of odd length, and, the other camp which includes those that consciously sidestep the issue. In this thesis, we consciously sidestep them, as discussed in the previous section.

2.9 Summary

In this chapter, we presented the backings of arguments as the minimal admissible sets that accept or reject an argument. In doing so, due to the importance of the grounded admissible sets, we decided to distinguish between the grounded and the not-grounded backings.

An important contribution of this chapter is the formulation of the relationship between the backings of an argument of those of its attackers, and, the recursive function that captures this relation. There are still many follow up queries on the backings of argument. For instance, the impact of an argument in relation to the acceptance of an argument. An important query of such, is how the backings of arguments propagate through the attack relations, or simply the *propagation of backings*.

For instance, if b is an in/direct defender of a, then, for any positive backing S of a, would we find some positive backing T of b such that $T \subseteq S$? The propagation of backings also provides a certain insight into the relations between arguments. The next chapter is therefore dedicated to the investigation the propagation of admissibility backings.

Chapter 3

The Role of one Argument in Admissibility of another and the Propagation of Admissibility Backings

3.1   Introduction

A central question in abstract argumentation frameworks is whether or not an argument is admissible. One aspect of this question is, what roles arguments play in regard to the admissibility of another. In this chapter, we employ the help of *admissibility backings* to identify and characterize a number of such roles. Since the *admissibility backings* localize the admissibility of arguments, they provide a suitable ground for investigating the functions that arguments may serve with respect to the admissibility of others.

The identified roles are termed *intercepted, critical, incompatible*, and, *redundant* arguments. Following the tradition in literature, these roles are presented in terms of relations among arguments.

Furthermore, we identify these roles in a manner that directly relates to how the admissibility backings propagate along the sequences of attack relations. Thus, in parallel, we address the propagation of the *admissibility backings*, as well.

The presented results can also be used for partitioning a framework into sub-frameworks with distinct features. Hence, as a show case and the final part of our contribution, we demonstrate how the *intercepts* can be used to partition a framework into *independent* sub-frameworks. This would also serve as a prelude to the future work on the applications of the *admissibility backings* in partitioning and merging of argumentation frameworks.

A number of relationships among arguments are already defined that identify whether or not an arguments can have a role in the admissibility of another. Naturally, all these relations are defined based on the attack relation. The classifications of current argument relations, e.g. attacker of an argument, do not however identify whether or not the argument does contribute to admissibility of the target argument.

In Dung's framework, the manner by which arguments are considered to be admissible governs a certain relation between the backings of an attacker of an argument and the backings of that argument. The theorem 2.5.3 shows this relation. By the propagation of admissibility backings of arguments, we mean how this relation propagates along the attack sequences. Consequently, the role that arguments play in admissibility of another is directly related to propagation of the backings of arguments.

The basic question that we are considering is,

Note 3.1.1.   if $R \in \langle\, b\,\rangle^+ \cup \langle\, b\,\rangle^-$ and $a \in R$, then, given any $S \in \langle\, a\,\rangle^+ \cup \langle\, a\,\rangle^-$, is there some $T \in \langle\, b\,\rangle^+ \cup \langle\, b\,\rangle^-$ such that $S \subseteq T$?

In answering this question we can in turn identify certain relations among arguments that characterize the roles that arguments play in admissibility of other arguments. This is the topic of this chapter.

3.2 Active arguments and attack sequences

Following the discussion above, not every argument, or, even every attacker or defender of an argument is relevant to the admissibility of that argument. On the other hand, it can be said that,

Note 3.2.1. an argument that takes part in some backing of some argument plays a role in the admissibility or the dismissibility of that argument.

Hence, we can expand on the role of in/direct attackers and in/direct defenders of an argument by looking at whether or not they take part in some backing of that argument. We call these in/direct attackers or in/direct defenders of an argument the *active attackers* and the *active defenders* of that argument.

Definition 3.2.2. For arguments a, b, we say b is an *active defender* of a if and only if $b \in S$, for some $S \in \langle a \rangle^+$. We say b is an *active attacker* of a if and only if $b \in S$, for some $S \in \langle a \rangle^-$. We refer to b as an *active argument* for a if and only if b is an active defender or an active attacker for a.

The claim (3.2.1) above says that an active argument b of an argument a is relevant to the admissibility of argument a. If this is the case then b must be a member of some minimal sub-argumentation framework for which the admissibility of a is preserved under sup-framework relation. The theorem 2.6.3 on page 64 shows this conclusion to be correct.

The converse of claim (3.2.1) is, however, not necessarily correct. Not every argument that does not appear in some backing of an argument can be deemed irrelevant in regard to the admissibility of the argument. The converse is not correct in the case of controversial arguments, as shown by the following example.

Example 3.2.3. In framework AF_1 below, argument d is attacked by some admissible set and belongs to no admissible set, i.e., $\langle d \rangle^+ = \varnothing$, $\langle d \rangle^- = \{\{a, b\}\} \neq \varnothing$. The minimal sub-framework generated for $\{a, b\}$, is shown by AF_{1a}. However, AF_{1a} is not the only minimal

79

sub-framework for which the admissibility status of d is preserved under sup-framework relation. AF_{1b} is another of such minimal sub-framework. Hence, we have another minimal set of argument that is important for the admissibility of d which is $\{b, c, y_1\}$, making the arguments c, y_1 relevant for the *admissibility status* of d.

$$a \overset{x \longrightarrow b}{\underset{y_1 \to y_2 \to c}{}} d \qquad a \overset{x \longrightarrow b}{} d \qquad \overset{b}{\underset{y_1 \to y_2 \to c}{}} d$$

$$AF_1 \qquad\qquad\qquad AF_{1a} \qquad\qquad\qquad AF_{1b}$$
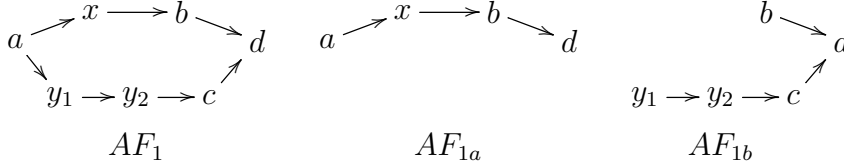
One may then suggest that the notion of active arguments for an argument to be extended to the attackers of active defenders or the active attackers of an argument, as well. The reason being, since the defeat of attackers of active arguments is a necessary condition for admissibility of the active arguments, they too must be directly relevant for admissibility of the original argument. However, we opt to keep the definitions simple and leave them as they are. We leave the handling of the controversial arguments as well as the dependency relation between arguments as future work of this Thesis. ∎

From the definition of active arguments, it is easy to see that there is a direct relation between the transitivity of the *active argument relation*, presented by the question,

Note 3.2.4. if a is an active argument for b and b is an active argument for c, then, is a an active argument for c?

and the propagation of backings, presented by the question (3.1.1) on page 78. In fact, it can be shown that the questions (3.2.4) and (3.1.1) are related. The answer to both questions is, generally, 'yes', unless, there are mitigating conditions. In sections 3.5 and 3.6, we classify and characterize the special conditions under which transitivity of the active argument relation fails. Although, the answer to question (3.1.1) is not always 'yes', the answer to the question,

Note 3.2.5. if a is an active argument for b, then, are there some admissibility backings S for a and T for b such that $S \subseteq T$?

is always 'yes', as presented by the following observation.

Observation 3.2.6. *For arguments a, b, c, if $a \in T$ for some $T \in \langle b \rangle^+$ or $T \in \langle b \rangle^-$, then, there is some $S \in \langle a \rangle^+$ such that $S \subseteq T$, and, if $c \in \overline{a}$, then, there is some $R \in \langle c \rangle^-$ such that $R \subseteq S \subseteq T$.*

The observation 3.2.6, states that once we know that an argument a is an active argument of an argument b, we then know that some positive backing of a, as well as, some negative backings of all its attackers are passed (with respect to set inclusion) to the backings of b.

The theorem 2.5.3 is a good point for studying the propagation of backings of arguments. Theorem 2.5.3 shows the basics of how the backings of an attacker of an argument propagate into the backings of the argument. Hence, we can say that a necessary condition for an argument a to be an active argument for some argument b is that a has to be an active argument for some attacker of b, as shown by the following observation. The converse is not necessarily true. There are cases where the transitivity of active argument relation fails.

Observation 3.2.7. *For two arguments a, b where $a \notin \overline{b}$, if a is not an active argument for any $c \in \overline{b}$ then a is not an active argument for b.*

Theorem 2.5.3 and observation 3.2.7, therefore, suggests that we can follow the chain by which the backings propagate with respect to the attack sequences. Hence, we define the *active attack sequences* as follows.

Definition 3.2.8. In an argumentation framework AF, an attack sequence $\pi = (a_0, a_1, \cdots)$ of a finite or countable length is said to be,

- *positively active* if and only if it is non empty (i.e, has a length greater than zero) and every a_i where i is an even number, is accepted by some admissible set in AF.

- *negatively active* if and only if it is non empty (i.e, has a length greater than zero) and every a_i where i is an odd number, is accepted by some admissible set in AF.

- *active* if and only if it is positively active or negatively active.

- *intercepted* if and only if it is not active.

- *partially intercepted* if and only if it is an active attack sequence that has a non trivial subsequence which is both positively and negatively active.

A sequence can be either or neither or both positively and negatively active. In definition 3.2.8, we differentiate between the positively and negatively active sequences. A positively active sequence can have a sub-sequence that is both positively and negatively active.

For instance, in framework AF_2, the attack sequence $\pi_1 = (b, a, e)$ is intercepted where its subsequence $\pi_2 = (a, e)$ is negatively active. In framework AF_{2a}, the attack sequence $\pi_3 = (c, f, d, b, a, e)$ is partially intercepted where it is negatively active and its subsequence $\pi_4 = (c, f, d, b)$ is both positively and negatively active. In framework AF_{2b}, the attack sequence $\pi_5 = (b, a, c, a)$, too, is partially intercepted where it is negatively active and its subsequence $\pi_6 = (a, c, a)$ is both positively and negatively active. In framework AF_{2c}, all possible non empty attack sequences are not intercepted where all the non empty attack sequences are both positively and negatively active sequences.
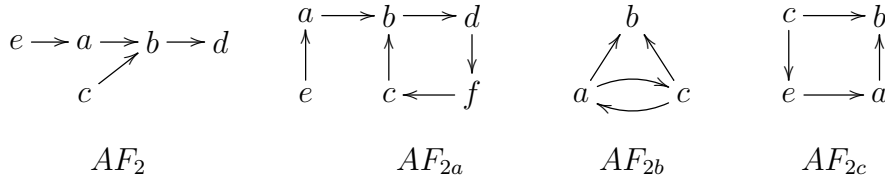
$$AF_2 \qquad AF_{2a} \qquad AF_{2b} \qquad AF_{2c}$$

In addition, in all frameworks AF_2, AF_{2a}, AF_{2b}, all the instances of intercepts and partial intercepts of sequences happen in regard to the a change in the admissibility status of the consecutive arguments a, b. The admissibility status ϵ of a, b in frameworks AF_2, AF_{2a}, AF_{2b} are in turn $\epsilon(a) = 0, \epsilon(b) = 0$, and, $\epsilon(a) = 0, \epsilon(b) = 1/2$, and, $\epsilon(a) = 1/2, \epsilon(b) = 0$. Hence, we use the change in admissibility status of arguments along an attack sequence to characterize the active and the intercepted sequences, as shown by the following lemma. [1]

---

[1] The characterization of partially intercepted attack sequences is more complicated. We classify the partially intercepted sequences in the following sections. Since our main interest is with the active and the intercepted

Lemma 3.2.9. *An attack sequence is not active if and only if there are some consecutive elements a_i, a_{i+1} on the sequence such that $\langle a_i \rangle^+ = \langle a_{i+1} \rangle^+ = \varnothing$.*

Many studies on the abstract argumentation frameworks either employ the graphical representation of frameworks, or, are based on the graph theoretical terms and methods. The attack sequences are directly parallel to the *paths* between the nodes (arguments) in a graph representation of a framework. Hence, the use of attack sequences allows for a straightforward translation of the results here into the graph theoretical terms.

Following observation 3.2.6, it easy to see that every active argument a for an argument b must be on some active attack sequence from b to a, as shown by the following lemma.

Lemma 3.2.10. *If b is an active argument for a, then, there is some active attack sequence from a to b, such that b is an active argument for all arguments in the sequence.*

From this point on, unless specified otherwise, by an *active sequence* we mean an active attack sequence. In the following sections, we will look more closely at the relation between the active arguments and the active attack sequences. First, we identify the points along an attack sequence where the sequence stops being an active attack sequence. In such situations, we say that the active sequence is *intercepted*. Hence, we first start with the *intercepts*.

3.3   Intercepts

In this section, we look at the lines where the propagation of backings fully stops. In such incidences, we say that propagation of backings is *intercepted*. For an argument a, an attacker of argument b, we normally expect that a backing of b to contain some backing of a. However, if the attack sequence $\pi = (b, a)$ is intercepted at b, then, this expectation shall not be met. Hence, we define the notion of *intercepts* with respect to the attack sequences.

Definition 3.3.1. In an argumentation framework AF, we say an argument a is *intercepted* for

---

sequences, we only characterized the active and the intercepted sequences.

an argument b if and only if all the attack sequence from b to a are intercepted or there are no attack sequences from b to a.

We expect the intercepted arguments to be irrelevant to the admissibility of that argument. If there is no attack sequence from b to a then a is vacuously irrelevant to the admissibility of b. Definition 3.3.1 adheres to this expectation and says that if there are no attack sequences from b to a then a is vacuously intercepted for b. Furthermore, following lemma 3.2.10 we see that if all attack sequences from b to a are intercepted, then a cannot be an active argument for b, as shown by the following observation.

Observation 3.3.2. *For two arguments a, b, if a is intercepted for b then a is not an active argument for b.*

If we follow the characterization of active sequences, the lemma 3.2.9, it is clear that the intercepts are related to the admissibility status of consecutive arguments a_i, a_{i+1} along the attack sequence. For an active attack sequence, if a_i is not admissible then a_{i+1} should be admissible. For an intercepted sequence, this however not the case. The following lemma uses this property to characterize the intercepts.

Lemma 3.3.3. *For an argument a where $a \in \bar{b}$, a is intercepted for b if and only if $\langle a \rangle^+ = \varnothing$ and there is some $S \subseteq \bar{b} - \{a\}$ such that $\prod_{c \in S} \gamma(\langle c \rangle^-) = \varnothing$.*

Following the above lemma, we see that for an intercept to occur we need some third argument, namely c, to stop propagation of the negative backings of a for the defense of b. In such situations, we say that *a is intercepted for b by c*. Moreover, there can be more than one argument c that intercepts a for b.

The propagation of intercepts happen both forward and backward under the attack relation. The following lemma presents this by breaking the propagation of intercepts in terms of the attack relation. The simple version of this lemma is such that if c is an attacker of b and a is intercepted for c, then, a is also intercepted for b, the forward propagation. Similarly, if a is an attacker of

c and c is intercepted for b then a is also intercepted for b, the backward propagation.

Lemma 3.3.4. *For arguments a, b where $a \notin \bar{b}$, let C be the set of all $c \in \bar{b}$ for which a is an in/direct attacker or an in/direct defender, and, D be the set of all $d \in \overset{+}{a}$ where d is an in/direct attacker or an in/direct defender of b.*

1. *If a is intercepted for all $c \in C$ then a is intercepted for b.*

2. *If all $d \in D$ are intercepted for b then a is intercepted for b.*

Lemma 3.3.3 presents the base case for occurrences of the intercepts and lemma 3.3.4 describes the propagation of the intercepts under the attack relation. We can therefore put the two lemmas together and characterize the occurrences of intercepts, as shown by the theorem below.

Theorem 3.3.5. *For arguments a, b, let D be the set of all arguments $d \in \overset{+}{a}$ where d is an in/direct attacker or an in/direct defender of b. Argument a is intercepted for b if and only if one of the followings hold.*

1. *If $a \in \bar{b}$ then $\langle a \rangle^- \neq \varnothing$ and there is some $S \subseteq \bar{b} - \{a\}$ such that $\langle S \rangle^- = \varnothing$.*

2. *If $a \notin \bar{b}$ then for any $d \in D$, either a is intercepted for d or d is intercepted for b or both.*

It is obvious that where the propagation of backings halts, the transitivity of the active argument relation fails as well. The notion of intercepts, therefore, applies to both the active argument relation and the active attack sequences. Hence, we can define the intercepts either in regard to the backings of arguments or with respect to the attack sequences. For the reasons explained before, we opt to define the intercepts as a relation between arguments using the attack sequences. The first stop in investigation of the role that arguments play with respect to the admissibility of others is the critical argument relation.

3.4 The critical arguments

For an argument a, arguments that do not belong to, or attacked by, any of the backings of a, play no significant role in admissibility of a. By the same token, an argument that appear in many of the backings of a, is expected to be very significant for the admissibility of a. An argument becomes crucial for a, if its inclusion or removal will result in a total change in the acceptability of a. We define the *critical* arguments as the arguments on which the admissibility of an argument almost totally depends.[2]

For instance, in argumentation framework, AF_3 below, if we remove either of arguments, b_{11} or b_{12}, argument a stays accepted by some admissible set in the resulting framework. On the other hand, if we remove b_2 then a is no longer accepted by any admissible set. Hence, we point out that b_2 is *critical* for the admissibility of a in AF_3.

$$AF_3$$
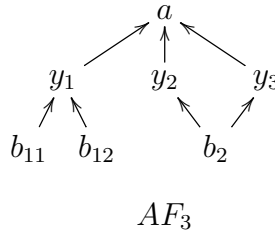
We do not deal with the dependency relations at this point. Hence we define, the notion of a *critical* argument by means of the set membership by all admissible sets that accept or attack an argument.

Definition 3.4.1. For two arguments a, b, argument b is said to be a *critical defender* for a if and only if b is in every admissible set that accepts a. Respectively, b is said to be a *critical attacker* of a, if b is in every admissible set that attacks a.

The above definition presents the critical arguments as the arguments whose acceptance is *necessary* condition for the acceptance (or rejection) of an argument. In the same fashion, the

---

[2] The word *almost* concerns the treatment of controversial arguments as presented by the example 3.2.3 of this chapter and example 2.6.4 of chapter 2.

defeat of all attackers of the critical arguments, is a *must condition* for the acceptance of that argument. The following observation makes this case.

Observation 3.4.2. *If a is a critical defender (or respectively critical attacker) of b, then, every admissible set that accepts (or respectively attacks) b, attacks any attacker of a, as well.*

3.4.1 The characterization of critical argument relation

The backings of an argument, are simply the minimal admissible sets that either accept or attack an argument. We therefore expect, for a critical defender of an argument to belong to all the positive, and, for a critical attacker of the argument, to belong to all the negative backings of that argument, as stated by the following observation.

Observation 3.4.3. *For arguments a, b in an argumentation framework AF,*

1. *b is a critical defender for a if and only if b is in every positive backing of a,*

2. *b is a critical attacker for a if and only if b is in every negative backing of a.*

Let us say that we have a set of arguments T where every argument in T is both admissible and has some critical defender a. Then we expect every admissible subsets S of T, $T \subseteq S$ to contain all such critical defenders a. The following observation states this finding.

Observation 3.4.4. *Let i be a natural number $1 \leq i \leq n$ for some natural number n. If a_i is a critical defender of an argument $b_i \in S$, then, if $S \in \gamma^{-1}(\prod_{1 \leq i \leq n} \gamma(\langle b_i \rangle^+))$ then $a_i \in S$. Respectively, if a_i is a critical attacker of b_i, then, if $S \in \gamma(\prod_{1 \leq i \leq n} \gamma(\langle b_i \rangle^-))$ then $a_i \in S$.*

On the other hand, not every argument has a critical defender or a critical attacker. Yet, there are some combinations of arguments that are indispensable to the admissibility or dismissibility of an argument. That is, if we remove that set of arguments, then there will be a change in the admissibility of the argument. Hence, we are tempted to revise the definition of critical arguments to the *critical set* of arguments.

Definition 3.4.5. In an argumentation framework AF, let C, S be two sets of arguments and a

be an argument in AF. C is a said to be a *critical set* for a if and only if C is a minimal set of arguments such that for any admissible set T that accepts (respectively attacks) a, $C \cap T \neq \varnothing$. Accordingly, C is a critical set for S if and only if C is a minimal set of arguments such that for any admissible set T that accepts (respectively attacks) every $a \in S$, $C \cap T \neq \varnothing$.

The issue with critical sets of arguments is that there are potentially many such critical sets of arguments. To list every critical set, is counter productive and against the (practical) motivation behind the formulation of critical arguments. On the other hand, we need the critical set relation to characterize the *redundant argument relation*, section 3.6. The following lemma states that the critical set relation unde critical argument relation.

Theorem 3.4.6. *Let a be some critical defender of argument b and S a critical set for a. Then, there is a critical set W for b such that $W \subseteq S$.*

Another property of critical argument relation is that the critical attacker relation is never reflexive while the critical defender relation can be reflexive. On the other hand, both the critical attacker and defender relations can be symmetric or asymmetric relations, depending on the particular framework.

We may be interested to know whether or not an argument can be at simultaneously both a critical defender and a critical attacker of some argument. The answer is yes as shown by the following example.

Example 3.4.7. In framework AF_4, below, a is both a critical defender and a critical attacker for d. In AF_4, d has one positive and one negative backing, $\langle d \rangle^+ = \{\{a, c\}\}, \langle d \rangle^- = \{\{a, b, x_1\}\}$. It is easy to see that in cases where an argument, a, is both a critical defender and critical attacker of some argument d, no backings of d attacks the argument, a. The reason is that since a is in all backings of d, any backing of d that attacks a automatically becomes a self attacking set, which leads to a contradiction. To test this, we can add an argument e where e and a symmetrically attack each other. In the resulting framework, a is no longer a critical attacker of d.

$$AF_4 \qquad\qquad AF_5$$

We may ask ourselves, how sensible are the cases as the one in AF_4? In answer, we make a reference to the simpler framework AF_5. In AF_5, argument a, for all intentional purposes, defeats the defender of d, where otherwise without a, argument d would be fully accepted. AF_5 has two preferred extensions $P_1 = \{a, c\}$, $P_2 = \{a, d\}$. P_1 is of course the sensible extension. In P_2, however, we have argument d and its *active indirect attacker* a, both in the same extension.

∎

The analysis so far covers only the base case for occurrences of the critical argument relation. Next, we discuss the *transitivity* of critical argument relation.

3.4.2   The propagation of critical argument relation

In general, we expect, the critical argument relation to be a transitive relation. That is if we are told an item a is critical for an item b, and, b is critical for c, we will then expect for a to be critical for c. The presented notions of critical attacker and critical defender do meet this expectation, as shown by the following theorem.

Theorem 3.4.8 (Transitivity property of critical argument relations). *Let a, b, c, d be four arguments in some $AF = \langle\, AR\,,\, ATT\, \rangle$, and, the relations $a\,\mathfrak{R}_1\,b$, $a\,\mathfrak{R}_2\,b$ stand for a is critical defender b, and, a is a critical attacker for b. Then, the following transitive properties hold over $\mathfrak{R}_1, \mathfrak{R}_2$ and $\hookrightarrow$ relations.*

1. *$a\,\mathfrak{R}_1\,b$, $b\,\mathfrak{R}_1\,c$ then $a\,\mathfrak{R}_1\,c$.*

2. *$a\,\mathfrak{R}_1\,b$, $b\,\mathfrak{R}_2\,c$ then $a\,\mathfrak{R}_2\,c$.*

3. *$a\,\mathfrak{R}_2\,d$, $d\hookrightarrow b$, $b\,\mathfrak{R}_1\,c$ then $a\,\mathfrak{R}_1\,c$.*

4. $a\,\mathfrak{R}_2\,b$, $d \hookrightarrow b$, $b\,\mathfrak{R}_2\,c$ *then* $a\,\mathfrak{R}_2\,c$.

Finally, we may like to know, how the acceptance of an argument depends on its critical defenders and attackers. In general, the acceptability of an argument fully depends on its critical defenders and attackers. That is, if we remove the critical defender of an argument, the argument will no longer be admissible. The controversial critical arguments are however exception to this rule, as illustrated by the following example.

Example 3.4.9. In framework AF_6, below, a is the critical attacker for c, and, the critical defender for d. We can see that the removing of a has no effect on the admissibility of c, d. This outcome is the result of a certain property of controversial arguments. The independence relation is an important topic and deserves our full attention in it own rights. Hence, we leave the discussion on the controversial arguments and its relation to the independence relation as a future topic.

$$x \longrightarrow y$$
$$a \longrightarrow b \longrightarrow c \longrightarrow d$$

$$AF_6$$

∎
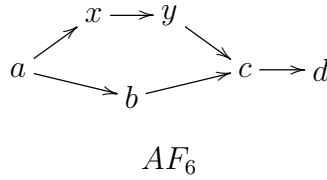
In the following sections, we will see the role that critical arguments play in identifying the various relations among arguments.

3.5 Incompatible arguments

In Dung's framework, an argument and its attacker are regarded to be *conflicting arguments*. They are conflicting in the sense that they can never belong to the same admissible set. [3] There are, however, instances where a two *non-conflicting* arguments cannot belong to the same

---

[3] The case that conflicting arguments cannot belong to the same admissible set is independent of any particular framework.

admissible set either.

For instance, in framework AF_8 below, the arguments a_1, a_2 are each is accepted by some admissible set, yet no admissible set can accept both arguments. We call these arguments the *incompatible arguments*. We define the incompatible arguments in terms of admissible sets, and, characterize them by means of their critical arguments. Moreover, this characterization is directly related to the attack cycles of even length. In addition, we investigate the incompatible argument relation in regard to the propagation of admissibility backings.

3.5.1 The Characterization of incompatible argument relation

We usually speak of the incompatibility of objects not in absolute terms, but, either in regard to some specific context, or under some quantifying measure. However, in this work, we define the incompatibility of arguments in the boolean terms, yes or no.

The conflict relation between arguments is characterized by the attack relation between arguments. Accordingly, we define the incompatible arguments relation by extending the conflict relation between arguments. We extend the criteria, *one argument is attacked by another*, to the criteria that *the admissible sets that accepts each argument attack another*. In other words, the two admissible sets accepting each argument cannot form (with respect to set union) a conflict free set.

The frameworks AF_7, AF_8, below illustrate the characterization of the incompatible arguments relation. In framework AF_8, argument a_1 is accepted by admissible set $S_1 = \{b_2\}$ and argument a_2 is accepted by admissible set $S_2 = \{b_1\}$. The admissible sets S_1, S_2, however, conflict with each other, and so, $S_2 \cup S_2$ cannot be a conflict free set. As a result $\{a_1, a_2\}$ is not accepted by any admissible set.

We test our characterization of incompatible arguments in framework AF_7. Similar to the framework AF_8, there is no admissible set that accepts $\{a_1, a_2\}$. Can we call the two arguments

a_1, a_2 incompatible ? Our characterization says otherwise. The reason is that to be regarded incompatible arguments, each argument should be accepted by some admissible set. But, no admissible set accepts a_2. Another way to look at this is that independent arguments cannot be considered incompatible, and, arguments a_1, a_2 are clearly independent of each other.

We can further elaborate on the incompatible argument relation. We have so far define the incompatible argument relation based on each argument to be accepted by some admissible set. We can also define an incompatible argument relation based on each argument to be attacked by some admissible set. It is possible to have two arguments such that while each argument is attacked by some admissible set, no admissible set attacks both. Hence, we distinguish between the two forms of incompatible argument relation. We call the former, the *positively incompatible arguments relation*, and, the latter, the *negatively incompatible arguments relation*. This way, we have a more refined notion of incompatibility that covers more cases.

The frameworks AF_8, AF_9, AF_{10} demonstrate that two forms of incompatible arguments relations. In AF_8, the two arguments a_1, a_2 are both positively and negatively incompatible, whereas in AF_9, a_1, a_2 are only positively incompatible, and, in AF_{10}, a_1, a_2 are only negatively incompatible.
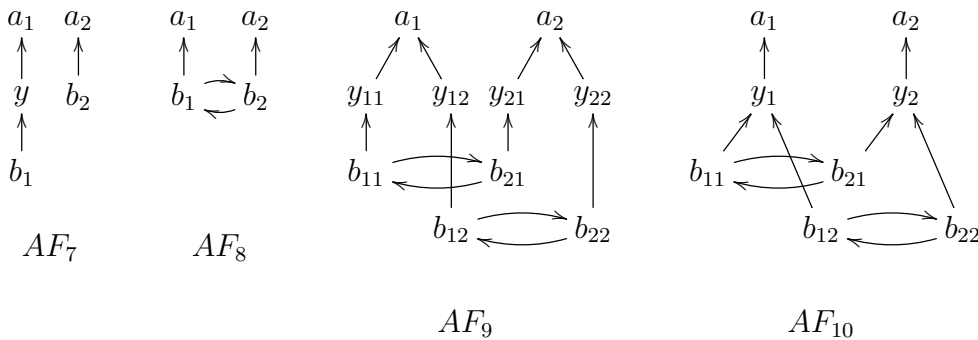
In addition, we need to remember that any two conflicting arguments are by default incompatible arguments. Hence, we define the incompatible set of arguments as follows.

Definition 3.5.1. For an argument a in an argumentation framework $AF = \langle\, AR\,,\, ATT\,\rangle$, let

$\mathcal{M}_a, \mathcal{N}_a$, respectively denote the class of admissible sets that accepts and attack argument a. Next, for a set S of arguments, let M denote some maximal $M \subseteq S$ such that for every $a \in M$, $\mathcal{M}_a \neq \varnothing$, and, respectively let ψ denote the selection functions over M where $\psi(a) \in \mathcal{M}_a$. Similarly, let N denote some maximal $N \subseteq S$ such that for every $a \in N$, $\mathcal{N}_a \neq \varnothing$, and, ϕ denote the respective selection functions over N where $\phi(a) \in \mathcal{N}_a$.

- We say a set S of arguments is *positively incompatible*, if and only if, S is not conflict free, or, for some $a \in S$, $\mathcal{M}_a \neq \varnothing$ and there is no selection functions ψ, for which $W = S \cup \bigcup_{a \in M} \psi(a)$ is conflict free.

- We say a set S of arguments is *negatively incompatible*, if and only if, S is not conflict free, or, for some $a \in S$, $\mathcal{N}_a \neq \varnothing$, and, there is no selection functions ϕ, for which $W = \bigcup_{a \in M} \phi(a)$ is conflict free.

- We say, arguments $a_1, a_2, \cdots, a_n$ are positively (respectively negatively incompatible) if and only if $S = \{a_1, a_2, \cdots, a_n\}$ is positively (respectively negatively) incompatible.

Following the definition above, we see that a *not* conflict free set is by default both positively and negatively incompatible. On the other hand, the empty set is vacuously both positively and negatively compatible. In addition, it follows that if two arguments are incompatible then the admissible sets to which one belongs are in conflict with all the admissible sets to which the other belongs. This mans that the incompatible arguments can essentially be traced back to the attack cycles of even length. The following observation points out this finding.

Observation 3.5.2. *For two admissible sets S, T, if $a \in S$ is positively or negatively incompatible with $b \in T$, then, S, T symmetrically attack each other.*

We can test the definition 3.5.1 against set $S = \{a_1, a_2\}$ in framework AF_9. for argument a_1, a_2 we have

$$\mathcal{M}_{a_1} = \{\ \{b_{11}, b_{12}\}\ \}, \mathcal{N}_{a_1} = \{\ \{b_{21}, y_{11}\}, \{b_{22}, y_{12}\}\ \}.$$

$$\mathcal{M}_{a_2} = \{ \, \{b_{21}, b_{22}\} \, \}, \mathcal{N}_{a_2} = \{ \, \{b_{11}, y_{21}\}, \{b_{12}, y_{22}\} \, \}.$$

The maximal set $M, N \subseteq S$ which for every $a \in M, N$, there are some $M_a \in \mathcal{M}_a$, $N_a \in \mathcal{N}_a$ where $M_a \neq \varnothing$, $N_a \neq \varnothing$ are $M = S, N = S$. For M there is no selection function ψ for which $S \cup \bigcup_{a \in M} \psi(a)$ is conflict free. Because, there is only one selection function ψ where $\psi(a_1) = \{b_{11}, b_{12}\}$, $\psi(a_2) = \{b_{21}, b_{22}\}$, but, $\psi(a_1) \cup \psi(a_2) = \{b_{11}, b_{12}, b_{21}, b_{22}\}$ is not conflict free. On the other hand, there is one selection ϕ where $\phi(a_1) = \{b_{21}, y_{11}\}$, $\phi(a_2) = \{b_{12}, y_{22}\}$, for which $\bigcup_{a \in M} \phi(a) = \phi_{a_1} \cup \phi_{a_2} = \{b_{21}, y_{11}, b_{12}, y_{22}\}$ is conflict free. Therefore, S is positively incompatible, but, it is not negatively incompatible.

Next, we provide an *algebraic* means for finding whether or not a set S of arguments is an incompatible set. Since admissibility backings are the minimal admissible sets that accept or reject arguments, there is an evident link between the backings of arguments and the incompatibilities of arguments. Following the definitions of incompatibilities, we see that two arguments a, b are positively incompatible if and only if $\langle a \rangle^+ \cup \langle b \rangle^+ \neq \varnothing$ and $\langle a \rangle^+ \circ \langle b \rangle^+ = \varnothing$. Respectively, a, b are negatively incompatible if and only if $\langle a \rangle^- \cup \langle b \rangle^- \neq \varnothing$ and $\langle a \rangle^- \circ \langle b \rangle^- = \varnothing$. The following theorem generalizes this finding.

Theorem 3.5.3. *For a set S of arguments, let M, N denote the maximum subsets of S where for every $a \in M$, $\langle a \rangle^+ \neq \varnothing$, and, for every $a \in N$, $\langle a \rangle^- \neq \varnothing$. A set S of arguments is positively incompatible if and only if $\sum_{a \in S} \gamma(\langle a \rangle^+) \neq \varnothing$ and $\{S\} \circ \prod_{a \in M} \gamma(\langle a \rangle^+) = \varnothing$, and, is negatively incompatible if and only if $\sum_{a \in S} \gamma(\langle a \rangle^-) \neq \varnothing$ and $\prod_{a \in N} \gamma(\langle a \rangle^-) = \varnothing$.*

The theorem 3.5.3, offers a simple way to determine whether or not a set S of arguments is incompatible. For instance, in frameworks $AF_7, AF_8, AF_9, AF_{10}$, we can draw the same conclusions for the incompatibility of arguments a_1, a_2, as before.

Example 3.5.4. Use theorem 3.5.3 to determine the incompatibility of arguments a_1, a_2 in frameworks $AF_7, AF_8, AF_9, AF_{10}$.

In the following, sets S, M, N correspond to the sets S, M, N in theorem 3.5.3. All four frame-

works share the same $S = \{a_1, a_2\}$, and, while for AF_7, $M = \{a_1\}$, $N = \{a_2\}$, for the rest of frameworks, $M = N = \{a_1, a_2\}$.

In AF_7, S is neither positively nor negatively incompatible.

$$\{S\} \circ \prod_{a \in M} \gamma(\langle a \rangle^+) = S \circ \gamma(\langle a_1 \rangle^+) = S \circ \{\{b_1\}\} = \{\{a_1, a_2, b_1\}\} \neq \varnothing.$$

$$\prod_{a \in N} \gamma(\langle a \rangle^-) = \gamma(\langle a_2 \rangle^-) = \{\{b_2\}\} \neq \varnothing.$$

In AF_8, S is both positively and negatively incompatible.

$$\{S\} \circ \prod_{a \in M} \gamma(\langle a \rangle^+) = \{S\} \circ \gamma(\langle a_1 \rangle^+) \circ \gamma(\langle a_2 \rangle^+) = \{S\} \circ (\{\{b_2, \star\}\} \circ \{\{b_1, \star\}\}) =$$
$$\{S\} \circ \varnothing = \varnothing.$$

$$\prod_{a \in M} \gamma(\langle a \rangle^-) = \gamma(\langle a_1 \rangle^-) \circ \gamma(\langle a_2 \rangle^-) = \{\{b_1, \star\}\} \circ \{\{b_2, \star\}\} = \varnothing.$$

In AF_9, S is only positively incompatible.

$$\{S\} \circ \prod_{a \in M} \gamma(\langle a \rangle^+) = \{S\} \circ \gamma(\langle a_1 \rangle^+) \circ \gamma(\langle a_2 \rangle^+) = \{S\} \circ \{\{b_{11}, b_{12}, \star\}\} \circ \{\{b_{21}, b_{22}, \star\}\} =$$
$$\{S\} \circ \varnothing = \varnothing.$$

$$\prod_{a \in N} \gamma(\langle a \rangle^-) = \gamma(\langle a_1 \rangle^-) \circ \gamma(\langle a_2 \rangle^-) = \{\{b_{21}, y_{11}, \star\}, \{b_{22}, y_{12}, \star\}\} \circ \{\{b_{11}, y_{21}, \star\}, \{b_{12}, y_{22}, \star\}\}$$

$$= \{\{b_{11}, b_{22}, y_{12}, y_{21}, \star\}, \{b_{12}, b_{21}, y_{11}, y_{22}, \star\}\} \neq \varnothing.$$

In AF_{10}, S is only negatively incompatible.

$$\{S\} \circ \prod_{a \in M} \gamma(\langle a \rangle^+) = \{S\} \circ \gamma(\langle a_1 \rangle^+) \circ \gamma(\langle a_2 \rangle^+) = \{S\} \circ \{\{b_{11}, \star\}, \{b_{12}\}, \star\} \circ \{\{b_{21}, \star\}, \{b_{22}, \star\}\}$$

$$= \{S\} \circ \{\{b_{11}, b_{22}, \star\}, \{b_{12}, b_{21}, \star\}\} = \{\{a_1, a_2, b_{11}, b_{22}, \star\}, \{a_1, a_2, b_{12}, b_{21}, \star\}\} \neq$$
$$\varnothing.$$

$$\prod_{a \in N} \gamma(\langle a \rangle^-) = \gamma(\langle a_1 \rangle^-) \circ \gamma(\langle a_2 \rangle^-) = \{\{b_{21}, b_{22}, y_1, \star\}\} \circ \{\{b_{11}, b_{12}, y_2, \star\}\} = \varnothing.$$

$\blacksquare$

We can use theorem 3.5.3 to draw a more direct connection between the incompatibility of

arguments and the attack cycles of even length, than the one stated by observation 3.5.2. If two arguments are incompatible then they both have some positive and some negative backings. Hence, the root of their admissibility lies within some attack cycles of even length. The following observation states this finding.

Observation 3.5.5. *For two distinct arguments a, b, if $\langle a \rangle^+ \neq \varnothing$, $\langle b \rangle^+ \neq \varnothing$ and $\langle a \rangle^+ \circ \langle b \rangle^+ = \varnothing$ then $\langle a \rangle^- \neq \varnothing$, $\langle b \rangle^- \neq \varnothing$. Moreover, if $\langle a \rangle^- \neq \varnothing$, $\langle b \rangle^- \neq \varnothing$ and $\langle a \rangle^- \circ \langle b \rangle^- = \varnothing$ then $\langle a \rangle^+ \neq \varnothing$, $\langle b \rangle^+ \neq \varnothing$.*

We defined incompatibility of arguments with respect to sets of arguments, while at the same time discussed and extended the notion of incompatibility as a relation among arguments. This treatment of incompatibility allows us to study different aspects of compatibility.

3.5.2 The propagation of incompatible arguments relation

The incompatibility as a relation is not transitive. If a is incompatible (or compatible) with b, and, b is incompatible (or compatible) with c, then, a is not necessarily incompatible (or compatible) with c. On the other hand, we can say, the incompatibility over sets of arguments is monotonic under set inclusion. This feature of incompatibility follows the property of *not being conflict free* over sets of arguments which is monotonic under set inclusion.

Observation 3.5.6. *The properties positively incompatible and negatively incompatible are monotonic under set inclusion.*

The incompatibility of arguments serves three basic roles. One is in relation to the acceptance of arguments. The other is in conjunction with critical arguments. The last one is in relation to the propagation of admissibility backings. In the followings, we will look at these three roles.

The first of these roles is in relation to the acceptance of arguments. To illustrate this function of incompatibility of arguments, let us extend frameworks AF_8, AF_9, AF_{10} by argument d and attack relations (a_1, d), (a_2, d). In all these frameworks, if d is to be accepted by some admissible

set S, S must attack both a_1 and a_2. In AF_8, AF_{10}, arguments a_1, a_2 are negatively incompatible, so, no admissible set attacks both a_1, a_2. Hence, d cannot belong to any admissible set. On the other hand, in AF_9, there is no such issue where a_1, a_2 are not negatively incompatible. Hence, there is some admissible set that accepts d. The following observation generalizes this finding.

Observation 3.5.7. *For an argument a, $\langle a \rangle^+ = \varnothing$ if and only if either for some $b \in \bar{a}$ where $\langle b \rangle^- = \varnothing$ or there is some subset of $\bar{a}$ which is negatively incompatible.*

We set to categorize the relation between arguments with respect to active attack sequences. As part of categorization we can show that for two incompatible arguments there is always some active attack sequence between some of their active arguments. That is if all the attack sequences between their active arguments are intercepted then the two arguments cannot be incompatible. The following lemma states this finding.

Lemma 3.5.8.   *1. If two admissible sets S, T attack each other, then, there are some $S' \subseteq S$, $T' \subseteq T$ for which there is an active attack sequence between every $a \in S'$ and $b \in T'$.*

   2. If a, b are incompatible then there is some active attack sequence between an active argument for a and an active argument for b.

The second function of incompatibility of arguments is in regard to the critical argument relation; how the critical argument relation gives rise to the incompatible arguments relation, and, how the incompatible arguments relation gives rise to the critical argument relation. Let us first look at how the critical argument relation gives rise to the incompatible arguments relation.

We start with the simplest case. Let b be some critical defender of an argument a. Then, any argument, c, that attacks b, will automatically conflict with any admissible set that accepts a. Now, if c belongs to some admissible set itself, then, c becomes positively incompatible with a. For example, in AF_9, argument b_{12} attacks b_{11} where b_{11} is a critical defender of a_1. Hence, b_{11}

is positively incompatible with a_1. The following observation captures this finding.

Observation 3.5.9. *For arguments a, b, c where $a \hookrightarrow b$ and a belongs to some admissible set, if b is a critical defender of c, then, a is positively incompatible with c; and, if b is a critical attacker of c, then, a is negatively incompatible with c.*

We can generalize the above observation by replacing the clause, $a \hookrightarrow b$ *and a belongs to some admissible set* by a *and b are positively incompatible*, which is the lemma below.

Lemma 3.5.10. *For arguments a, b, c where a is positively incompatible with b, if b is a critical defender of c, then a is positively incompatible with c, and, if b is a critical attacker of c, then a is negatively incompatible with c.*

We can further generalize the above lemma, and say that, if the critical defenders of arguments are incompatible, then, so are the arguments, as presented by the following theorem. For example, in AF_9, since b_{11}, b_{21}, the critical defenders of a_1, a_2, are positively incompatible, a_1, a_2 are positively incompatible, too. The following theorem can also serve for the *propagation of incompatible arguments under the critical argument relation.*

Theorem 3.5.11. *For arguments a_1, a_2, b_1, b_2 where b_1 is a critical defender for a_1, and, b_2 a critical defender for a_2, if b_1 is positively incompatible with b_2, then, a_1 is positively incompatible with a_2, and, if b_1 is negatively incompatible with b_2, then, a_1 is negatively incompatible with a_2.*

So far, we only showed, how critical arguments can give rise to the incompatible arguments. We still need to discuss how the incompatible arguments can give rise to the critical arguments. However, it is best if we discuss the third key function of the incompatibility of arguments first. The last of the key functions of incompatibility of arguments is in regard to the propagation of admissibility backings.

There are a few special cases that prevent us to answer *yes* to the question (3.2.4). One of these special cases is in relation to the incompatibility of arguments. The others will be discussed in

the following sections. To visualize the problem, we present the following framework.

In framework AF_{11} below, b_1, b_2 are both critical defenders for argument a. The argument b_1 itself has one critical defender, namely, c_1. Argument b_2 has two positive backings, $S_1 = \{c_2\}$, $S_2 = \{d\}$. On the other hand, c_1, c_2 are both positively and negatively incompatible arguments.

$$d \searrow$$
$$y_2 \rightarrow b_2 \rightarrow x_2$$
$$c_2 \nearrow \qquad\qquad \searrow$$
$$\qquad\qquad\qquad a$$
$$c_1 \rightarrow y_1 \rightarrow b_1 \rightarrow x_1 \nearrow$$
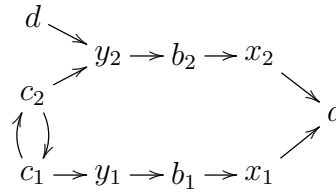
$$AF_{11}$$

The question in (3.2.4) asks, given that $\{c_2\}$ is a positive backing for b_2 and b_2 is a critical defender of a then does c_2 appear in some backing of a ? Normally, we expect the answer to be *yes*, but, on this occasion, the answer is *no*. The reason is that c_2 is incompatible with b_1, so, b_1 and c_2 cannot belong to any admissible set. On the other hand, b_1 is a critical defender of a, hence, any backing of a must have c_2. Consequently, c_2 cannot belong to any admissible set that accepts a. The following lemma presents this finding.

Lemma 3.5.12. *For arguments a, b, c, let $a \in S$ for some $S \in \langle b \rangle^+$ and $b \in T$ for some $T \in \langle c \rangle^+$, then, $a \in R$ for some $R \in \langle c \rangle^+$ only if a is not positively incompatible with some critical defender of c.*

The above lemma is in the form of *only if* statement, and, does not characterize the propagation of backings. The reason is, the incompatibility of arguments is only one of the conditions that interferes with the propagation of backings. A gradual and then a full characterization of propagation of backings will be presented in the following sections.

Next, to the question that how the incompatibility of arguments can give rise to the critical arguments. From the previous section, we know that the critical argument relation is transitive.

Let us consider the scenario where an argument a belongs to some positive backings of b, and, b belongs to some positive backings of c, while a is not a critical defender for b. Argument a is not a critical defender for b. Hence, the transitivity of critical argument relation does not apply here. This however does not mean that a cannot be a critical argument for c.

As an example, in framework AF_{11} above, b_2 is a critical defender for a, and d while being an active defender of b_2, is not a critical defender of b_2. Yet, due to the incompatibility of c_2 and b_2, c_2 cannot belong to any positive backings of a. As a result, d becomes a critical defender for a. This concludes our investigation of incompatibility of arguments.

3.6 Redundant Arguments

We follow up on the central question of (3.1.1) on page 78, that under what conditions, apart from the incompatibility of arguments, an active defender of an argument b where b itself is an active defender of some argument c, does not take part in any of the backings of argument c. These, otherwise, to be active defenders (or attackers) of an argument, are then phrased to be *redundant* for the admissibility of that argument.

All the instance of the redundant argument relation are due to the minimality clause in the definition of backings of arguments. The frameworks AF_{15}, AF_{16} below show the two basic cases where an argument becomes redundant for the admissibility of another. In AF_{15}, argument b which is accepted by some admissible set does not take part in any of the negative backings of argument a despite being a direct attacker of a. In AF_{16}, b is this time a defender of a, but, it still does not belong to any of the positive backings of a.

$$AF_{15} \qquad AF_{16} \qquad AF_{17}$$
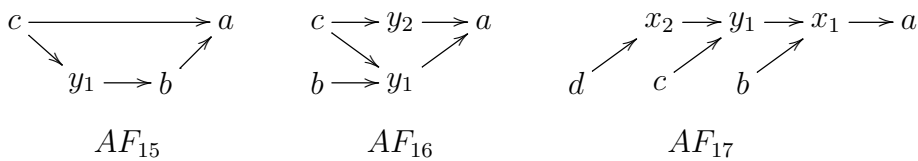
The reason for argument b becoming redundant for the admissibility of a is simple. It is due to the minimality clause of the admissibility backings. To show that b is not required for the admissibility of a, we can remove b. If we do so, we see that the admissibility situation of none of the arguments will change. On the other hand, if we remove c, then, the admissibility situation of many arguments including a, will change. Accordingly, we define the *positive redundant* and the *negative redundant* argument relations.

Definition 3.6.1. For two compatible arguments a, b, argument b is said to be *positively redundant* for a if and only if there is a positively active attack sequence from a to b, but, b is not in any positive backings of a. Respectively, b is said to be *negatively redundant* for a if and only if b is an in/direct attacker of a and there is a negatively active attack sequence from a to b and b is not in any negative backings of a.

The above definition uses the positive and negative active attack sequences. The active argument relation pertains some strong *dependency* relation, such that if a is an active argument for b and b an active argument for c then a is an active argument for c. On the other hand, the redundant argument relation states a lack of a strong dependency relation in places where we normally expect to find one. In this work, we do not address the *dependency chains*. Therefore, we need to appeal to other means to state the role of *dependency chains* in regard to the redundant argument relation. Our other means is to use the positive and negative active attack sequences. From now on, unless specified otherwise, by the *redundant relation* or *redundant argument*, we are referring to the redundant argument relation.

In our customary fashion, we are primarily interested in two aspects of the redundant relation between arguments, the base case for how they are formed, and, how they propagate along the attack sequences.

3.6.1 The characterization of redundant argument relation

There is a strong connection between the redundant argument relation and the critical argument relation, as illustrated by frameworks AF_{15}, AF_{16}. To identify, elaborate and expand on this connection, we employ the help of additional frameworks, $AF_{15a}, AF_{15b}, AF_{15c}$ and $AF_{16a}, AF_{16b}, AF_{16c}$ that all are the extended forms of frameworks AF_{15}, AF_{16}.

Occurrences of the redundant argument relation can be traced back to two basic forms, shown in AF_{15}, AF_{16}. There is a certain condition under which an argument becomes redundant for another. Frameworks AF_{15a}, AF_{16a} show how this required condition is preempted. Frameworks AF_{15b}, AF_{16b} show the extended forms of the two basic cases of redundant argument relation shown in AF_{15}, AF_{16}. The extended forms of the redundancy relation rely on the critical set relation instead of the critical argument relation. Finally, frameworks AF_{15c}, AF_{16c} show the backward propagation of redundancy. This also includes the redundancy by transitivity property of the redundant relation, under the attack relation.

$$AF_{15} \qquad AF_{15a} \qquad AF_{15b} \qquad AF_{15c}$$

$$AF_{16} \qquad AF_{16a} \qquad AF_{16b} \qquad AF_{16c}$$
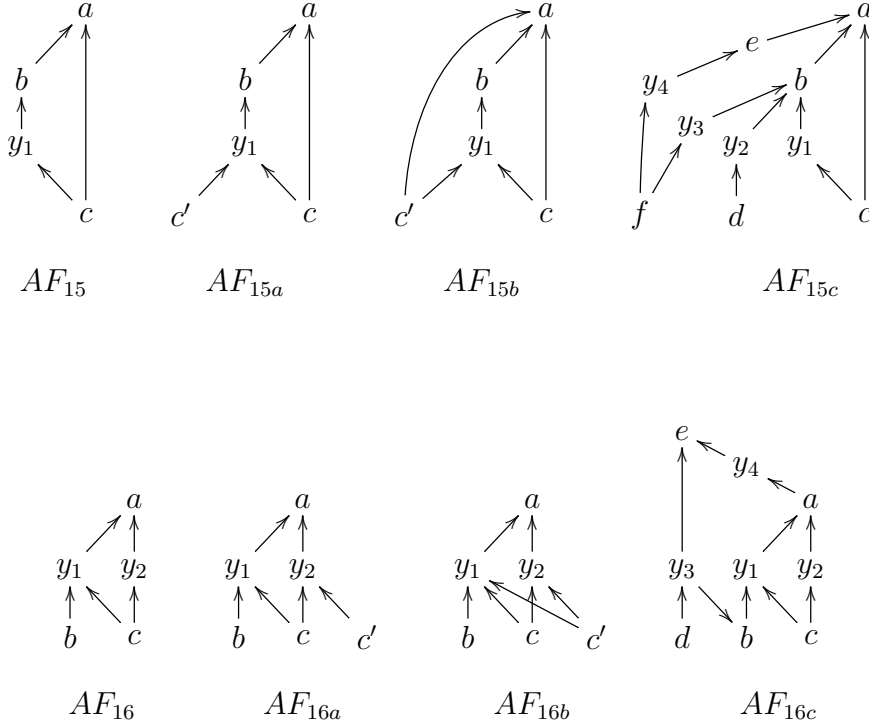
We first look at framework AF_{15}, showing how an argument becomes negatively redundant for

another. In AF_{15}, argument c is a critical defender for b. Hence, any admissible set that includes b, is bound to include both c and some positive backing of c (in this instance, c has one positive backing which the empty set). We then conclude that any admissible set that includes b cannot be a minimal admissible set that attacks a. Hence, we hypothesize that,

> a necessary condition for b to be a *negatively redundant argument* for a, is for some attacker of a to be a critical defender for b.

We can test this hypothesis by making c, not a critical defender for b, as is shown in framework AF_{15a}. We see that, now, a has a negative backing that includes b, namely $\{c', b\}$. The following observation presents this.

Observation 3.6.2. *For arguments a, b, c where $b, c \in \bar{a}$, if c is a critical defender for b, then, b is negatively redundant for a.*

We can generalize observation 3.6.2, by expanding on the condition, c *is a critical defender for* b. Looking at AF_{15}, we see that all that is required to render argument b redundant for a is some combination of attackers of a be critical for the admissibility of b. The framework AF_{15b} illustrates this finding. In AF_{15b}, either c or c' makes b admissible. On the other hand, either c or c' makes a automatically dismissible. Hence, b is redundant for the dismissibility of a. The following lemma formulates this finding.

Lemma 3.6.3. *For arguments a, b where $\langle a \rangle^- \neq \varnothing$ and $b \in \bar{a}$, argument b is negatively redundant for a, if and only if, there is some critical set S for b such that $S \subseteq \bar{a}$.*

Next, we look at how the positive redundant relations are formed. In a similar fashion, we start with the basic case, the framework AF_{16}. In AF_{16}, any admissible set that accepts a, must attack both y_1, y_2. Moreover, any admissible set that attacks y_2, must include c, as well. Now, since c attacks all the arguments that b attacks, effectively c makes b redundant for the admissibility of a. Hence, we may hypothesize that,

> a necessary condition for b to be a positively redundant argument for a is for some critical defender of a to defend a against all the attackers of a that b attacks.

Again, we can test this hypothesis, by making c, not a critical defender for a. In AF_{16a}, we add argument c' where c' attacks y_2. We then see that b is no longer positively redundant for a, as a now has a positive backing $\{b, c'\}$. The following observation states this finding.

Observation 3.6.4. *For arguments a, b, c where b, c are two defenders of a, if c is a critical defender of a and any not intercepted in/direct attacker of a that b attacks, c attacks that argument as well, then, c makes b positively redundant for a.*

Observation 3.6.4 captures the simplest form of positive redundant argument relation. We can then generalize observation 3.6.4 by expanding the conditions in the antecedent of 3.6.4. Framework AF_{16b} shows a simple case of this generalization. In AF_{16b}, instead of the critical defender c, there is a set of defenders of a, namely $\{c, c'\}$ that defend a against all its attackers. Again, the presence of either c or c' renders b redundant for the admissibility of a. The following lemma formulates this generalization of the primary case in observation 3.6.4. The generalization of primary cases of positive redundant relation is, however, more complicated than the generalization of the primary cases of negative redundant relation.

Lemma 3.6.5. *For two admissible and compatible arguments a, b, let D denote the set of all critical defenders of a, and, Y denote the set of all $y \in \bar{a}$ for which b is an active attacker of y where $Y \neq \varnothing$, and, Π denote the set of all active attack sequences form $y \in Y$ to b. Then, b is positively redundant for a if and only if there is some critical set C for $W = \bar{a} - Y$ such that for any $\pi \in \Pi$, there is some c on π where $D \hookrightarrow c$, or, $d \hookrightarrow c$ for every $d \in C$.*

We can test the above lemma against frameworks AF_{17a}, AF_{17b} below. In both frameworks, argument a is admissible, where in AF_{17a}, $\langle a \rangle^+ = \{\{c_{11}\}, \{c_{21}\}, \{c_{12}\}, \{c_{22}\}\}$, and, in AF_{17b}, $\langle a \rangle^+ = \{\{c_{11}, c_{22}\}, \{c_{12}, c_{21}\}\}$. In both frameworks, b is an active attacker of z_2 where $z_2 \in \bar{a}$. Argument b however does not belong to any backings of a. Hence,we can say that in both frameworks b is positively redundant for a.

Example 3.6.6. In the following frameworks AF_{17a}, AF_{17b}, use theorem 3.6.5 to show that b is positively redundant for a.

$$AF_{17a} \qquad\qquad AF_{17b}$$
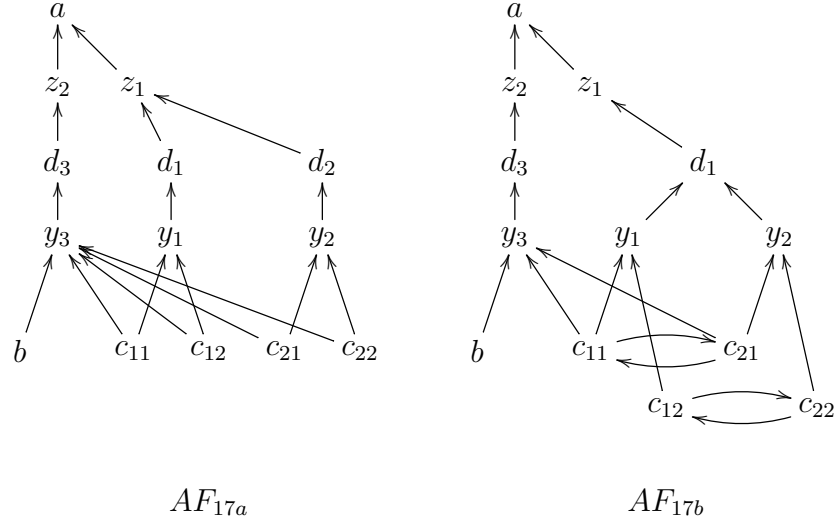
The parameters of lemma 3.6.5 for a in each framework are as follows. In both frameworks, $a = a, b = b$, D the set of critical defenders of a is $D = \varnothing$, $Y = \{y \in \bar{a} \mid b \text{ is an active attacker of } y\} = \{z_2\}$, $\Pi = \pi \mid$ active attack sequence from $y \in Y$ to $b = \{(z_2, d_3, y_3, b)\}$, and, $W = \bar{a} - Y = \{z_1\}$. Although, z_1 has no critical defender in both frameworks, z_1 has some critical set C, definition 3.4.5, in both frameworks. In AF_{17a} the critical set for W is $C = \{c_{11}, c_{12}, c_{21}, c_{22}\}$, and, in AF_{17b} the critical set for W is $C = \{c_{11}, c_{21}\}$. As a note, C by definition of critical sets is also a critical set for a. Moreover, if we take any of the critical sets out of each framework, argument d will be admissible and argument a will be rejected. We can see that in both framework, there is some c on $\pi = (z_2, d_3, y_3, b) \in \Pi$ where $c = y_3$ such that for every $d \in C$, $d \hookrightarrow c$. Hence, we can say that b is positively redundant for a. ∎

Next, we look at how the redundant argument relation propagates along the attack sequences.

3.6.2   The propagation of redundant argument relation

The propagation of redundant argument relation can be both forward and backward under the attack relation. In simple terms, if c is redundant for a, then, any active argument b for c will also be redundant for a, unless, b can reach a by another active sequence. The reason is that any active role that b is to play with respect to the admissibility of a, stops at c.

For instance, in frameworks AF_{15c}, since b is negatively redundant for a, d is also negatively redundant for a. Respectively, in AF_{16c}, since b is positively redundant for a, d is also positively redundant for a.

For the propagation of redundant relations to work, there are certain strings attached. The required condition is that there should be no other active attack sequences from a to b. Otherwise, there are some active paths by which b can participate in the admissibility, or the dismissibility, of a.

For instance, in AF_{15c}, while d is negatively redundant for a, argument f is not. The reason is that there is some attack sequence π, namely $\pi = (a, e, y_4, f)$, via which f becomes an active attacker of a. Similarly in AF_{16c}, although, d is positively redundant for a, d is not positively redundant for e. This is so because, there is an active positive attack sequence $\pi = (e, y_3, d)$.

The following theorem formulates these findings in regard to the propagation of redundant argument relations. This formulation is, however, done with respect to the defense relation, instead of the attack relation. [4]

Lemma 3.6.7. *For arguments a, b, let C be some set of arguments c such that $C \neq \varnothing$ and for every active attack sequence π from a to b, there is some $c \in C$ on π.*

1. *If b is an active defender for all $c \in C$, then,*

   (a) *if all $c \in C$ are positively redundant for a, then, b is positively redundant for a,*

   (b) *if all $c \in C$ are negatively redundant for a, then, b is negatively redundant for a.*

2. *If b is positively redundant for all $c \in C$, then,*

   (a) *if all $c \in C$, are active defenders for a, then b is positively redundant for a,*

---

[4] To formulate the propagation of redundancies under the attack relations, the defeated arguments on active paths must be termed and accounted. A number of additional definitions are required to account for these defeated arguments. To avoid the additional definitions, we formulate the propagation of redundancy relations under the defense relation.

(b) if all $c \in C$ are active attackers for a, then b is negatively redundant for a.

The first part of lemma 3.6.7 addresses the backward propagation of redundancy relation along the sequences of attack relations. In simple terms, if c is redundant for a, then, the active in/direct attackers or defenders b of c will be redundant for a as well. The second part of lemma 3.6.7 deals with the forward propagation of redundancy relation. That is, if b is redundant for c, then, b is redundant for all arguments a where c is an in/direct attacker or defender for a.

We can now put the three lemmas 3.6.5, 3.6.3, 3.6.7, the two base cases and the propagation case, together and completely characterize the redundant argument relation. The following theorem formulates this characterization.

Theorem 3.6.8. *For two compatible arguments a, b, let D denote the set of all critical defenders of a, Y denote the set of all $y \in \bar{a}$ for which b is an active attacker of y, and, Z, W denote the sets of all $z, w \in \bar{a}$ for which there is, respectively, some positive attack sequence from z to b, and, from w to b.*

1. *b is negatively redundant for a, if and only if, one of the lemmas 3.6.3, 3.6.7 applies with respect to a and b.*

2. *b is positively redundant for a, if and only if, one of the lemmas 3.6.5, 3.6.7 applies with respect to a and b.*

One last issue is left to complete our discussion on how some arguments make other arguments redundant for admissibility of an argument, and, that is in relation to the arguments that take part in their own defense.

3.6.3 Redundancy by self-defense

We presented the backings as the minimal admissible sets of arguments that are sufficient for the acceptance, or the rejection of an argument. In doing so, we intentionally overlooked one

issue, and that is in regard to the arguments that take part in their own defense.

Whenever we contemplate the admissibility of an argument, by default we presume that the argument is present. Hence, if we discuss the admissibility of an argument a that defends itself against some attackers, we always assume that a is present. However, there is a chance that a defends itself against some attacker b that a defender c of a also defends a against that attacker, argument b. Now, since a is always present, a effectively makes c redundant for the defence of a.

For instance, in framework AF_{18a} below, argument a defends itself against b where b also defends itself against c. Under the current formulation of backing of arguments, a has two positive backings, $\langle a \rangle^+ = \{\{a\}, \{c\}\}$.

$$AF_{18a} \qquad AF_{18b} \qquad AF_{18c}$$
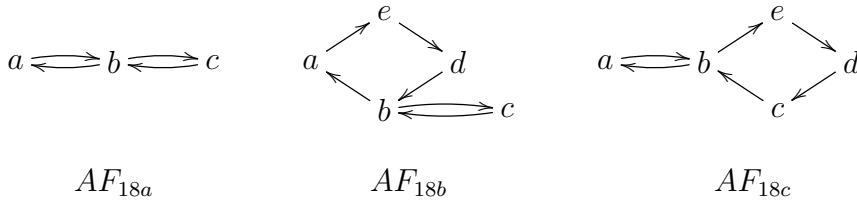
However, upon a closer inspection, we realize that argument c is effectively made redundant for a. If a is present then there is no need for c to defend a against b. Hence, a makes c *categorically redundant* for its admissibility, regardless of the sub-framework of choice.

The frameworks AF_{18b}, AF_{18c} show that the symmetric attack relation between a, b and between b, c, as in AF_{18a}, is a necessary condition for making c redundant for admissibility of a. The reason is, in either framework AF_{18b}, AF_{18c}, we can find a sub-framework, for which c makes a difference in the admissibility of a.

An underlying motivation behind the admissibility backings is for the backings to capture the dependency of one argument for its admissibility upon the admissibility of another. Under this motivation, $\{c\}$ cannot be regarded as positive backing for a, as it serves no function in making a admissible.

To make sure that the backings are relevant to the minimal required condition for the admissibility arguments, we are then left with two choices. Either amend the definition of backings of arguments, or, leave the admissibility backings as they are and address the redundancy of arguments by self defending arguments in another way.

We chose the latter option. Had we gone with the former option, we needed to define something in the form of a *proper backing* of an argument, something along the lines, *a positive backing S for a is a proper positive backing for a if and only if it is either a grounded backing for a, or, $S \cup \{a\}$ is the minimal member of $\{B \mid B = A \cup \{a\}, A \in \langle a \rangle^+\}$*. To leave the formulation of backings simple, we chose the latter option. The adverse effects of not changing the definition of backings is little, as the redundancies not caught by the current definition of the backings do not propagate forward. For instance, in AF_{18a} whenever argument a participates in the attack or the defense of an argument, the minimality clause in definition of the backings of arguments, automatically makes c redundant for the admissibility of that argument. As an example, in framework AF_{19a} below, argument a makes c a redundant argument for both d and e.

$$e \longleftarrow d \longleftarrow a \rightleftharpoons b \rightleftharpoons c \qquad\qquad e \longleftarrow d \longleftarrow a \rightleftharpoons b \longleftarrow c$$
$$\Updownarrow$$
$$f$$

$$AF_{19a} \qquad\qquad\qquad AF_{19b}$$
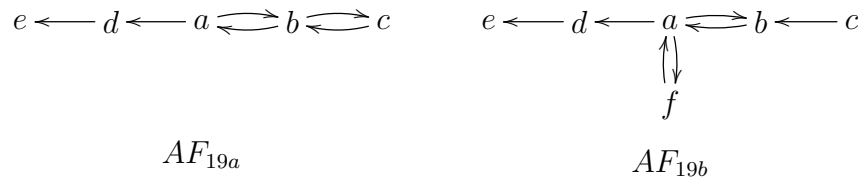
The same mechanism also stops the propagation of certain grounded backings. For instance, in framework AF_{19b}, $\{c\}$ is a positive grounded backing for a. But, a, f make c a redundant argument for both d, f. In other words, argument c is stopped to be an active argument for d.

There is more to be said about the redundant argument relation, e.g. the *conditional redundant argument relation*. By *conditional redundant argument relation* we mean given a sub-framework $AF' \sqsubseteq AF$, which arguments remain redundant for some argument a under the sup-

framework relation. But, we stop our analysis of the redundant argument relation here.

3.6.4 Characterization of attack sequences by the roles of arguments

With the introduction of redundant argument relation, we can finally put the propagation of backings and the active and intercepted attack sequences under one picture. The positive and negative active attack sequences help us to distinguish between the redundant argument relation and the intercepted argument relation. The redundant argument relation requires the existence of some active attack sequence while the intercept relation requires the lack of such existence. For instance, in framework AF_{14}, all the attack sequences form a to d are intercepted. Hence, the explanation for why d, an indirect defender of a, is not an active argument for a is that d is intercepted for a.

With the introduction of redundant argument relation, we can finally put the propagation of backings and the active and intercepted attack sequences under one picture as presented by theorem 3.6.9 below. Theorem 3.6.9 classifies the roles that arguments play in relation to the admissibility or dismissibility of other arguments in terms of the classification of attack sequences.

Theorem 3.6.9. *For arguments a, b, let $\langle b \rangle^+ \neq \varnothing$ and π an attack sequence from a to b.*

1. *π is positively active if and only if b is either an active defender or a positively redundant argument for a.*

2. *π is negatively active if and only if b is either an active attacker or a negatively redundant argument for a.*

3. *π is active if and only if b is either an active or a redundant or an incompatible argument for a.*

This concludes our analysis of the role of arguments and its relation with the active attack se-

quences. In the next, section, we use a distinct feature of the intercepts in splitting a framework into pseudo disjoint sub-frameworks.

3.7 Intercepts and the disjoint sub-frameworks

The intercepts have one distinct feature which is most useful for the partitioning and merging of argumentation frameworks in that they split a framework into pseudo disjoint frameworks. We next briefly investigate this feature of the intercepts. [5]

In order to facilitate our discussion, we introduce the two operations $+^N, -^N$ over an argumentation framework. The operations $+^N, -^N$ are the simplest of operations for the merging and partitioning frameworks into the minimal normal sub-frameworks. Our attention here is solely to establish the independence relation among frameworks. The establishing of independence relation requires considering all the possible attack relations. For this reason, we define the operations $+^N, -^N$. The superscript 'N' is in regard to the closure by the minimal normal sub-framework. The operations $+^N, -^N$ are then the minimal normal sub-frameworks that cover the otherwise usual merging and splitting operations $+, -$.

The operations $+^N, -^N$ over an argumentation framework AF are given such that for any $AF_1, AF_2 \sqsubseteq AF$,

$AF_1 +^N AF_2 = AF_3$ if and only if AF_3 is the minimum $AF_3 \sqsubseteq^N AF$ where $AF_1, AF_2 \sqsubseteq AF_3$.

$AF_1 -^N AF_2 = AF_3$ if and only if AF_3 is the minimum $AF_3 \sqsubseteq^N AF$ where $AF_4 \sqsubseteq AF_3$ and AF_4 is the maximum $AF_4 \sqsubseteq AF_1$ for which there is no non empty AF_5 where $AF_5 \sqsubseteq AF_4$, $AF_5 \sqsubseteq AF_2$.

It is easy to see that, for two sub-frameworks AF_1, AF_2, there are always two unique AF_3, AF_4 for which $AF_1 +^N AF_2 = AF_3$ and $AF_1 -^N AF_2 = AF_3$. Furthermore, for a sub-framework

---

[5] Since in this work, we do not address the splitting and merging of argumentation frameworks, we only give a brief attention to the function of intercepts in splitting a framework into pseudo disjoint frameworks.

AF', the sub-framework $AF'' = AF' +^N AF'$ is the minimum normal sub-framework that covers AF'.

Next, given a set $\mathcal{AF}$ of sub-frameworks $AF_i \sqsubseteq^N AF$, $1 \leq i \leq n$ where n is some natural number, we can define the operation $\sum$ over the sets of such $\mathcal{AF}$ in the usual way, such that,

$$\sum_{AF' \in \mathcal{AF}}^{N} AF' = AF_1 +^N AF_2 +^N \cdots +^N AF_n.$$

Furthermore, for two sub-frameworks AF_1, AF_2, we can define

their *shared sub-framework* $AF_3 \sqsubseteq^N AF$ as the minimum normal sub-framework AF_3 that covers the maximum sub-framework AF_4 where $AF_4 \sqsubseteq AF_1, AF_4 \sqsubseteq AF_2$. We call two sub-frameworks AF_1, AF_2 *intersecting* if have some *shared* sub-framework, otherwise, we say, they are *not-intersecting* sub-frameworks.

The two sub-argumentation frameworks are said to be *completely disconnected* if there is no undirected attack path from one to another. Hence, for two sub-frameworks $AF_1, AF_2 \sqsubseteq AF$, we say,

AF_1, AF_2 are said to be *disjoint* if and only if there is no undirected attack path from any argument in AF_1 to any argument in AF_2.

Any two disjoint sub-argumentation frameworks, both act like, and, can be treated like totally independent frameworks. That is, any information regarding one sub-framework reveals no information regarding the other.

In this work, we do not address the dependence/independence relations among the arguments, and among the argumentation frameworks. However, to make our discussion clear, we provide a more targeted readings of the term *information* and the *independence relation* between sub-frameworks. We cap the domain of term *information* to the information on whether or not an argument or a set of arguments is admissible.

Hence, the reading of *independence relation* between two sub-frameworks AF_1, AF_2 can be

narrowed to,

Note 3.7.1. AF_1, AF_2 are said to *act like independent frameworks*, if the admissibility of any set in AF_1 is independent of any argument in AF_2, and vice versa.

We can further refine the notion of independence relation given in (3.7.1). We can replace the expression *admissibility of any set in AF_1 is independent of any argument* $\cdots$ in (3.7.1) by, *if we merge AF_1 with AF_2, then, all the admissible sets in the original AF_1 remain admissible after the merge.*

The following lemma puts this reading of independence between AF_1, AF_2 into a more succinct relation that if AF_1, AF_2 are to be considered independent then the set of admissible sets $\mathcal{A}_3$ in the merged sub-framework AF_3 is the product of set of admissible sets $\mathcal{A}_1$ in AF_1 with the set of admissible sets $\mathcal{A}_2$ in AF_2, i.e., $\mathcal{A}_3 = \mathcal{A}_1 \dot{\circ} \mathcal{A}_2$.

Lemma 3.7.2. *For two sub-frameworks $AF_1 = \langle AR_1, ATT_1 \rangle$, $AF_2 = \langle AR_2, ATT_2 \rangle$ of a framework AF, let $AF_3 = AF_1 +^N AF_2$ and $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_{31}, \mathcal{A}_{32}$ denote the sets of admissible sets in AF_1, AF_2, AF_3, $\mathcal{A}_{31} = \{S \mid S = T \cap AR_1, T \in \mathcal{A}_3\}$, $\mathcal{A}_{32} = \{S \mid S = T \cap AR_2, T \in \mathcal{A}_3\}$. Then, $\mathcal{A}_{31} = \mathcal{A}_1, \mathcal{A}_{32} = \mathcal{A}_2$ if and only if $\mathcal{A}_3 = \mathcal{A}_1 \dot{\circ} \mathcal{A}_2$.*

In the following passages, we present lemma 3.7.5, an equivalent reading of the above lemma that is in terms of the active argument and the intercepted relations. The formulation of the sub-frameworks independence relation AF_1, AF_2 in lemma 3.7.2 is made in terms of the independence among the admissible sets in AF_1 against those in AF_2. However, the state of relation $\mathcal{A}_3 = \mathcal{A}_1 \dot{\circ} \mathcal{A}_2$ in lemma 3.7.2 may only be accidental, and, an outcome of circumstance with respect to the current state of AF_1, AF_2. A more solid claim to the independence of AF_1, AF_2 can be made in regard to the independence of admissible sets against all possible sub-frameworks of AF_1, AF_2. Hence, we formulate the notion of independence relation in (3.7.1) as,

Note 3.7.3.   for all $AF_1' \sqsubseteq AF_1$, $AF_2' \sqsubseteq AF_2$, if $AF_3' = AF_1' + AF_2'$, then, $\mathcal{A}_3' = \mathcal{A}_1' \dot{\circ} \mathcal{A}_2'$ where $\mathcal{A}_1', \mathcal{A}_2', \mathcal{A}_3'$ denote the set of admissible sets in AF_1', AF_2', AF_3'.

We can test the above formulation of independence relation with respect to the disjoint sub-frameworks such that, for a framework AF,

> if AF is comprised of n disjoint sub-frameworks AF_i, then, any $AF' \sqsubseteq AF$ is also comprised of n number of disjoint frameworks $AF'_i \sqsubseteq AF_i$, and that, the set of admissible sets $\mathcal{A}'$ in AF' can be written as a product of the sets of admissible sets $\mathcal{A}'_i$ in AF'_i such that $\mathcal{A}' = \dot{\prod}_{1 \leq i \leq n} \mathcal{A}'_i$.

The above claim can be shown by a simple derivation form theorem 2.2.5. However, upon a closer inspection of theorem 2.2.5, we also realize that any two *closed* sub-framework that do not intersect also exhibit the same independency property as stated in the note 3.7.3. Hence, for two sub-frameworks to be considered acting independently, they need not be completely disconnected. All that is required is that they should not *influence* another.

We may then extrapolate that,

Note 3.7.4. for two independent sub-frameworks AF_1, AF_2, neither any argument in AF_2 can be an active argument for any argument in AF_1, nor, any argument in AF_1 can be an argument for any argument in AF_2.

Hence, following note (3.7.4) above a new reading of lemma 3.7.2 can be given that is based on the active argument relation. This new formulation is given as the first part of the following lemma, 3.7.5. There is, though, a difference between the two lemmas. This time, the focus is solely on the active argument relation with respect to the arguments in AF_{12}, AF_{21}, the non shared sub-frameworks, i.e. $AF_{12} = AF_1 -^N AF_2$, $AF_2 -^N AF_1$. Hence, we relax the constraint $\mathcal{A}_3 = \mathcal{A}_1 \dot{\circ} \mathcal{A}_2$ on AF_1, AF_2.

Furthermore, to stay true to our analysis so far, in order to correctly consider two sub-frameworks AF_{12}, AF_{21} independent we need to apply the criteria (3.7.4) to all the pairings of AF'_1, AF'_2 where $AF'_1 \sqsubseteq AF_{12}$ and $AF'_2 \sqsubseteq AF_{21}$, given that the shared sub-framework of AF_1, AF_2 remains fixed and is always in play. Indeed if we do so, then the converse of the first part of lemma 3.7.5 will also be true, as shown by the second part of lemma 3.7.5.

Lemma 3.7.5. *For two sub-frameworks $AF_1 = \langle AR_1, ATT_1 \rangle$, $AF_2 = \langle AR_2, ATT_2 \rangle$ of a*

framework AF, let $AF_3 = AF_1 +^N AF_2$, AF_{12}, AF_{21} denote $AF_{12} = AF_1 -^N AF_2$, $AF_{21} = AF_2 -^N AF_1$, and AF_4 be the shared sub-framework of AF_1, AF_2. For AF_1, AF_2 then let AR_{12}, AR_{21} be $AR_{12} = AR_1 - AR_2$, $AR_{21} = AR_2 - AR_1$, and, $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ each in turn denote the set of admissible sets in AF_1, AF_2, AF_3 and $\mathcal{A}_1^*, \mathcal{A}_2^*, \mathcal{A}_{13}^*, \mathcal{A}_{23}^*$ each be $\mathcal{A}_1^* = \{S \mid S = T \cap AR_{12}, T \in \mathcal{A}_1\}$, $\mathcal{A}_2^* = \{S \mid S = T \cap AR_{21}, T \in \mathcal{A}_2\}$, $\mathcal{A}_{13}^* = \{S \mid S = T \cap AR_{12}, T \in \mathcal{A}_3\}$, $\mathcal{A}_{23}^* = \{S \mid S = T \cap AR_{21}, T \in \mathcal{A}_3\}$.

1. *If no argument of AF_{12} is an active argument for any argument of AF_{21} in AF_3, and vice versa, then $\mathcal{A}_1^* = \mathcal{A}_{13}^*$ and $\mathcal{A}_2^* = \mathcal{A}_{23}^*$.*

2. *The conserve is true provided that the result $\mathcal{A}_1^* = \mathcal{A}_{13}^*$, $\mathcal{A}_2^* = \mathcal{A}_{23}^*$ holds with respect to all the sub-frameworks AF_1', AF_2' of AF_1, AF_2 where $AF_4 \sqsubseteq AF_1' \sqsubseteq AF_1$, $AF_4 \sqsubseteq AF_2' \sqsubseteq AF_2$, i.e. the shared sub-framework AF_4 remains fixed.*

One way to ensure that the condition (3.7.4) is met is by applying observation 3.3.2, and see whether or not all argument in AF_{12} are intercepted for any argument in AF_{21}, by some set of arguments in AF_4, and vice versa. In case this condition is met, we say the two sub-frameworks, e.g. AF_{12}, AF_{21}, are to be *disjointed by intercept*.

Definition 3.7.6. Two sub-argumentation framework AF_1, AF_2 of a framework AF are said to be *disjointed by intercept* in AF if and only if every argument in AF_1 is intercepted for every argument in AF_2 and vice versa. Moreover, let AF_3 be a non-intersecting sub-framework for one or both sub-frameworks AF_1, AF_2. AF_3 is said to *disjoint AF_1, AF_2 by intercept* if and only if every sequence of attack relations between any argument in AF_1 and AF_2 is intercepted by some set S of arguments in AF_3.

It is then, that we can apply observation 3.3.2 and obtain the initial intent that for two disjointed by intercept sub-frameworks, no argument in one is an active argument for another and vice versa, as stated by the following observation.

Observation 3.7.7. *Let AF_1, AF_2 be two sub-framework of AF. If AF_1, AF_2 are disjointed by*

intercept in AF, then, no argument of AF_1 is an active argument for any argument of AF_2 and vice versa.

In light of the above observation, all that is required to ensure that the criteria (3.7.4) is met, is to keep all the pairings of sub-frameworks $AF'_1 \sqsubseteq AF_1$, $AF'_2 \sqsubseteq AF_2$ disjointed by intercept. Hence, we give the final version of lemma 3.7.2 that is based on the *disjointed by intercept* sub-frameworks.

Lemma 3.7.8. *Let two sub-frameworks AF_1, AF_2 be as described in lemma (3.7.5). Then, the claims (3.7.5.1) and (3.7.5.2) hold for all the sub-frameworks AF'_1, AF'_2 of AF_1, AF_2 where $AF_4 \sqsubseteq AF'_1 \sqsubseteq AF_1$, $AF_4 \sqsubseteq AF'_2 \sqsubseteq AF_2$ if and only if AF_4 disjoints AF_{12} and AF_{21} by intercept in AF_3.*

To demonstrate the properties of the disjointed by intercept sub-frameworks we present the following example 3.7.9. The framework AF_{12} is split into three sub-frameworks $AF_{12a}, AF_{12b}, AF_{12c}$ where AF_{12b} disjoints AF_{12a}, AF_{12c} by intercept. We will be looking at the working of lemmas 3.7.2, 3.7.5, 3.7.8 with respect to the three sub-frameworks.

Example 3.7.9. The argumentation framework AF_{12}, below, is split into three sub-frameworks $AF_{12a}, AF_{12b}, AF_{12c}$, such that $AR_{12a} = \{e, a\}$, $AR_{12b} = \{c, b, d, f\}$, $AR_{12c} = \{h, g, p\}$, and, $AF_{12} = AF_{12a} +^N AF_{12b} +^N AF_{12c}$.

(i) The three sub-frameworks $AF_{12a}, AF_{12b}, AF_{12c}$ are obtained by removing the attack relations (a, b), (f, g) from ATT_{12}. The two attack relations (a, b), (f, g), each in turn correspond to the arguments b and g where the corresponding attack sequences $\pi_1 = (b, a, e)$ and $\pi_2 = (g, f, d, b, c)$ are intercepted by the set of arguments $S_1 = \{c\}, S_2 = \{h\}$.
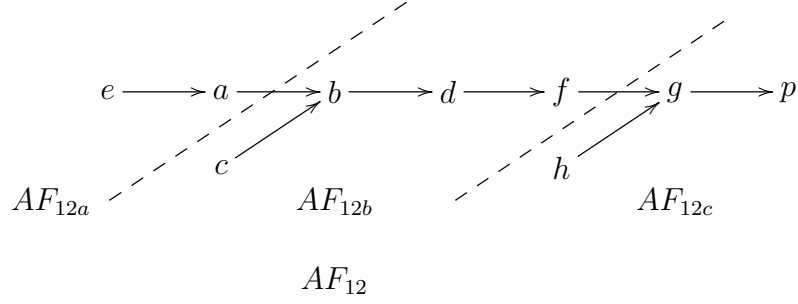
(ii) It is easy to check that the three sub-frameworks AF_{12a}, AF_{12b}, AF_{12c} are all disjointed by intercepts in regard to another where AF_{12b} disjoints AF_{12a}, AF_{12c} and AF_{12c} disjoints itself from AF_{12b}.

Next, to test the results in lemmas 3.7.2, 3.7.5, 3.7.8, we construct the sub-frameworks $AF_{12ab} = AF_{12a} +^N AF_{12b}$ and $AF_{12bc} = AF_{12b} +^N AF_{12c}$. The respective set of non-trivial admissible sets $\mathcal{A}$, $\mathcal{A}_{12a}$, $\mathcal{A}_{12b}$, $\mathcal{A}_{12c}$, $\mathcal{A}_{12ab}$, $\mathcal{A}_{12bc}$ for the sub-frameworks AF, AF_a, AF_b, AF_c, AF_{ab}, AF_{bc} are —

$$\mathcal{A} = \{\{e\}\} \mathbin{\dot\circ} \{\{c\}, \{c, d\}\} \mathbin{\dot\circ} \{\{h\}, \{h, p\}\},$$
$$\mathcal{A}_{12a} = \{\{e\}\},$$
$$\mathcal{A}_{12b} = \{\{c\}, \{c, d\}\},$$
$$\mathcal{A}_{12c} = \{\{h\}, \{h, p\}\},$$
$$\mathcal{A}_{12ab} = \{\{e\}\} \mathbin{\dot\circ} \{\{c\}, \{c, d\}\},$$
$$\mathcal{A}_{12bc} = \{\{c\}, \{c, d\}\} \mathbin{\dot\circ} \{\{h\}, \{h, p\}\}.$$

(iii) From the above values, we can easily see the claim of lemma 3.7.2 that, $\mathcal{A} = \mathcal{A}_{12ab} \mathbin{\dot\circ} \mathcal{A}_{12bc}$.

(iv) Next, we look at the claim of theorem 3.7.12 below. The claim is that since all sub-frameworks AF_{12a}, AF_{12b}, AF_{12c} are disjointed by intercept for another, the set of admissible sets $\mathcal{A}_{12}$ in AF_{12} is a simple product of $\mathcal{A}_{12a}$, $\mathcal{A}_{12b}$, $\mathcal{A}_{12c}$, i.e., $\mathcal{A}_{12} = \mathcal{A}_{12} \mathbin{\dot\circ} \mathcal{A}_{12b} \mathbin{\dot\circ} \mathcal{A}_{12c}$.

To verify the claims of lemmas 3.7.5, 3.7.8, the sub-frameworks $AF'_{12} \sqsubseteq AF_{12}$, $AF'_{12ab} \sqsubseteq AF_{12ab}$, $AF'_{12bc} \sqsubseteq AF_{12bc}$ are constructed such that $AF'_{12} = AF'_{12a} +^N AF_{12b} +^N AF'_{12c}$, $AF'_{12ab} = AF'_{12a} +^N AF_{12b}$, $AF'_{12bc} = AF_{12b} +^N AF'_{12c}$, where $AF'_{12a} \sqsubseteq AF_{12a}$, $AF'_{12c} \sqsubseteq$

AF_{12c}, and, the set of all admissible sets in AF'_{12}, AF'_{12ab}, AF'_{12bc} be denoted by $\mathcal{A}'_{12}$, $\mathcal{A}'_{12ab}$, $\mathcal{A}'_{12bc}$. In addition let $\mathcal{A}'^{*}_{12a}$, $\mathcal{A}'^{*}_{12c}$, $\mathcal{A}'^{*}_{12ab}$, $\mathcal{A}'^{*}_{12bc}$ denote $\mathcal{A}'^{*}_{12a} = \{S \mid S = T \cap AR_{12a}, T \in \mathcal{A}'_{12}\}$, $\mathcal{A}'^{*}_{12c} = \{S \mid S = T \cap AR_{12c}, T \in \mathcal{A}'_{12}\}$, $\mathcal{A}'^{*}_{12ab} = \{S \mid S = T \cap AR_{12a}, T \in \mathcal{A}'_{12ab}\}$, $\mathcal{A}'^{*}_{12bc} = \{S \mid S = T \cap AR_{12c}, T \in \mathcal{A}'_{12bc}\}$.

(v) We do not go through all sub-frameworks AF'_{12}, AF'_{12ab}, AF'_{12bc}, but, it is easy to check that for all such sub-frameworks, $\mathcal{A}'_{12} = \mathcal{A}'_{12a} \dot{\circ} \mathcal{A}'_{12b} \dot{\circ} \mathcal{A}'_{12c}$ holds, as presented in the lemmas 3.7.5, 3.7.8.

(vi) Next to the theorem 3.7.14, below, which is a generalization of lemma 3.7.8. The theorem presents a variation of the claim in lemma 3.7.8. The claim here is that for all, AF'_{12}, $\mathcal{A}'_{12} \subseteq \mathcal{A}'_{12ab} \dot{\circ} \mathcal{A}'_{12bc}$.

(vii) Finally, we can show case the result of theorem 3.7.12, below, which is also a generalization of lemma 3.7.8. The theorem states that since AF_{12a} is intercepted for both AF_{12b}, AF_{12c}. The admissibility of any set in any sub-framework of AF_{12a} does not affect the admissible sets in AF_{12b}, AF_{12c}. That is, in all sub-frameworks $AF''_{12} = AF'_{12a} +^{N} AF_{12bc}$, the set of admissible sets $\mathcal{A}''_{12}$ in AF''_{12} is a simple product of $\mathcal{A}'_{12a}$, $\mathcal{A}_{12bc}$, i.e., $\mathcal{A}''_{12} = \mathcal{A}'_{12a} \dot{\circ} \mathcal{A}_{12bc}$. ∎

In the above example, the framework AF is partitioned into disjointed sub-frameworks. This partitioning of AF has a number of distinct properties that are highlighted in the example. In the followings, we can generalize the splitting of AF_{12} into *disjointed by intercept sub-frameworks* and formulate the ensuing properties of such partitioning. First, we define the partitioning of a framework into the *disjointed by intercept sub-frameworks*.

Definition 3.7.10. Let $\mathcal{AF}$ be a countable set of sub-frameworks of a framework AF. We say $\mathcal{AF}$ to be a *partition* of AF into *disjointed by intercept sub-frameworks* if and only if $AF = \sum_{AF' \in \mathcal{AF}}^{N} AF'$, and, any two distinct AF', AF'' in $\mathcal{AF}$ are disjointed by intercept.

The following framework AF_{13} presents a schematic view of partitioning of a framework to disjointed by intercept sub-frameworks $AF_1, AF_2, \cdots$. The nodes here present the disjointed

sub-frameworks, and, the arrows show the intercepted attack sequences. Any two non-adjacent sub-framework is then intercepted by some in-between sub-frameworks. For instance, AF_4 is disjointed from AF_1 by sub-frameworks AF_2, AF_3.

$$AF_{13}$$
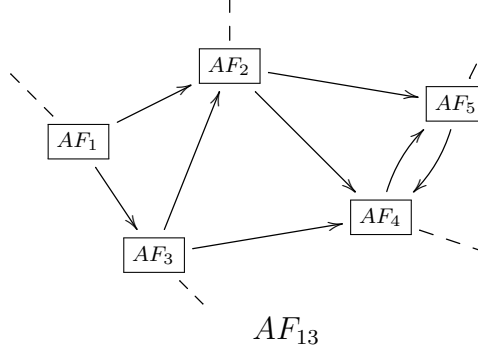
Obviously, any two sub-frameworks that are disjointed by intercept cannot be intersecting. Hence, any partition of a framework to disjointed by intercept sub-frameworks deals with the non-intersecting sub-frameworks. Due to this property, we can construct a partial order $\preceq$ over the class $\mathbb{AF}$ of all disjointing by intercept partitions of AF such that $\mathbb{AF}$ is a partial order with a maximum and a minimum element. The following lemma states these two properties of $\mathbb{AF}$.

Lemma 3.7.11. *Let $\mathbb{AF}$ be the class of all partitions of AF to disjointed by intercept sub-frameworks. The order $\preceq$ over $\mathbb{AF}$ is defined such that for $\mathcal{AF}_1, \mathcal{AF}_2 \in \mathbb{AF}$, $\mathcal{AF}_1 \preceq \mathcal{AF}_2$ if and only if for every $AF_1 \in \mathcal{AF}_1$ there is some $AF_2 \in \mathcal{AF}_2$ such that $AF_1 \sqsubseteq AF_2$.*

1. *For any $\mathcal{AF} \in \mathbb{AF}$, all sub-frameworks in $\mathcal{AF}$ are non-intersecting.*

2. *The order $\preceq$ over $\mathbb{AF}$ is a partial order with a maximum element.*

3. *Let $\mathbb{AF}^N$ be a subclass of $\mathbb{AF}$ such that for every $\mathcal{AF} \in \mathbb{AF}^N$ is a normal sub-framework of AF. $\mathbb{AF}^N$ then has a minimum element under $\preceq$.*

The section (iv) of example 3.7.9 illustrates that the admissibility of any admissible set in any of the sub-frameworks of a disjointed by intercept partition does not depend on another sub-

framework. Thus, the set of admissible sets of the whole framework is the product of the admissible sets of each sub-framework. The following theorem presents this finding.

Theorem 3.7.12. *Let $\mathcal{AF}$ be a partition of AF into disjointed by intercept sub-frameworks, then, $\mathcal{A} = \prod_{AF' \in \mathcal{AF}} \mathcal{A}'$ where $\mathcal{A}, \mathcal{A}'$ each is the set of admissible sets in AF, AF'.*

Following the results of example 3.7.9, we can see that AF_{12b} plays a distinct role in regard to a number of properties of the partitioning. All these properties stem from the fact that AF_{12b} is the *adjacent intercepting sub-framework* for AF_{12a} where it separates AF_{12a} from AF_{12c}. The following definition labels such sub-frameworks as the *adjacent intercepting sub-frameworks*.

Definition 3.7.13. For two sub-frameworks AF', AF'' of a framework AF, we say AF' is the *adjacent intercepting sub-framework* for AF'' in AF if and only if there is some attack sequence π from some argument in AF'' for which some set of arguments in AF' intercepts π, and, no set of arguments in a non-intersecting sub-framework with AF' intercepts any nontrivial subsequence of π.

The primary role of the *adjacent intercepting sub-frameworks AF^** is that they isolate the sub-frameworks AF' for which they are intercepting, from those that intercept them AF_0. As a result, the admissibility of sets in AF' is shielded against any changes in the admissibility of sets in AF_0. For instance, in framework AF_{13}, as long as AF_2, AF_3 are in play, the changes in AF_4 have no affect on the admissibility of any set in AF_1. The following theorem is formulates this property which is, in a manner, a generalization of lemma 3.7.8.

Theorem 3.7.14. *Let $\mathcal{AF}$ be a partition of AF into disjointed by intercept sub-frameworks, and, $AF_0 \in \mathcal{AF}$. For $\mathcal{AF}, AF_0$, let $\mathcal{AF}^* \subseteq \mathcal{AF}$ be the set of all sub-frameworks $AF' \in \mathcal{AF}^*$ that are the adjacent intercepting sub-framework for AF_0. For $\mathcal{AF}^*$, then, let $AF^* = \sum_{AF' \in \widehat{\mathcal{AF}^*}}^N AF'$, and, $\widehat{\mathcal{AF}} = \mathcal{AF} - (\mathcal{AF}^* \cup \{AF_0\})$, and, $\widehat{AF}$ denote the corresponding sub-framework $\widehat{AF} = \sum_{AF' \in \widehat{\mathcal{AF}^*}}^N AF'$ for $\widehat{AF}$. In addition, For every $AF'_0 \sqsubseteq AF_0$, the sub-framework $AF^\circ \sqsubseteq AF$ with*

respect AF_0' is $AF^\circ = AF^{*'} +^N \widehat{AF}$ where $AF^{*'} = AF^* +^N AF'$.

For every AF' and its corresponding AF°, $\widehat{AF}$, $AF^{*'}$, and, their corresponding set of all admissible sets, $\mathcal{A}^\circ$, $\widehat{\mathcal{A}}$, $\mathcal{A}^{*'}$, it then holds that $\mathcal{A}^\circ = \widehat{\mathcal{A}} \dot\circ \mathcal{A}^{*'}$.

One of the required conditions in the above theorem is that the isolated frameworks $\widehat{\mathcal{AF}}$ are kept unchanged. In section (vi) of example 3.7.9, this constraint is relaxed with respect to AF_{12a}, AF_{12c}. However, their separating sub-framework AF_{12b} is still kept fixed. Under this relaxed setting, the admissible sets of the *separated* sub-frameworks are no longer shielded against change. However, the admissibility of arguments in one framework is shielded against the changes in another framework. That is, an argument that is admissible in its sub-framework will be admissible in the final merged framework. The following theorem formulates this finding.

Theorem 3.7.15. *Let $\mathcal{AF}$ be a partition of AF into disjointed by intercept sub-frameworks. For $\mathcal{AF}$, let $\mathcal{AF}^* \subseteq \mathcal{AF}$ be such that any two distinct AF', AF'' in $\mathcal{AF}^*$ are disjointed by intercept by some set of arguments in $\widehat{AF}$ where $\widehat{AF} = \sum_{AF' \in \widehat{\mathcal{AF}}}^N AF'$ and $\widehat{\mathcal{AF}} = \mathcal{AF} - \mathcal{AF}^*$. For $\mathcal{AF}^*$, then, let $\mathcal{AF}^{*'}$ denote some arbitrary set of sub-frameworks $AF' \sqsubseteq^N AF$ such that for each AF' there is one and only one $AF'' \in \mathcal{AF}^*$ where $AF' \sqsubseteq AF''$. For $\mathcal{AF}'$, then let AF^* be the sub-framework formed from all $AF' \in \mathcal{AF}^{*'}$ where $AF^* = \left(\sum_{AF' \in \mathcal{AF}'}^N AF' \right) +^N \widehat{AF}$. For every AF' and its corresponding AF', AF^*, $\widehat{AF}$, and, their corresponding set of all admissible sets, $\mathcal{A}'$, $\mathcal{A}^*$, $\widehat{\mathcal{A}}$ it then holds that $\mathcal{A}^* \subseteq \left(\prod_{AF' \in \mathcal{AF}'} \mathcal{A}' \right) \circ \widehat{\mathcal{A}}$.*

In section (i) of example 3.7.9, the disjointed sub-framework AF_{12a}, AF_{12b}, AF_{12c} are formed by removing the links, i.e., the attack relations, in the intercepted attack sequences at the nodes where the intercept occur. It is easy to see that if we remove all such attack relations from a framework, we will then not have any intercepted path in the resulting sub-framework.

Removing of these attack relations may or may not result in more than one disjointed sub-framework. Regardless, all the resulting sub-frameworks,

Note 3.7.16.
1. are disjointed by intercept,
2. are comprised of only the active attack sequences, and therefore, cannot be further divided into more disjointed sub-frameworks,
3. and, partition the framework into maximum (with respect to number) disjointed by intercept sub-frameworks.

In the followings, we formulate the construction of these sub-frameworks and show that the properties (1)-(4) above hold for them. We first begin by defining the sub-frameworks that cannot be divided into disjointed by intercept sub-frameworks. We call such sub-frameworks *biased* frameworks, because, all the defenders (respectively attackers) of the arguments are *by large* either all admissible or all dismissible.

Definition 3.7.17. An argumentation framework AF is said to be *biased* if and only if all attack sequences in AF are active.

Next, we show the relation between the second and the third properties in (3.7.16). In lemma 3.7.11, we defined the order $\preceq$ over $\mathbb{AF}$, the class of partitions of AF into disjointed by intercept sub-frameworks. It is easy to see that under $\preceq$, the minimal elements of $\preceq$ are the maximal elements of $\mathbb{AF}$ in terms of size of a set. The following theorem shows, if there is a partition $\mathcal{AF}$ of AF, as is obtained in example 3.7.9, then, $\mathcal{AF}$ is the maximum, in terms of number of elements, possible partition of AF into disjointed by intercept sub-frameworks that we can have.

Theorem 3.7.18. *Let $\mathbb{AF}$ be the class of partitions of AF into disjointed by intercept sub-frameworks, and, $\preceq$ be the order defined over $\mathbb{AF}$ in lemma 3.7.11.*

1. *If $\mathcal{AF}$ is the minimal element of $\mathbb{AF}$ under $\preceq$ then every $AF' \in \mathcal{AF}$ is biased.*

2. *There is a unique $\mathcal{AF} \in \mathbb{AF}$ such that all $AF' \in \mathcal{AF}$ are biased and for all $\mathcal{AF}' \in \mathbb{AF}$, $\mathcal{AF}' \succ \mathcal{AF}$, there is some $AF'' \in \mathcal{AF}'$ that is not biased. Furthermore, $\mathcal{AF}$ is equal in size with any minimal element of $\mathbb{AF}$ under $\preceq$.*

Finally, we can deal with construction of the partitioning of AF that meets the desired properties, presented in theorem 3.7.18. The following theorem states that if we follow the same

method of partitioning of a framework as is carried in example 3.7.9, we will have a partition of AF that holds all the properties described in (3.7.16) above.

Theorem 3.7.19. *For and argumentation framework $AF = \langle\, AR\,,\, ATT\,\rangle$, let $AF^* \sqsubseteq AF$ be constructed such that for all attack sequences $\pi = (a, b)$ where π is intercepted at a, the attack relation (b, a) is removed from ATT. The set $\mathcal{AF}$ of all disjointed sub-frameworks of AF^* then partitions AF into disjointed by intercept sub-frameworks such that every $AF' \in \mathcal{AF}$ is a biased sub-framework of AF.*

This concludes our discussion on the role of intercepts in partitioning a framework into *disjointed* sub-frameworks. Obviously, there can be many ways to split a framework into subframeworks with distinct properties. But, we leave that discussion as future work.

3.8   Summary

In this chapter we utilized the relationship between the admissibility backings of an argument and the admissibility backings of its attackers so to address three important lines of inquiries within argumentation theory. The three lines of inquiries are

- the relevance of an argument in regard to the admissibility of other arguments;

- the propagation of backings along the attack sequences; and,

- how to split a framework into sub-frameworks of distinct characteristics.

In this regard, we defined the *active attack sequences* and the *intercepts*. The active attack sequences mark the lines on which the admissibility backings propagate. The intercepts correspond to where the propagation of backings along the attack sequences halts. Hence, the intercepts show the arguments arguments that are made irrelevant for the admissibility of some argument, and, how they are made irrelevant.

Other presented arguments relations are the active argument relation, the critical argument re-

lation, the incompatible argument relation and the redundant argument relation. The presented argument relations are in one way or another related to each other. Accordingly, these argument relations are shown to be sufficient for determining whether or not admissibility backings propagate along the attack sequences.

Not all the arguments on an active attack sequence of an argument play a part for the admissibility of that argument. Such arguments are identified by the *incompatible* and *redundant* argument relations. The identification of incompatible and redundant arguments are done with the help of *critical arguments*. A critical argument for an argument is an argument that is indispensable for the admissibility of that argument. Hence, an argument that is critical for another argument can be neither incompatible or redundant for that argument.

The intercepts play a most distinct role in that they split an argumentation framework into *independent* sub-frameworks. We marked these independent sub-frameworks as *disjointed by intercept* sub-frameworks. Accordingly, the independence is meant that any change including the addition of new arguments that happens to one sub-framework does not affect the other disjointed frameworks.

As far as we know, there are no other research that address the same issues presented in this chapter. Hence, we cannot site any related research. All the related research are in relation to the admissibility backings of arguments that are discussed in chapter two. A list of possible future work is discussed in the conclusion chapter of this dissertation. Hence, we conclude chapter three.

Chapter 4

Context Sensitive Defeasible Rules

4.1 Introduction

Defeasible argumentation systems are used to model commonsense and defeasible reasoning. Current argumentation systems assume that an argument that appears to be justified also satisfies our expectation in relation to the correct outcome, and, vice versa. In this chapter we present an alternative representation of defeasible rules that adheres more to this assumption. The proposed inference rules are called *context sensitive* rules. The context sensitive defeasible rules are tailored for argumentation based reasoning. Effectively, we assume that a mechanism exists that given an arbitrary inference rule, the mechanism tells us whether in a given situation the rule is applicable. This mechanism is usually presented in terms of an abnormality condition for the rule. Accordingly, we provide a mapping between our argumentation system and Dung's abstract argumentation theory to show the efficacy of the presented argumentation system.

The defeasibility of reasoning is captured in different ways in different frameworks. Roughly, in the framework of default logic it is captured by assuming that the rules are defeasible, allowing for alternative extensions depending on which set of defaults get activated. [1] Furthermore, in case of conflicting rules, often rules of thumb such as *specificity* are used to break the tie

---

[1] Informally, a default rule is of the form: If A is known as a matter of fact, and B can be assumed without courting inconsistency, then C may be inferred. Thus, given a knowledge base, the rule itself tells, as it were, whether or not it can be "fired".

[Poo91, SL92, DS01]. [2] In circumscription, it is achieved by minimizing the extension of *abnormal* predicates. In the case of argumentation, it is achieved by allowing some arguments to *defeat* other arguments [McC86].

Most of the proposed inference rules are however context independent, in the sense that the condition that makes a rule applicable is context independent. In our representation of defeasible rules, we refer to the context independent *abnormality* conditions, the *conclusive defeaters* of a rule.

This chapter is structured such that in the next section we present our motivation for the proposed context sensitive rules, accompanied by two running examples. We continue by elaborating on the function of the proposed rules. We then develop an argumentation theory, in the usual way, that is based on the attack and reinstatement relation between arguments. We then follow to discuss the semantics of the formulated argumentation system. The semantics is provided by means of a translation from the proposed system to Dung's abstract argumentation framework. As part of our discussion, we show that when only the *conclusive defeaters* are in play, the proposed argumentation system produces the same outcome as systems based on conventional default rules.

4.2 Motivation

From the outset we assume a propositional language $\mathcal{L}$ composed of countably many literals (both positive and negative) and a set of defeasible inference rules $\mathcal{R}$. Technically a rule is a relation between a set of literals called premises and another literal called a conclusion.

Notation: An inference rule, d, is represented as $d : a_1, a_2, ..., a_n \rightarrow a$ where $a_1, a_2, ..., a_n, a \in \mathcal{L}$. We call $bd(d) = \{a_1, a_2, ..., a_n\}$ the *body* of the rule, and,

---

[2] *Specificity* reads as given two rules, applicable to a given context, the one that makes use of more specific information takes precedence over the other. Thus, if we know Tweety is a penguin, and given the rules that birds in general fly and penguins don't, we should conclude that Tweety does not fly.

$hd(d) = a$ its *head*.

An argument is usually defined as a sequence of inferences from known premises (the *contingent knowledge*) to a conclusion. Alternatively it is represented as an inference-tree-structure embedded in the premises, or, as a pair of premises and conclusion. We use all these three representations dictated by convenience.

We note that unlike the truth-based classical logical systems, argumentation systems are founded upon justification [Nut01]. An argument is accepted in the absence of a justified counter argument. The counter argument against an argument is called the attacker [Dun95b] or the defeater [Pol87] of an argument. We will be using the terms *defeat* and *attack* pretty much interchangeably. We say that a defeated argument is *reinstated* if its defeater gets defeated by an accepted argument.

Furthermore, we use the *Rebuttal* and *Undercutting* defeats to model the defeat relations [Pol94, PV01]. In argumentation systems, the *accepted (justified)* arguments that automatically assumed to adhere to two general constraints.

1. An acceptable argument (including its conclusion) should meet our expectation in regard to the available information,

2. An acceptable argument should be justified within the logic of the corresponding argumentation system.

It is obvious that the two constraints are imperative to any argumentation system. A defeasible inference rule is *well crafted* only if it preserves and conveys the relation that its antecedent is a reason for believing its consequent. In this regard, we present two motivating examples that show case, there are cases where "well crafted" defeasible inference rules cannot always accommodate the requirement set by the two constraints. That is, we either have to abandon one or the other. Otherwise, we need to appeal to a more flexible presentation of defeat relation.

The new presentation, while still being based on the *rebutting* and *undercutting* defeats, it also shows how the two forms of defeats can get toggled by the context.

In the first example we advocate the need to expand the notion of reinstatement of argument to allow arguments to be reinstated without defeating the defeater. In the second example we argue the need to allow asymmetrical-provisional-defeat relationship. We use the result of these examples to introduce the representation of defeasible rules that we advocate. We argue that our representation while simple, provides a more explanatory model of "pragmatic reasoning". Moreover, the simplicity of the rules allows an argumentation system to meet one of the main objectives of argumentation reasoning, namely to provide an explanation in line with human reasoning [RG01, KR04, RW06, Ver01a].

We note in passing that since we have not yet introduced our own definition of an argument, in the following examples, we represent arguments as sequences of inferences from premises, standard in the literature. We also depict arguments as triangles where the base represent premises and top vertex conclusion of the argument. The attack relationship is shown by an arrow form attacking to the attacked argument.

Example 4.2.1. A physiologist is studying a system that involves the secretion of hormones and enzymes in presence of other hormones and enzymes. Let us assume that states of the system are all describable in terms of atoms a, b, c, v, x, y, z where a, b, c mnemonically stands for, *the enzyme A, B, C is present*, and, the atoms v, x, y, z stand for *the hormone V, X, Y, Z is present*.

Tables (4.1.a), (4.1.b) represent the results of careful experimentation for two alternative scenarios. Each table has two sections, *Known Facts* and *Can be Believed*. For instance, the row one in table (4.1.a) states that *if all we know is that the enzyme A is present then we are allowed to believe that hormone Z is present, as well*. Furthermore, the only difference between the two scenarios is that in scenario (a) all hormones are detectable whereas in scenario (b) the hormone Y is not detectable. Our question then is *can we model both these scenarios in terms*

Known Facts					Can be Believed				
a					a	z			
a	y				a	y			
b					b	y			
a	b				a	b	y		
a	b	x			a	b	x	z	
a	b	c			a	b	c	x	z
.	.	.	.	.	.	.	.	.	.

(a) y: detectable

Known Facts					Can be Believed				
a					a	z			
b					b				
a	b				a	b			
a	b	x			a	b	x	z	
a	b	c			a	b	c	x	z
.	.	.	.	.	.	.	.	.	.

(b) y: not detectable

Visualization of tables (4.1.a), (4.1.b) in terms of defeasible rules			
Table	Name	Rule	Defeater
Table 1(a)	d_1	$a \rightarrow z$	y
	d_2	$b \rightarrow y$	x
	d_3	$c \rightarrow x$	
Table 1(b)	d_1'	$a \rightarrow z$	b
	d_2'		
	d_3'	$c \rightarrow x$	

(c) The inferenc rules for (a) and (b)

Table 4.1: The results for the relation between hormons and enzymes in example (4.2.1)

of an argumentation system?

Furthermore, the knowledge that *if all we know is that enzyme A is present then we are allowed to believe hormone Z is present too* is interpreted as *presence of enzyme A is the primary explanation for secretion of hormone Z* (Enzymes generally act as catalysts), and is represented in terms of defeasible inference rule $a \rightarrow z$. The second row of table 1(a) is interpreted as the presence of hormone Y acts as a suppressant for secretion of hormone Z. In argumentation terms, the presence of hormone Y is therefore interpreted as an undercutting defeater for the reasoning, from a and $a \rightarrow z$ to z.

The result of construction of such defeasible inference rules and their associated undercutting defeaters is given in table (4.1.c). It can be shown that the observed system, as expressed in Table 1(a) can be modeled in terms of an argumentation system using the rules in table (4.1.c).

An argumentation system carries as follows.

1. An argument is justified if it has no defeater, or, all its defeaters are defeated by justified arguments.

2. An argument that has a justified defeater is overruled.

3. Only the conclusions supported by justified arguments are justified.

In the second scenario we assume that the hormone Y is not detectable. The result of experimentation for this scenario is shown in table (4.1.b). Base on table (4.1.b), the presence of enzyme B now acts as the undercutting defeater for the rule $a \to z$. The construction of defeasible rules for table (4.1.b) is given in table (4.1.c). We would now like to ask the same central question *could we still model this system in terms of an argumentation system*?

The answer this time is far from obvious. For instance, if the contingent knowledge is $\{a, b, c\}$, i.e. the enzymes A, B, C are known to be present, are we allowed to believe in presence of hormone Z?

Let an argument be represented as a sequence $\langle s_1, s_2, ..., s_n \rangle$ of statements where the last statement, s_n is the conclusion. The sequence $\langle s_n \rangle$ is then an argument with an empty set of premises representing a single fact.

The arguments in relation to the secretion of hormone Z are as follows.

1. *enzyme A is present* is denoted as $arg_0 = \langle a \rangle$.

2. "enzyme A is present and since enzyme A is the reason for secretion of hormone Z, so hormone Z is present" is denoted as $arg_1 = \langle a, \ a \to z, \ z \rangle$.

3. *enzyme B is present* is denoted as $arg_2 = \langle b \rangle$.

4. "enzyme C is present and since enzyme C is the reason for secretion of hormone X, so hormone X is present" is denoted as $arg_3 = \langle c, \ c \to x, \ x \rangle$.

The only attack relationship is arg_2 undercutting arg_1 i.e. presence of enzyme B undercuts the

reasoning $a \to z$. Therefore, arg_1 is defeated by arg_2, figure (4.1). Furthermore, since arg_2 has no defeater arg_1 stays defeated (note: arg_2 is a fact so it cannot have any defeaters). Yet, table (4.1.b) indicates that despite the presence of enzyme B, if enzyme C is present we are allowed to believe that enzyme A results in secretion of hormone Z.

As it can be seen in figure (4.1), the problem lies in arg_3 being unable to reinstate arg_4. The only way to reinstate arg_4 is to defeat arg_2. But, as it is already noted arg_2 is in essence an observation and cannot be defeated. It is as if there is a missing argument arg_4 as shown in figure (4.1) by dotted lines (arg_4 can be constructed from table (4.1.a)) where arg_4 attacks arg_1 and arg_3 reinstates arg_1 against this attack.
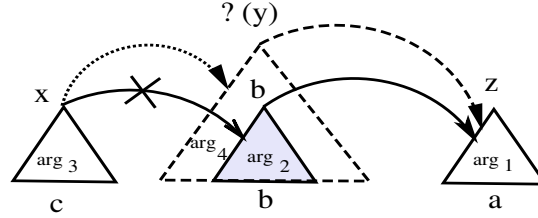

Figure 4.1: Arguments interaction for the scenario ($\{a, b, c\}$ is given facts) in example 1

To fix this problem we could introduce an new defeasible rule $(a, c) \to z$ to independently derive z. However, $(a, c) \to z$ is an artificial construct. The explanation for belief in z lies only in a. Hence, unless we introduce this rather artificial rule, our argumentation system will fall short of matching the real system. To short, we would like to allow arguments to be reinstated through context (not necessarily by attacking the attacker). We will also use the idea of missing arguments in our translation to Dung's Argumentation framework. ∎

Our next example is a common example presented in [Pol87]. The idea is, under normal circumstances, an object that appears red can regarded red. However, if one learns that object is seen under the red lighting, then, one an no longer argue that the object is red. The information "red lighting" is the regarded as an undercutting defeater against the argument for *the object being red*.

The following example shows that this situation can be correctly formulated such that the argument for *the object being red* is only provisionally defeated. That is to say an intended undercutting attack may result in a provisional defeat.

Example 4.2.2. In this example we argue that the common representation of this scenario in terms of an argumentation system, figure (4.2.a), does not yield the expected outcome. However, an alternative representation, figure (4.2.b), gives the expected outcome. We take the contingent knowledge and rule base to be $\{A_{red}, L_{red}\}$ and $\{A_{red} \rightarrow I_{red}\}$ where A stands for *appears*, I for *is* and L for *lighting*.
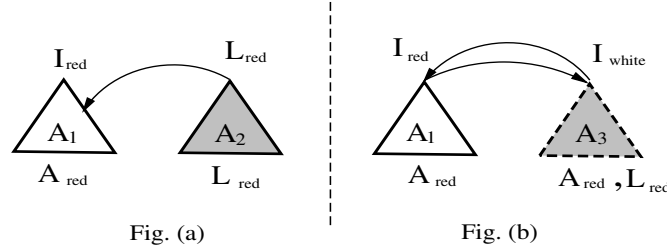

Figure 4.2: Argumentation representation of the two alternative scenarios in example (4.2.2)

The argument for *object is red* in both depiction of this scenario, figures (4.2.a), and (4.2.b), is $A_1 = \langle A_{red}, A_{red} \rightarrow I_{red}, I_{red} \rangle$. In the first representation, figure (4.2.a), the argument A_1 is undercut by the argument $A_2 = \langle L_{red} \rangle$, and, consequently I_{red} gets status *overruled* [Pol87]. Yet, the expected answer is that the statement I_{red} is *defensible* as object is either red or white (the status *defensible* is also referred to as *provisionally defeated*).

In an alternative representation where rule $(A_{red} \wedge L_{red}) \rightarrow I_{white}$ (or alternatively $(A_{red} \wedge L_{red}) \rightarrow \neg I_{red}$) is added to the rule base, the resulting argument interaction, figure (4.2.b), gives the expected outcome, i.e I_{red} is defensible. In figure (4.2.b), the two arguments A_1 and $A_3 = \langle A_{red}, L_{red}, (A_{red} \wedge L_{red}) \rightarrow I_{white}, I_{white} \rangle$ rebut each other leading to both arguments being *provisionally defeated*. The argument for *object being white* in figure (4.2.b), shown by dotted lines, could again be viewed as a missing argument in the first depiction of this scenario. We therefore would like to allow undercutting attacks (or asymmetrical attacks in general) result

in provisional defeat. ∎

We take an inference rule to be like a black box with some *underlying explanation* similar to the notion of *conveyance* given in [KR04]. [3] In every rule, antecedents are considered as the primary reason for belief in the consequent. In addition, there are other ancillary reasons that either strengthen or weaken a given rule.

In relation to how a rule works a justification function is provided that describes the conditions under which a rule gets activated. The justification function maps a given context (represented as a set of literals) into the operability space $\{0, 1/2, 1\}$, the values in question signaling, respectively, whether the rule is acceptable right away, is provisionally defeated, or is outright defeated.

As formal theories of argumentation get matured there is a growing interest to adapt these theories for modeling various forms of human reasoning. We believe our approach is in line with this goal. One approach is by the modeling of the natural language argumentation schemes [RG01, KR04, RW06]. This approach involves,

1. The characterization and classification of stereotypical patterns of reasoning is theorized in form of argumentation schemes [RW06].

2. The translation of formal arguments into natural language dialectic arguments [RG01].

In order to adapt formal argumentation theory to model argumentation schemes, the proposed approaches extend the current theory [AC02, WMP05]. Amgoud and Cayrol [AC02] propose a preference based argumentation framework that augments preferences among premises with the attack relationship in Dung's framework, while Wooldridge *et al.* [WMP05] propose a hierarchical metalogical argumentation framework. Our approach, too, is an extension of current argumentation theory with a central theme that each inference rule should have an underlying

---

[3] A black-box view of inference rule allows for an element of intentionality in the rules [AP97] as longs as there is an underlying explanation or a notion of conveyance.

explanation. Explanations are a supposition in argumentation schemes [KR04] and a requirement for translation into natural language arguments. Argumentation schemes are a top down approach, while our approach is a bottom up approach.

4.3 Defeasible reasoning system

By a defeasible reasoning system we mean a pair $\langle \mathcal{L}, \mathcal{R} \rangle$ where $\mathcal{L}$ is formal language, and $\mathcal{R}$ is a set of inference rules [BDKT97].

In the last section we stated that if $p \to q$ is an inference rule then our belief in p is the primary reason for our belief in q and the circumstances affecting our belief in applicability of $p \to q$ are the ancillary reasons.

Consider a rule d. If a context/circumstance do not affect d then d is applicable by default. However, if the circumstance affects the applicability of d then we should determine its effect. We represent each circumstance C_i as a set of literals. Suppose $C_1, C_2, ..., C_n$ are the circumstances that affect applicability of d. Now, not all literals in a circumstance C_i affect the applicability of d. Let for each circumstance C_i, its subset $J_i \subseteq C_i$, be the set of literals that affects rule d. Then $J = \bigcup_{i=1}^{n} J_i$ is the set of all literals that affects the applicability of rule d. We would therefore like to define a justification function for defeasible rule d by partitioning the space 2^J into three equivalence classes (equivalence w.r.t. the degree of *acceptability* of the rule d). The degree of *acceptability* of the rule d is represented by values in the operability space $\{0,\ 1/2,\ 1\}$.

4.3.1 Defeasible inference rule

Accordingly, we assume every rule d is associated with three families of $\mathbf{T}_d$, $\mathbf{U}_d$, and $\mathbf{F}_d$ of sets of literals such that,

1. $\mathbf{T}_d$, $\mathbf{U}_d$, and $\mathbf{F}_d$ partition a set 2^{J_d} where $\mathbf{J}_d$ is called the *justification domain* of d.

The assumption says that the three families, $\mathbf{T}_d, \mathbf{U}_d$ and $\mathbf{F}_d$ jointly exhaust all possible *observable states* that determine applicability of rule d.

2. x and $\neg x$ are not both in $\mathbf{J}_d$.

3. If $a \in bd(d)$, then $a \notin \mathbf{J}_d$. The antecedent of a rule must already be believed for a rule to be fired. Therefore, there is no special need to have the antecedent in $\mathbf{J}_d$.

Definition 4.3.1. Let d be a rule with its three associated sets, $\mathbf{T}_d, \mathbf{U}_d, \mathbf{F}_d$, given above. The *justification function* of a rule d is a function $\mathbf{H}_d \colon 2^{\mathcal{L}} \longrightarrow \{0, 1/2, 1\}$ where [4]

$$
\mathbf{H}_d(X) = \begin{cases} 1 & \text{if} \quad \mathbf{R}_d(X) = \varnothing \text{ or } \mathbf{R}_d(X) \in \mathbf{T}_d \\ 0 & \text{if} \quad \mathbf{R}_d(X) \in \mathbf{F}_d \\ 1/2 & \text{if} \quad \mathbf{R}_d(X) \in \mathbf{U}_d \end{cases}
$$

and the *relevance factor* of X w.r.t. the rule d, denoted $R_d(X)$, is the largest subset of X that is also a member of $(\underline{\mathbf{T}}_d \cup \mathbf{U}_d \cup \underline{\mathbf{F}}_d)$ i.e. $R_d(X) = X \cap \mathbf{J}_d$.

The parameter X is intended to represent a circumstance. A tabular representation of justification function is called *justification matrix* of the rule. It can be seen that if a context has no relevance to the applicability of a rule then the rule is applicable. In other words, if $R_d(X) = \varnothing$ then $\mathbf{H}_d(X) = 1$.

We next define what it means to say whether a rule is accepted or defeated, as well as, classifying rules based on the justification function.

Definition 4.3.2. Let $A \subseteq \mathcal{L}$ be a set of sentences, and $d \in \mathcal{R}$ a rule.

1. A is said to

   (a) *accept* d if and only if $\mathbf{H}_d(A) = 1$.

   (b) *outright defeat* d if and only if $\mathbf{H}_d(A) = 0$.

---

[4] While determining $\mathbf{T}_d$, $\mathbf{U}_d$, and $\mathbf{F}_d$ require some effort, it is no more arduous than assigning strength to arguments. However, in case of *Justification Function* there is the advantage of having a point of reference, i.e. the *circumstance*.

(c) *provisionally defeat* d if and only if $\mathbf{H}_d(A) = 1/2$.

(d) *conclusively defeat* d if and only if $\mathbf{H}_d(A) = 0$ and if $A \subseteq B$ then $\mathbf{H}_d(B) = 0$.

2. A rule $d : a_1, a_2, ..., a_n \rightarrow a$ in $\mathcal{R}$ is a *normal* rule if and only if $\{\neg a\}$ conclusively defeats d.

3. A normal rule $d \in \mathcal{R}$ is a *default* rule if and only if $\mathbf{U}_d = \varnothing$, and, every $A \in \mathbf{F}_d$ conclusively defeats d. In addition, we call $\mathbf{F}_d^k = \{A \mid A \text{ is the minimal set in } \mathbf{F}_d\}$, the *justification base* of the default rule.

4. A rule $d \in \mathcal{R}$ is said to be *indefeasible* if and only if $\mathbf{F}_d = \mathbf{T}_d = \mathbf{U}_d = \varnothing$.

The terms outright-defeat and provisional-defeat are adopted from [Pol94]. An indefeasible rule d is always acceptable, i.e. $\mathbf{H}_d(X) = 1$ is always true. This makes indefeasible rules synonymous to the material conditionals or the necessary knowledge in other argumentation systems [DS01, SL92].

The *conclusive defeat* relation is an important property that is implicitly assumed in other argumentation systems. In a conclusive defeat the defeat condition is context independent viz. a rule is always inapplicable in presence of a conclusive defeater.

We use the conclusive defeat relation to define normal rules. A normal rule is always defeated in light of contrary evidence to its conclusion. Normal rules implicitly capture the rebuttal attacks. They also ensure that no two arguments with contradictory conclusions are simultaneously justified. The above default rules are a special class of normal rules. In default rules all the defeat conditions are conclusive defeat conditions. It can be argued that within our semantics default rules would be equivalent to Reiter's default rules.

Theorem 4.3.3. *Given any set of sentences $A \subseteq \mathcal{L}$ and rule $d \in \mathcal{R}$,*

1. $\mathbf{H}_d(A)$ *has one and only one value.*

2. *A normal rule d with $\mathbf{U}_d = \varnothing$ is a default rule if and only if no A in $\mathbf{F}_d$ is a subset of B*

in $\mathbf{T}_d$.

3. *If d is a default rule and $\mathbf{F}_d^k$ is its justification base then,*

   (a) *for every $A \in \mathbf{F}_d^k$, A is either a singleton or $A = \{a \mid a \in B \text{ and } B \in \mathbf{T}_d\}$.*

   (b) *for $X \subseteq \mathcal{L}$, $\mathbf{H}_d(X) = 0$ if and only if $\exists A \in \mathbf{F}_d^k$ such that $A \subseteq X$.*

The theorem 4.3.3 is the first step to ensure that the proposed argumentation theory is well defined. The results in theorem 4.3.3.2, 4.3.3.3 draw a parallel between default rules (as defined in this work) and the conventional default rules e.g. Reiter default rules. Theorem 4.3.3.3 states that if a context includes any member of $\mathbf{F}_d^k$ then rule d is in applicable. Hence, one can say that members of $\mathbf{F}_d^k$ are similar to negation of grounded justification assumptions in Reiter default rules. For instance, if $d : b \to f$ and $\mathbf{F}_d^k = \{\neg f, p, e\}$, then the equivalent Reiter default rule is $d^r = \frac{b : f, \neg p, \neg e}{f}$, as shown by the followig examples.

Example 4.3.4. (the standard example in non-monotonic reasoning) *A Bird can usually fly unless it is a penguin or an emu.* Let literals b, f, p, e mnemonically stand for Tweety is a bird, can fly, is a penguin, is an emu. The example can be represented as a defeasible rule $d : b \to f$ is a default rule with the following justification function.

$$\mathbf{T}_d = \mathbf{U}_d = \varnothing$$
$$\mathbf{F}_d = \{\{p\}, \{e\}, \{\neg f\}, \{p, e\}, \{p, \neg f\}, \{e, \neg f\}, \{p, e, \neg f\}\}$$
$$\mathbf{F}_d^k = \{\{p\}, \{e\}, \{\neg f\}\}.$$

∎

Example 4.3.5. Sam's friends usually like ethnic foods unless they are hot and spicy. Though, Thai green curry is hot and spicy, they still like it. Let the literals sf, ef, hs, tg mnemonically represent Sam's friends, like ethnic food, food is hot and spicy, and food is Thai green curry. *Sam's friends usually like ethnic foods* could therefore be represented as, $d : sf \to ef$ with the justification function,

$$\mathbf{U}_d = \varnothing \quad \mathbf{T}_d = \{\{tg\}, \{hs, tg\}\}$$
$$\mathbf{F}_d = \{\{hs\}, \{\neg ef\}, \{\neg ef, tg\}, \{\neg ef, tg, hs\}\}.$$

∎

Example 4.3.6. The behavior of the table (4.1.b), to the extent it is specified, can be captured by assuming two inference rules d_1, d_2 where $d_1 : a \to z$ is a normal rule, and, $d_2 : c \to x$ is a default rule.

$d_1 : a \to z$		
$\mathbf{T}_{d1}$	$\mathbf{U}_{d1}$	$\mathbf{F}_{d1}$
(1)	(1/2)	(0)
x		b
x, b		$\neg z$
		$\neg z, b$
		$\neg z, x$
		$\neg z, x, b$

$d_2 : c \to x$		
$\mathbf{T}_{d2}$	$\mathbf{U}_{d2}$	$\mathbf{F}_{d2}$
(1)	(1/2)	(0)
		v
		$\neg x$
		$v, \neg x$

Table 4.2: The justification matrices of inference rules for table (4.1.b)

The respective justification functions, see table (4.2), are determined by,

$$\mathbf{T}_{d1} = \{\{x\}, \{x, b\}\}, \mathbf{U}_{d1} = \varnothing, \mathbf{F}_{d1} = \{\{b\}, \{\neg z\}, \{\neg z, b\}, \{\neg zx\}, \{\neg z, x, b\}\}.$$
$$\mathbf{T}_{d2} = \varnothing, \ \mathbf{U}_{d2} = \varnothing, \ \mathbf{F}_{d2} = \{\{v\}, \{\neg x\}, \{v, \neg x\}\}.$$

∎

4.3.2 Context sensitive arguments

We define an argument by a set of contingent facts, a set of inference rules and conclusion of the argument. This definition of an argument is in line with those given in [SL92] and [PV01]. Though, unlike the latter we do not include the length and the size of an argument as its properties. The size of an argument might indicate its strength [Lou87], but, our position is that information regarding strength of an argument should be within its inference rules.

Definition 4.3.7. Let $D \subseteq \mathcal{R}$, $A \subseteq \mathcal{L}$ and $a \in \mathcal{L}$, an argument Å is a tuple $\langle A, D, a \rangle$ such that there exists a sequence of rules $d_1, \ldots, d_m \in D$ where,

1. $a = hd(d_m)$, and,

2. $\forall d_{i, \, 0 < i \leq m}$, either

   (a) $bd(d_i) = \emptyset$, or

(b) $\forall a_j \in bd(d_i)$, either $a_j \in A$ or there exists $d_{k,\,0<k<j}$ such that $a_j = hd(d_k)$.

3. No proper subsequence of $A' \subset A$ and $D' \subset D$ satisfy the two conditions above.

We denote A, D and a by $A_{\text{Å}}, D_{\text{Å}}, a_{\text{Å}}$, and, call $A_{\text{Å}}$ the evidence and $a_{\text{Å}}$ the conclusion of the argument. Furthermore, we say an argument $\text{Å}_2 = \langle A_2, D_2, a_2 \rangle$ is a *subargument* of $\text{Å}_1 = \langle A_1, D_1, a_1 \rangle$, denoted by $\text{Å}_2 \sqsubseteq \text{Å}_1$, if and only if $D_2 \subseteq D_1$.

In our system, arguments interact with other arguments indirectly through context. Arguments create the context in which other arguments are accepted or rejected. The natural contribution of an argument to a context is its conclusion. On the other hand, when we accept an argument we implicitly accept all its subarguments. Therefore, in a set of arguments, the effective contribution of an argument to context is the conclusions of all its subarguments.

The set $Cn(\text{Å}) = \{x \in \mathcal{L} \mid x = a_{\text{Å}_i}, \text{Å}_i \sqsubseteq \text{Å}\}$ is called the *consequences* of the argument Å. If $\mathcal{A}$ is a set of arguments then $Cn(\mathcal{A}) = \bigcup_i Cn(\text{Å}_i)$ where $\text{Å}_i \in \mathcal{A}$.

We are now in a position to extend the concept of justification function to that of defeasible arguments. The justification function of an argument is defined by applying the *weakest link principle* to its inference rules.

Definition 4.3.8. Given a defeasible argument $\text{Å} = \langle A, D, a \rangle$, its justification function $G_{\text{Å}}(X)$ is defined as $G_{\text{Å}} \colon 2^{\mathcal{L}} \longrightarrow \{0, 1/2, 1\}$, $G_{\text{Å}}(X) = \min_{d \in D}(\mathbf{H}_d(X))$.

Furthermore, since we deal with sets of arguments we need to have a justification function of an argument with respect to a set of arguments $\mathcal{A}$. For that we substitute conclusions of the arguments $\mathcal{A}$ for the set of sentences X in $G_{\text{Å}}(X)$.

Definition 4.3.9. For an argument $\text{Å} = \langle A, D, a \rangle$, its justification function with respect to a set of arguments is defined as $\mathbf{H}_{\text{Å}}(\mathcal{A}) = \mathbf{G}_{\text{Å}}(Cn(\mathcal{A}))$ where $Cn(\mathcal{A})$ is the set of consequences of all arguments in $\mathcal{A}$.

Observation 4.3.10. *For an argument $\text{Å} = \langle A, D, a \rangle$ in a defeasible reasoning system $\langle \mathcal{L}, \mathcal{R} \rangle$,*

1. $\mathbf{G}_{\mathring{A}}(X)$, and $\mathbf{H}_{\mathring{A}}(\mathcal{A})$ are well defined.

2. If $B \subseteq Cn(\mathcal{A})$ conclusively defeats $d \in D$ then $\mathbf{H}_{\mathring{A}}(\mathcal{A}) = 0$.

The above observation states that justification function of an argument is well defined. Furthermore, any set of arguments that conclusively defeats some rule of an argument, conclusively defeats that argument.

4.4 Context sensitive defeat and reinstatement relationships

The defeasiblity of arguments is captured by *defeat relation* between arguments. From presented defeat relationships, we are interested in undercutting attacks and rebuttal attacks. We capture rebuttal attacks through defeasible property of normal inference rules, without explicitly defining rebuttal attacks.

Unlike most argumentation systems where defeat is a direct binary relationship between individual arguments, in this system a group of arguments can cause or remove the defeat-condition for an argument, indirectly, via context. This property makes defeat a binary relationship between a group of arguments and an argument.

The phenomenon of separate arguments with same conclusion reenforcing each other is called accrual of arguments. Whether accrual of arguments is a valid argumentation concept or not is debatable [Pol02]. Nonetheless, since our defeat and reinstatement relationship is between a group of arguments and an argument the intended meaning of accrual of arguments [Ver01a] can be easily represented in this model of defeat relationship.

In a set of arguments $\mathcal{A}$, context is set by consequences (conclusions of all subarguments) of all arguments in $\mathcal{A}$. In order to show an argument set $\mathcal{A}_c$ attacks an argument Å in $\mathcal{A}$, we need to establish given an initial context $Cn(\mathcal{A}')$ where $\mathcal{A}' \subset \mathcal{A}$, addition of $Cn(\mathcal{A}_c)$ results in Å being defeated. The notion of defeat is connected to a decrease in degree of acceptability of Å that is

a decrease in $\mathbf{H}_{\text{Å}}(\mathcal{A})$. Accordingly, if $\mathbf{H}_{\text{Å}}(\mathcal{A}')$ is reduced to 0 it is said $\mathcal{A}_c$ outrightly defeats Å, and, if reduced to $1/2$ provisionally defeats Å. In the same token, in order to reinstate Å, $\mathcal{A}_c$ (or $\{\text{Å}_1\}$) has to increase the degree of acceptability of Å.

An example of attack and reinstatement is shown in Figure 4.3 below. The large arrows from arguments to context show contribution of arguments (their consequences) to the context. Let us assume an initial scenario where $\mathcal{A}' = \{\text{Å}, \text{Å}_1\}$, and Å is acceptable with respect to $\mathcal{A}'$. If we add Å_2 to $\mathcal{A}'$ a portion of $Cn(\mathcal{A}_{c1}) \subset Cn(\mathcal{A}')$ where $\mathcal{A}_{c1} = \{\text{Å}_1, \text{Å}_2\}$ attacks Å (shown by a circle). Yet, if add Å_3 to the mix, Å becomes acceptable with respect to the new argument set $\mathcal{A}$. It is as if a portion of context of $\mathcal{A}_{c2} = \{\text{Å}_2, \text{Å}_3\}$ reinstates Å against $\mathcal{A}_{c1}$.
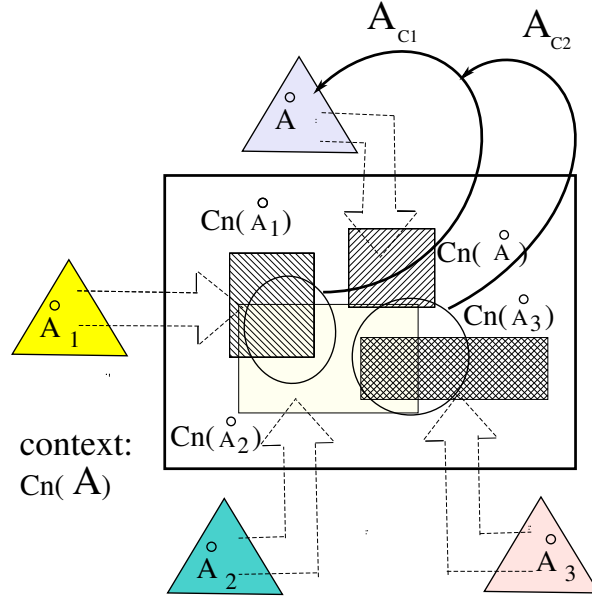

Figure 4.3: A schematic of arguments interaction

Definition 4.4.1. Let $\mathcal{A}$, $\mathcal{A}_c = \{\text{Å}_1, \text{Å}_2, ..., \text{Å}_n\} \subseteq \mathcal{A}$ be sets of arguments, and $\text{Å} \in \mathcal{A}$ an argument. We will say that the argument set $\mathcal{A}_c$ defeats the argument Å in $\mathcal{A}$

1. *outright* if and only if $\exists \mathcal{A}' \subseteq \mathcal{A}$ such that $\mathcal{A}_c \subseteq \mathcal{A}'$ and $\mathbf{H}_{\text{Å}}(\mathcal{A}') = 0$, and $\mathcal{A}_c$ is a maximal subset of $\mathcal{A}'$ where $\mathbf{H}_{\text{Å}}(\mathcal{A}' \setminus \{\text{Å}_j\}) \neq 0$ for all $\text{Å}_j \in \mathcal{A}_c$,

2. *provisionally* if and only if $\exists \mathcal{A}' \subseteq \mathcal{A}$ such that $\mathcal{A}_c \subseteq \mathcal{A}'$ and $\mathbf{H}_{\text{Å}}(\mathcal{A}') = 1/2$, and $\mathcal{A}_c$ is a maximal subset of $\mathcal{A}'$ where $\mathbf{H}_{\text{Å}}(\mathcal{A}' \setminus \{\text{Å}_j\}) = 1$ for all $\text{Å}_j \in \mathcal{A}_c$.

3. Furthermore, we say,

    (a) $\mathcal{A}_c$ is a defeat scenario for Å in $\mathcal{A}$,

    (b) the argument set $\mathcal{A}' \setminus \mathcal{A}_c$ is the *defeat context* of the corresponding defeat relation,

    (c) a defeat scenario $\mathcal{A}_c$ is a *conclusive defeater* of Å in $\mathcal{A}$ if and only if for every $\mathcal{A}'' \subseteq \mathcal{A}$, if $\mathcal{A}_c \subseteq \mathcal{A}''$ then $\mathbf{H}_{\text{Å}}(\mathcal{A}'') = 0$. If $\mathcal{A}_c$ is conclusive defeater in any arbitrary $\mathcal{A}$ then it is called *TConclusive defeater* of Å [5]. If $\mathcal{A}_c$ outrightly defeats Å and is not a conclusive defeater, we say it is a *nonconclusive defeater* of Å.

In definition above, the condition $\mathbf{H}_{\text{Å}}(\mathcal{A}') = 0$ might have been enough to say $\mathcal{A}'$ defeats Å. However there is a section of $\mathcal{A}'$ that is responsible for the defeat and there is a section of $\mathcal{A}'$ that acts as the context for the particular defeat scenario. We name the first part the defeater and the second part the context for defeat. Now we are in a position to show the following results.

Theorem 4.4.2. *Let $\mathcal{A}, \mathcal{A}'$ be two sets of arguments, $\mathcal{A} \subseteq \mathcal{A}', \text{Å}, \text{Å}' \in \mathcal{A}, \text{Å}$ a subargument of Å$'$, and $\mathcal{A}_c$ a defeat scenario for Å in $\mathcal{A}$ then,*

1. *$\mathcal{A}_c$ is also a defeat scenario for Å in $\mathcal{A}'$.*

2. *if $\mathcal{A}_c$ is a conclusive defeater of Å then $\mathcal{A}_c$ is also a conclusive defeater of Å$'$ in $\mathcal{A}$.*

3. *If D, the set of all rules of arguments in $\mathcal{A}$, is comprised of only indefeasible and default rules then all defeaters of arguments in $\mathcal{A}$ are TConclusive defeaters where the context of defeat is $\varnothing$. In addition, if both $\mathcal{A}_c$ is the minimal set of arguments where $F \in \mathbf{F}_d^k$, $F \subseteq Cn(\mathcal{A}_c)$, and, F is singleton set, then $\mathcal{A}_c$ is singleton set.*

4. *If $Cn(\mathcal{A}_c)$ conclusively defeats at least one defeasible rule in argument Å then $\mathcal{A}_c$ is a conclusive defeater of Å (though, the reverse is not necessarily true.)*

The results 3 and 4 of theorem 4.4.2 are the continuation of our attempt to draw a parallel between argumentation systems that are built upon conventional default rules and this argu-

---

[5]'T' for True in every argument set or universal defeater

mentation system. In most argumentation systems defeat relation is a static relation. Theorem 4.4.2.1 states that the defeat relation between a defeat scenario $\mathcal{A}_c$ and an argument Å is a static relation. However, this claim is contrary to our original claim that defeat relation is subject to a context. The reason for this apparent conflict is that we want to keep the proposed argumentation system in line with the conventional argumentation theories. In order to account for the influence of context over defeat relation we define a reinstatement by context relation.

In current argumentation systems an argument is reinstated only when its defeater is defeated. In our system arguments can reinstate other arguments by context without defeating their defeaters. The parts 1(a) and 1(b) in the following definition are the conventional method of reinstatement whereas parts 1(b) and 2(b) are exclusive to our system representing *reinstatement by context*. It can be seen that in case of conclusive defeat scenarios there is no reinstatement by context.

Definition 4.4.3. Let $\mathcal{A}$, $\mathcal{A}_{c1} = \{\text{Å}_{11}, \text{Å}_{12}, \ldots, \text{Å}_{1n}\} \subseteq \mathcal{A}$ and $\mathcal{A}_{c2} = \{\text{Å}_{21}, \text{Å}_{22}, \ldots, \text{Å}_{2m}\} \subseteq \mathcal{A}$ be three sets of argument and arguments Å $\in \mathcal{A}$, and $\mathcal{A}_{c2}$ be defeat scenarios for Å in $\mathcal{A}$.

1. $\mathcal{A}_{c1}$ is said to *outrightly reinstate* Å in $\mathcal{A}$ against $\mathcal{A}_{c2}$ if and only if either,

   (a) $\exists \text{Å}_i \in \mathcal{A}_{c2}$ such that $\mathcal{A}_{c1}$ is an outright-defeat-scenario for Å_i in $\mathcal{A}$, or

   (b) both

      i. $\exists \mathcal{A}' \subseteq \mathcal{A}$ such that $\mathcal{A}_{c1}, \mathcal{A}_{c2} \subseteq \mathcal{A}'$, $\mathbf{H}_{\text{Å}}(\mathcal{A}') = 1$ and

      ii. $\mathcal{A}' \setminus (\mathcal{A}_{c2} \cup \{\text{Å}_j\})$ is a defeat context for $\mathcal{A}_{c2}$ defeating Å for all $\text{Å}_j \in \mathcal{A}_{c1}$.

2. $\mathcal{A}_{c1}$ *provisionally reinstates* Å in $\mathcal{A}$ against $\mathcal{A}_{c2}$ if and only if $\mathcal{A}_c$ is not an an outright defeat scenario for any argument $\text{Å}_i \in \mathcal{A}_{c2}$ and either,

   (a) $\exists \text{Å}_i \in \mathcal{A}_{c2}$ such that $\mathcal{A}_{c1}$ is a provisional defeat scenario for $\text{Å}_i \in \mathcal{A}_{c2}$ in $\mathcal{A}$, or

   (b) both

      i. $\exists \mathcal{A}' \subseteq \mathcal{A}$ such that $\mathcal{A}_{c2}, \mathcal{A}_{c1} \subseteq \mathcal{A}'$, $\mathbf{H}_{\text{Å}}(\mathcal{A}') = 1/2$, and

143

ii. $\mathcal{A}' \setminus (\mathcal{A}_{c2} \cup \{\AA_j\})$ is a defeat context for $\mathcal{A}_{c2}$ outright defeating $\AA$ forall $\AA_j \in$ $\mathcal{A}_{c1}$.

We define an argumentation theory comprised of a set of arguments and all types of attack and reinstatement relationships given above. In order to interpret this argumentation theory into a Dung's Argumentation framework, we require to identify our system with all classes of attack and reinstatement relationships. The reason for this requirement is given in the next section.

Definition 4.4.4. Given a defeasible reasoning system $(\mathcal{L}, \mathcal{R})$, an *argumentation theory* is a tuple $\mathbb{AT} = \langle \mathcal{A}, \otimes, \oplus \rangle$ where $\mathcal{A}$ is a set of arguments constructed in $(\mathcal{L}, \mathcal{R})$, and $\otimes, \oplus$ are the defeat and reinstatement relationships between a set of arguments and an argument as defined above. Furthermore, given (A, D) in $(\mathcal{L}, \mathcal{R})$, if $\mathcal{A}$ is all the possible arguments that can be constructed in (A, D) then the argumentation theory $\mathbb{AT} = \langle \mathcal{A}, \otimes, \oplus \rangle$ is called an *induced argumentation theory* from (A, D). In addition, $\mathbb{AT}$ is called *context insensitive* if all defeat scenarios $\mathcal{A}_c \subseteq \mathcal{A}$ are conclusive defeaters.

Theorem 4.4.5. *Let $\mathbb{AT}\langle \mathcal{A}, \otimes, \oplus \rangle$ be an argumentation theory and D the set of all rules in all arguments in $\mathcal{A}$. Then, if D is comprised of only indefeasible and default rules, $\mathbb{AT}$ will be context insensitive.*

The next example shows the attack and reinstatement relations at work. This example is a modified version of example given in [SL92, DS01]. The example also shows the role of primary and ancillary reasons in an inference rule.

Example 4.4.6. Let $\mathcal{L} = \{a, \neg a, s, \neg s, r, \neg r, c, \neg c, e, \neg e\}$ with the following readings,

a : Tom is a mature adult; s : Tom is a student;
r : Tom has very rich parents; c : Tom has a car; and e : Tom is employed.

All the inference rules are normal, constituting

$$\mathcal{R} = \{d_1 : s \to \neg a,\ d_2 : s \to \neg e,\ d_3 : a \to \neg s,\ d_4 : a \to c,\ d_5 : a \to e,$$
$$d_6 : c \to e,\ d_7 : \neg e \to \neg c,\ d_8 : r \to c,\ d_9 : e \to c\}.$$

$d_2: s \to \neg e$		
1	1/2	0
$\{r\}$,	$\{c\}$,	$\{a\}$,
$\{c,r\}$	$\{a,r\}$,	$\{a,c\}$
	$\{a,c,r\}$	

$d_4: a \to c$		
1	1/2	0
$\{r\}$, $\{e\}$,		$\{s\}$
$\{s,r\}$,		
$\{s,e\}$,		
$\{r,e\}$,		
$\{s,r,a\}$		

$d_5: a \to e$		
1	1/2	0
$\{s\}$,	$\{s,r\}$,	
$\{c\}$,	$\{s,r,c\}$	
$\{r\}$,		
$\{c,s\}$,		
$\{c,r\}$		

$d_6: c \to e$		
1	1/2	0
$\{a\}$,	$\{s\}$,	
$\{r\}$,	$\{s,r\}$,	
$\{s,a\}$,	$\{s,a,r\}$	
$\{a,r\}$		

Table 4.3: The justification matrices of inference rules d_2, d_4, d_5, d_6 in example 4.4.6

The rules d_1, d_3, d_8, and d_9 are all default rules with no other defeater except the negation of their consequents. The rule d_7 is also a default rule, but having an additional conclusive defeater, namely the scenario $\{r\}$. The justification matrices of rest of the rules are provided in the tables below.

Let $\mathring{A}_1 = \langle\{s\}, \varnothing, s\rangle, \mathring{A}_2 = \langle\{r\}, \varnothing, r\rangle, \mathring{A}_3 = \langle\{a\}, \{d_5\}, e\rangle$, and $\mathring{A}_4 = \langle\{c\}, \{d_6\}, e\rangle$ be arguments in an induced argumentation theory derived from $\langle\mathcal{L}, \mathcal{R}\rangle$. Then $\mathcal{A}_{c1} = \{\mathring{A}_1, \mathring{A}_2\}$ is a provisional defeat scenario for $\mathring{A}_3$ but not for $\mathring{A}_4$ and $\mathcal{A}_{c2} = \{\mathring{A}_1\}$ is a provisional defeat scenario for $\mathring{A}_4$ but not for $\mathring{A}_3$. ∎

4.5   Semantics

The semantics of an argumentation system is determined by the rules of interaction between arguments. There are a number of approaches to provide the semantics of argumentation systems [PV01], e.g., assigning status to arguments [Pol94], defining the acceptable set(s) of arguments [Dun95b, BDKT97] and using dialectic argumentation trees [SCG$^+$94]. While there are minor differences, the approaches are driven by the same intuition where a definition in one can be an observation in another [Dun95b, QBVT05]. For instance in [Dun95b] it is shown the set

of justified arguments in [Pol87] is equivalent to the grounded extension in [Dun95b]. Dung's Argumentation Framework is used as basis in a number of argumentation systems. We adopt Dung's framework in order to give an anatomical picture of this system's behavior. We first translate proposed argumentation theory into a Dung argumentation framework and then apply Dung's semantics to the interpreted arguments.

In general, systems that are built upon Dung's system deal with conclusive defeats, and re-instatement of arguments is by defeating of their defeaters, and the provisional-defeat is an interpretation of multiple preferred extensions.

Figure 4.4 shows the underlying idea in translation to a Dung's abstract argumentation framework. Let $\mathbb{AT}$ be an argumentation theory where $\mathring{A}$ attacks $\mathring{A}_2$ and $\mathring{A}_3$; and $\mathring{A}_1$ reinstates $\mathring{A}_3$ against $\mathring{A}$. For simplicity we use arguments instead of argument sets. If we translate $\mathbb{AT}$ by one-to-one mapping between arguments in $\mathbb{AT}$ to arguments in AF, we get $\mathring{A}_1$ reinstating both $\mathring{A}_3$ and $\mathring{A}_2$, figure 4.4.a. In figure 4.4.a $\mathring{A}_1$ defends $\mathring{A}_3$ against $\mathring{A}$ which also leads to defending $\mathring{A}_2$. This translation is however incorrect. The reason is $\mathring{A}_1$ should only reinstate $\mathring{A}_3$. To obtain the desired translation we borrow the idea of *missing arguments* from examples 4.2.1 and 4.2.2. We assume there is an imaginary argument $\mathring{A}_4$, $\mathring{A} \sqsubset \mathring{A}_4$. This imaginary argument is shown by dotted lines in figure 4.4.b. Figure 4.4.b shows, it is $\mathring{A}_4$ that attacks $\mathring{A}_3$ and $\mathring{A}_1$ reinstates $\mathring{A}_3$ by attacking $\mathring{A}_4$. This time a one-to-one translation to a Dung's framework would yield $\mathring{A}_3$ reinstated and $\mathring{A}_2$ defeated. This result is the intended result. Hence, we need to distinguish between various types of defeats and reinstatement in our translation.

Definition 4.5.1 (Translation). Let $AF = \langle\, AR\,,\ ATT\,\rangle$ be a Dung argumentation framework, and $\mathbb{AT} = \langle \mathcal{A}, \otimes, \oplus \rangle$ an argumentation theory in $\langle \mathcal{L}, \mathcal{R} \rangle$. An abstract argumentation framework AF is a translation of $\mathbb{AT}$ if and only if,

1. there is a surjective function $\mathbf{M1}\colon\ 2^{\mathcal{A}}\ \longrightarrow\ AR$ such that,

    (a) For every $\mathring{A} \in \mathcal{A}$ there is an $\alpha \in AR$ for $\{\mathring{A}\}$,

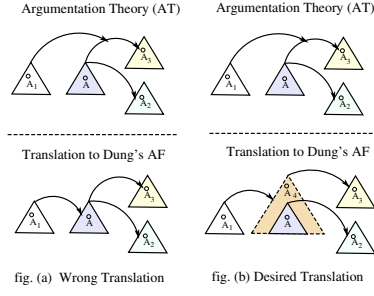fig. (a) Wrong Translation   fig. (b) Desired Translation

Figure 4.4: Translation to Dung's Argumentation Framework

(b) For every $\mathcal{A}_c$ where $\mathcal{A}_c \subseteq \mathcal{A}$ is a *defeat scenario* or *reinstatement scenario* for $\mathring{A} \in \mathcal{A}$ there is an $\alpha \in AR$ (if $\mathcal{A}_c$ is singleton then α is the same as α in 1(a)),

(c) if $\mathcal{A}_c$ is a reinstatement by *context scenario* or a non-conclusive defeat scenario there is one additional $\alpha_j^i \in AR$ for each defeat or reinstatement case (indexes i and j denote $\mathcal{A}_c$ and individual case j).

2. Given all $\alpha, \alpha_j^i \in AR$ as specified above then there is a surjective funcion $\mathbf{M2}\colon 2^{\mathcal{A}} \times 2^{\mathcal{A}} \longrightarrow ATT$ where ATT is determined such that,

   (a) if $\mathcal{A}_{c1}$ is an outright defeat scenario for an argument $\mathring{A}_k$, and β standing for any β mapped under $\mathbf{M1}$ for $\mathring{A}_k$ or any $\mathcal{A}_{ck}$ that $\mathring{A}_k$ is a member of then,

      i. if $\mathcal{A}_{c1}$ is a conclusive defeat scenario, then, $\alpha ATT \beta$ where α is the mapped $\alpha \in AR$ for $\mathcal{A}_{c1}$,

      ii. if $\mathcal{A}_{c1}$ is a non conclusive defeat scenario then $\alpha_j^i ATT \beta$ where α_j^i is the mapped α_j^i for the corresponding $\mathcal{A}_{c1}$ defeating $\mathring{A}_k$,

      iii. if $\mathcal{A}_{c1}$ is a provisional defeat scenario for an argument $\mathring{A}_k \in \mathcal{A}_{c2}$ then $\alpha_j^i ATT \beta$ and $\beta ATT \alpha_j^i$ where α_j^i is the mapped α_j^i for the corresponding $\mathcal{A}_{c1}$ provisionally defeating $\mathring{A}_k$.

   (b) if $\mathcal{A}_{c1}$ is a reinstatement by context scenario for an argument $\mathring{A}_k$ against the defeat scenario $\mathcal{A}_{c2}$, and α_j^i, β_j^i are the mapped arguments in AR for $\mathcal{A}_{c1}$ and $\mathcal{A}_{c2}$ for this

defeat reinstatement scenario then,

    i. if $\mathcal{A}_{c1}$ is an outright reinstatement scenario, then $\alpha_j^i ATT \beta_j^i$,

    ii. if $\mathcal{A}_{c1}$ is a provisional reinstatement scenario then $\alpha_j^i ATT \beta_j^i$ and $\beta_j^i ATT \alpha_j^i$.

To define the domain of operation we implicitly adopted values of status of arguments in [Pol94, PV01]. Assigning status to arguments is part of semantics in [Pol94]. We defined semantics based on Dung's semantics. We are yet to interpret semantics to values in the operaton doamin. The relation between BDKT-argumentation [BDKT97] semantics, and Pollock argumentation [Pol94] semantics is given in [QBVT05]. BDKT-argumentation semantics closely follows semantics given in [Dun95b].

Definition 4.5.2. Let $\mathbb{AT} = \langle \mathcal{A}, \otimes, \oplus \rangle$ be an argumentation theory in a defeasible reasoning system $\langle \mathcal{L}, \mathcal{R} \rangle$ and $AF = \langle AR, ATT \rangle$ its interpreted Dung framework, the status assignment function [6] $\mathbf{E}: \mathcal{A} \longrightarrow \{0, 1/2, 1\}$ is

$$\mathbf{E}(\text{Å}) = \begin{cases} 1 & \text{if } X_p \text{ is in all preferred extensions in AF} \\ 1/2 & \text{if } X_p \text{ is in at least one, but not all the preferred extensions in AF} \\ 0 & \text{if } X_p \text{ is in none of the preferred extensions in AF} \end{cases}$$

where X_p is the mapped argument α in AR for $\{\text{Å}\}$. Furthermore, The status of a literal $x \in \mathcal{L}$ is given by the status function $\mathbf{S}(x) = \max_{\text{Å} \in \mathcal{A}}(\mathbf{E}(\text{Å}))$ where $x \in Cn(\text{Å})$, for any $x \in \bigcup_{\text{Å} \in \mathcal{A}} Cn(\text{Å})$, otherwise $S(x) = 0$. The values $1, 1/2$ and 0 stand for *justified, defensible* and *overruled*.

The following observation ensures that semantics of an argumentation theory $\mathbb{AT}$ is well defined.

Theorem 4.5.3. *In an argumentation theory $\mathbb{AT}$, every argument Å and literal $a \in \mathcal{L}$ has one and only one status.*

---

[6] Instead of *preferred semantics* we could have given admissible semantics by saying $\mathbf{E}(\text{Å}) = 1$ *if X_p is in an admissible extension where no attacker of X_p is in any admissible extensions*, and $\mathbf{E}(\text{Å}) = 1/2$ *if at least one attacker of X_p is also in a admissible extensions.*

The following example applies the above translatiion to the framework given in example 4.4.6.

Example 4.5.4. We extend example (4.4.6) as follows. Let $A = \{s, a, r\}$ and $\mathbb{AT} = \langle \mathcal{A}, \otimes, \oplus \rangle$ be an induced argumentation theory from $(A, \mathcal{R})$ wher,

$$\mathcal{A} = \{ \text{Å}_1 = \langle \{s\}, \emptyset, s \rangle, \text{Å}_2 = \langle \{a\}, \emptyset, a \rangle, \text{Å}_3 = \langle \{r\}, \emptyset, r \rangle,$$

$$\text{Å}_{11} = \langle \{s\}, \{d_1\}, \neg a \rangle, \quad \text{Å}_{12} = \langle \{s\}, \{d_2\}, \neg e \rangle, \quad \text{Å}_{121} = \langle \{s\}, \{d_2, d_7\}, \neg c \rangle,$$

$$\text{Å}_{21} = \langle \{a\}, \{d_3\}, \neg s \rangle, \quad \text{Å}_{22} = \langle \{a\}, \{d_4\}, c \rangle, \quad \text{Å}_{221} = \langle \{a\}, \{d_4, d_6\}, e \rangle,$$

$$\text{Å}_{23} = \langle \{a\}, \{d_5\}, e \rangle, \quad \text{Å}_{231} = \langle \{a\}, \{d_5, d_9\}, c \rangle,$$

$$\text{Å}_{31} = \langle \{r\}, \{d_8\}, c \rangle, \quad \text{Å}_{311} = \langle \{r\}, \{d_8, d_6\}, e \rangle \ \}$$

The calculated state of $\text{Å} \in \mathcal{A}$ is:

$$\mathbf{E}(\text{Å}) = \begin{cases} 1 & \text{if } \text{Å} \in \mathcal{A}_2 \\ 1/2 & \text{if } \text{Å} \in \mathcal{A}_3 \\ 0 & \text{if } \text{Å} \in (\mathcal{A} \setminus (\mathcal{A}_2 \cup \mathcal{A}_3)) \end{cases}$$

where $\mathcal{A}_2 = \{\text{Å}_1, \text{Å}_2, \text{Å}_3, \text{Å}_{22}, \text{Å}_{221}, \text{Å}_{23}, \text{Å}_{231}, \text{Å}_{31}\}$,

and $\mathcal{A}_3 = \{\text{Å}_{12}, \text{Å}_{23}, \text{Å}_{231}, \text{Å}_{311}\}$. ∎

We define consistency in an argumentation theory based on its Dung's AF translation. For an argumentation theory to be consistent, no two justified arguments should have contradictory conclusions. Since the acceptability of an argumient is captured through admissible set(s) then no two arguments in any given admissible set should have contradictory consequences.

Definition 4.5.5. An argumentation theory $\mathbb{AT} = \langle \mathcal{A}, \otimes, \oplus \rangle$ is said to be consistent if and only if there is no $a \in \mathcal{L}$ such that $a \in Cn(\text{Å}_1)$, $\neg a \in Cn(\text{Å}_2)$, $\text{Å}_1, \text{Å}_2 \in \mathcal{A}$, and, the corresponding mapped arguments α_1, α_2 in the translation of $\mathbb{AF}$ belong to the same preferred extension.

Theorem 4.5.6. *Fr an argumentation theory* $\mathbb{AT} = \langle \mathcal{A}, \otimes, \oplus \rangle$,

1. *If* $\mathbb{AT}$ *is consistent then* $\forall a \in \mathcal{L}$ *if* $\mathbf{S}(a) = 1$ *then* $\mathbf{S}(\neg a) \neq 1$ *(note:* $\neg(\neg a) = a$*).*

2. *Let* $D = D_1 \cup D_2$ *be the set of defeasible rules of all arguments in* $\mathcal{A}$ *where* D_1 *is the set of indefeasible and* D_2 *the set of normal rules; and,* C *be the set of consequences of all arguments in* $\mathcal{A}$ *(i.e.* $C = \bigcup_{\text{Å} \in \mathcal{A}} Cn(\text{Å})$*). Then,*

> *(a) if the induced argumentation theory $\mathbb{AT}'$ from (C, D_1) is consistent then $\mathbb{AT}$ is consistent.*
>
> *(b) The mapping in the translation of $\mathbb{AT}$ to AF is a bijective mapping. Moreover, if $\mathbf{F}_d^k$ for all the rules $d \in D_2$ is composed of singleton sets then AF is isomorphic to the structure $\langle \mathcal{A}, R^* \rangle$ where $R^* = \{(\mathring{A}_1, \mathring{A}_2) \in \mathcal{A} \times \mathcal{A} \mid \mathring{A}_1 \text{ defeats } \mathring{A}_2\}$.*

The first part of theorem 4.5.5 states that for an argumentation theory $\mathbb{AT}$ whose indefeasible part is consistent, if its defeasible part is comprised of only normal rules, then, its is a consistent argumentation theory. The second part of theorem 4.5.5 draws a parallel between our argumentation system and argumentation systems that is based on default rules. The bijective mapping means that there are no imaginary arguments in translation AF. Furthermore, if the defeaters in default rules are singleton defeaters then $\mathbb{AT}$ has the same structure as its Dung's translation AF. By this theorem we conclude this chapter.

4.6 A Short Comparison with Other Defeasible Reasoning Systems

In this section we briefly discuss how the defeasible rules we proposed relate to other rule based defeasible reasoning systems, namely BDKT abstract assumption-based framework [BDKT97], and the argumentation based Defeasible Logic [GMAB04]. These defeasible rules are essentially assumption based defeasible rules applicable in the absence of any contrary evidence in $\mathbf{F}_d$ and $\mathbf{U}_d$. Our argumentation theory is therefore a form of assumption based default reasoning theory. The language $\mathcal{L}$ and defeasible rules are similar to the language and rules in Defeasible Logic where language, body and head of rules are comprised of literals.

In relation to BDKT, assuming that $\mathbf{U}_d$ in all rules is empty, expanding $\mathcal{L}$ to include sentences with logical conjunction '$\wedge$' makes the proposed defeasible rules a form of grounded Reiter defaults where the assumption is the negation of elements of $\mathbf{F}_d$. It is already shown that

an argumentation theory based on this modified version of rules can be captured by BDKT framework [BDKT97]. However, in BDKT there is no direct means of addressing the $\mathbf{U}_d$ part of a rule. In other words, in BDKT an acceptable set of assumptions can not to provisionally defeat an assumption in a direct fashion.

In relation to Defeasible Logic, assumption based rules and the corresponding undercutting attacks on the assumptions of rules can be represented in two ways depending on whether the assumptions are explicitly or implicitly expressed. In our approach assumptions are implicitly expressed. For a rule $r : bd(r) \Rightarrow hd(r)$ with an implicit assumption a, r is divided into two rules $r_1 : bd(r) \Rightarrow inf(r)$ and $r_2 : inf(r) \Rightarrow hd(r)$. [7] The undercutting attack on the assumption of the rule can be expressed by $r_{dft} : \neg a \rightsquigarrow \neg inf(r)$ (or, alternatively $r_{dft} : \neg a \Rightarrow \neg inf(r)$). [8] Though undercutting attacks can be expressed in Defeasible Logic, attacks themselves are invariant. So, we cannot represent reinstatement of a rule by context unless we expand the rule base by additional rules as discussed in example 4.2.1. Furthermore, even if we assume that there is no reinstatement by context (effectively making attacks invarient,) a translation of $\mathbf{U}_d$ in terms of undercutting defeater would only make sense for ambiguity blocking semantics [GMAB04] of Defeasible Logic. There is no direct means of addressing $\mathbf{U}_d$ in terms of a defeater for an ambiguity propagating semantics [GMAB04].

As it can be seen the two reasons why this argumentation system cannot be directly expressed in BDKT framework or in argumentation based Defeasible Logic are $\mathbf{U}_d$ and reinstatement by context. In light of the results and discussion above, it can be argued that with appropriate semantics, an induced argumentation theory from the proposed default rules can indeed be embedded in both BDKT framework and Defeasible Logic. It is possible to envisage a schema for translating a non-default rule to a set of new default rules, effectively constructing an argumentation theory that consists of only indefeasible and default rules. The translated argumentation

---

[7] Splitting rules in this fashion is originally proposed to express superiority relation among conflicting rules [GMAB01].

[8] The reverse translation, i.e. translating a Defeasible Logic rule to a Default Logic rule, is given in [?].

theory will be equivalent to the original theory with respect to the status of literals in $\mathcal{L}$. Such translation allows us to capture a given argumentation theory in BDKT framework (or in Defeasible Logic.) The schema is similar to the method used for translation to Dung's Framework.

9

4.7 Discussion and future direction

In this chapter we proposed a simple representation of defeasible rules consisting of only literals. Each defeasible rule is associated with a justification function. The justification function effectively describes under what circumstances a rule cannot be applied. The antecedents of a rule are the primary reasons for believing the consequent. The literals in the justification function are taken to be the ancillary reasons that strengthen or weaken a rule. Unlike most argumentation systems, in this system arguments attack or reinstate other arguments indirectly via context. The context is the collection of consequences of arguments in a set of arguments. We also provided a translation from this system to Dung's abstract argumentation theory in way of validating our approach.

Our investigation into defeasible rules in the context of argumentation systems is programmatic in character. There are some important issues that have not yet been addressed.

1. Our system shares a problem regarding the non-normal Reiter defaults [Poo91] in relation to two seemingly acceptable arguments that are built upon contradictory assumptions. We will address this problem along the lines suggested in [BDKT97] where the conceded assumptions are explicitly stated.

---

[9] We provide only the basic idea behind the schema. For every rule d, we construct all possible sequences of the form $J_1, J_2, \cdots, J_n$ where $J_i \subseteq \mathbf{J}_d$, $J_1 \in \mathbf{U}_d$ or $\mathbf{F}_d$, $J_i \subset J_{i+1}$, and J_{i+1} is the minimum J_k (with respect to $\subseteq$) that is not in the same class as J_i (class in terms of $\mathbf{T}_d, \mathbf{U}_d, \mathbf{F}_d$). For every distinct $J_{i+1} \setminus J_i$ we construct a rule d_i^* where $bd(d_i^*) = J_{i+1} \setminus J_i$, $hd(d_i^*) = a_i^*$, $\mathbf{J}_{d_i^*} = \mathbf{F}_{d_i^*} = \{b_i^*\}$, and $a_i^* = b_{i-1}^*$. In addition, $bd(d_1^*) = J_1$ and $\mathbf{J}_{d_n^*} = \mathbf{F}_{d_n^*} = \varnothing$; and, if $J_{i+1} \in \mathbf{U}_d$ then $a_{i-1}^* = b_i^*$. This way we extend $\langle \mathcal{L}, \mathcal{R} \rangle$ to $\langle \mathcal{L}^*, \mathcal{R}^* \rangle$ by newly introduced a_i^*, b_i^* and d_i^*. It can be shown that an argumentation theory from substitution of non-default rules with the corresponding set of default rules is equivalent to the original theory with respect to $\mathcal{L}$.

2. In general, the contrapositives of default rules are not automatically allowed in defeasible reasoning systems [BDKT97, Poo91]. Yet, it has been argued that the contrapositives of defaults can help avoiding certain counterintuitive results [CA05].

In our future work we aim to show how a rule can be explained by other rules, including expressing non-default rules as a set of default rules. We will address introduction of logical connectives in the antecedent of a rule, as well as, giving a more in depth comparison with other defeasible reasoning systems.

Chapter 5

Conclusion

In this dissertation we attempted to address two different aspect of formal argumentation theory. Both serve the two governing features of reasoning by argumentation, the localization of a reasoning to its relevant factors, and, the reasoning by inquiry. Accordingly, in chapters two and three we presented the admissibility backings of arguments. The admissibility backings of arguments localize the admissibility or dismissibility of arguments in an argumentation framework. In chapter four, we presented a new type of defeasible inference rules, called context sensitive rules. The context sensitive rules extend the reach of reasoning by argumentation to many instances of practical reasoning. The presented findings can follow many avenues, especially, in the case of the admissibility backings of arguments.

5.1 A summary of the thesis achievements

The following is a detailed summary of thesis achievement that was discussed in section 1.2.5.

5.1.1 Admissibility backings, its propagation, and the role that arguments play in admissibility of others

The issues in artificial intelligence usually tend to interlace together, for instance the issues of *learning, contradiction* and *change* are closely related. The classic example which is relevant

to our topic is the twin problems of the *searching* and the *sorting*. Searching, when things are unsorted, takes a lot more effort than when things are sorted. On the other hand, the sorting itself requires searching. However, once things are sorted, the subsequent search and sorts take the minimal effort.

The importance of the proper structuring of a knowledge base is therefore self evident, and, so a central subject in all the fields of computer science, from the database design to the mapping of Bayesian networks, to the constraint programming and constraint propagation. The formulation of admissibility backings of arguments is aimed to follow this general goal. In chapter two, we localized the admissibility of arguments in terms of the minimal admissible sets that accept or attack an argument. We called these the (admissibility) backings of arguments. we called the minimal admissible sets that accept an argument, the positive backings of an argument and the minimal admissible sets that attack an argument, the negative backings of an argument.

The *grounded admissible extensions* are an important class of sets of arguments. They are important because they identify arguments that have the property of being accepted beyond a reasonable doubt [GW09]. Consequently, we presented a class of *grounded admissible sets*, and distinguished between the grounded and the not-grounded admissibility backings.

We defined the admissibility backings free of any special requirement on the argumentation frameworks. However, a motivating principle behind the admissibility backings of an argument is for them to carry all the information regarding the admissibility situation of that argument.

This motivation sets certain expectations on the backings of arguments. One of such requirements is that if an argument has no positive backings then we should expect for that argument to have some negative backings. However, only a certain class of argumentation frameworks, reflect fully this intention behind the backings of arguments. As part of our analysis, we characterize this class of frameworks as the *normally stable* argumentation frameworks. Accordingly,

we showed that a normally stable framework is a framework that both itself and all its normal sub-frameworks are *coherent*.

Another major result of this thesis is the identification of the dependency relation between the arguments. A distinct property of the admissibility backings is that each backing operates independent of other arguments in the framework. Hence, whenever an admissibility backing of an argument is present, the argument is duly accepted or rejected irrespective of the other arguments in play. It is easy to see how this result sheds light on any notion of dependency between the admissibility of arguments.

As part of our analysis, we also identified the relation between the admissibility backings of an argument and the admissibility backings of its attackers. This relation is central to many of our other findings. We presented this relation by means of two algebraic operators $+, \circ$. The algebraic relation between the admissibility backings of an argument and of its attackers also lends itself to a recursive formulation. We also provided the operators $+, \circ$ and the accompanying recursive formula with a number of simplification results. The simplification results are intended to show how the process of finding the backings of arguments can be made efficient.

The relationship between the admissibility backings of an argument and the admissibility backings of its attackers then sets us up in three interlacing directions. The three directions are, (1) the relevance of an argument in regard to the admissibility of other arguments, (2) the propagation of backings along the attack sequences, and, (3) *active attack sequences* and the *intercepts* provide us the means to split a framework into sub-frameworks of distinct characteristics.

The active attack sequences, in general, mark the lines on which the admissibility backings propagate. In connection with the active attack sequences, we introduced the notion of intercepts. The intercepts correspond to arguments that are made irrelevant for the admissibility of some argument. An argument that is neither an in/direct attacker or an in/direct defender of some argument a is by default intercepted for argument a. Hence, the intercepts identify one

type of relation between arguments. The intercepts also play a distinct role in that they split an argumentation framework into *independent* sub-frameworks.

The other presented arguments relations are the active argument relation, the critical argument relation, the incompatible argument relation and the redundant argument relation.

Any argument that belongs to some backing of an argument is an active argument for that argument. The admissibility backings are need to adhere to a certain minimality condition. As a result, the propagation of active arguments along the active attack sequences is not guaranteed. Furthermore, not all the arguments on an active attack sequence of an argument are active arguments for that argument. Such arguments are identified by the *incompatible* and *redundant* argument relations. The identification of incompatible and redundant arguments are done with the help of *critical arguments*, arguments that are indispensable for the admissibility of an argument.

5.1.2 Context sensitive defeasible rules

In chapter four of this thesis we presented the context sensitive rules. The context sensitive rules are based on the practical consideration that many instances of inference not only involve primary reasons but also involve ancillary reasons. The role of primary reasons is to trigger whether or not an instance of a defeasible rule is applicable in the first place. The role of ancillary reasons is to fine-tune the applicability of the rule once the primary reason signals that the rule is relevant. In this sense, the ancillary reasons of a rule reflect the context for the applicability of the rule. The aim is therefore to present an argumentation system that covers a wider range of defeasible reasoning.

A feature of the introduced inference rules is that they subsume the conventional defeasible rules, the default rules. Therefore, the presented argumentation system can augment the systems based on the conventional default rules. The provided semantics for this argumentation

system is by means of translating it into Dung's abstract argumentation framework. Hence, the framework also meets the standard semantics in the literature.

An important achievement of chapter four is the introduction of the *missing arguments*. The missing arguments highlight the deficiencies in the conventional treatment of attack relation. These deficiencies were highlighted with two motivating examples. The examples show a most basic case where some otherwise legitimate arguments can be hidden or unknown to an agent. The role of missing argument are then explicated with respect to the attack and reinstatement relations. The chapter four finally shows how to account for the existence of missing arguments in terms of the current approaches to the formal argumentation theory. Under this interpretation of the role of the missing arguments we provided a concise reading of the presented framework in Dung's framework.

In order to address the role of *missing arguments* in relation to attack relation, we defined a new form of attack relation where an argument attacks an attack relation. A consequence of this is that arguments can reinstate other arguments without attacking their attackers. But, instead they attack the attack relation. The missing arguments provide the explanation that why this new form of attack relation is still in line with the conventional attack relation. The explanation is that the reinstating argument conventionally attacks the hidden missing argument.

We should note that the attack relation between an argument and an attack relation is already presented in [Mod06]. The difference between our new attack relation and the one presented in [Mod06] is that ours is presented and accounted for under the notion of missing arguments.

5.2 Future work

5.2.1 Splitting, merging, and the dynamics of argumentation theory

The main motivating factor behind the formulation of backings of arguments is to setup a background theory to address some of the important issues in argumentation theory, including the splitting, merging, the dependence relation, and, the dynamics of argumentation theory.

In section 3.7 we split a framework into *independent* sub-frameworks. Continuing on this work, we want to split a framework into sub-framework of distinct characteristics. Each characteristic is set to identify a class of argumentation frameworks. Accordingly, we want theses classes of argumentation frameworks to have certain properties of their own.

Each class of framework is set to preserve certain information that is efficiently accessible. There should also be certain relations between the classes of frameworks so that the frameworks of the same or different class can be formulated, merged, and split into frameworks of the same or a different class.

The goal is to formulate operations for the translation, merging, and split of a framework that are in terms of the operations between the backings of arguments in each class of argumentation framework. In this regard, we envisage to employ the positive and negative attack sequences, and the intercepts as well as their relations with respect to each other. The reason for this choice is that the positive and negative attack sequences, in one form or another, preserve the propagation of the backings where the fall out is mapped by the intercepts.

The split of a framework depends on what type of information we wish to preserve in each class of framework. In turn, the *type of information* is determined by what sort of question we want to answer. The *preservation of information* is then guided by where the lines of dependence or independence relations are drawn. Consequently, the future work should first address the *dependence* relation between arguments with respect to their admissibility.

5.2.1.1 The dependence relation between the arguments with respect to their admissibility

The domain of our inquiry is set by the questions we want to answer. Each domain of inquiry then identifies its own corresponding independence relation. For instance, we may set the dependence relation in regard to changes in the admissibility status of an argument, or, in the backings of an argument, or, the backings of attackers of an argument.

We can follow all such forms of independence relation between arguments along the lines of active attack sequences. One crude way to test for the dependence relation is removing an argument and look for the ensuing changes (with respect to the domain of inquiry). If any change is noticed then a dependence relation is in play.

To establish the independence relation requires careful considerations. To establish an independence relation, we need to consider all possible scenarios that involve the arguments in question. We denote all possible cases through the sub-framework relation, as we have done in this thesis (see sections 2.6 and 3.7). In a similar fashion it is also possible to address the conditional independence relation where the *condition* is represented by a sub-framework. We can then employ the statements of the form $AF_c \sqsubseteq AF' \sqsubseteq AF''$ where AF_c represents the *condition in place*. That is, given AF_c, the answers to certain inquiries do not change between AF', AF''.

5.2.1.2 The future work in regard to the sub-argumentation framework relation

The study of sub-argumentation framework relation ranges over many topics. One application of sub-argumentation framework is in formulation of the equivalence relation(s) between frameworks.

In formulating an equivalence relation, we generally use some mapping from one framework to another. The mappings are in general intended to preserve certain characteristics. Accordingly, the mappings can be admissibility preserving, admissibility status preserving, etc,. Regardless

of the mapping that we use, the equivalence relation under a mapping should not only hold over the two frameworks, but it should also hold over all their sub-frameworks, or at least over a class of their sub-frameworks.

Furthermore, the sub-argumentation framework relation forms a partial order. The range of operators one can define with respect to partial orders is well studied in the literature, all with their own distinct properties and utilities. Accordingly, one area of study of sub-argumentation framework relation is in regard to the operators we can define over them. In this thesis, we presented two of such operators, $+^N, -^N$ when we discussed the splitting of a framework into independent sub-frameworks..

5.2.1.3 The subargument relations

The subargument relation is well known within the literature, and it is recognized to be a fundamental relation between arguments with important applications. It is fundamental because, it directly relates to the semantics. Accordingly, it has important applications, as it has the potential to apply certain type of *cut* property.

The connection between subargument relation and the semantics of argumentation systems is generally expressed in terms of the *weakest link principle*. It says that an argument is accepted only if all its subarguments are accepted. This principle is generally agreed among the research community. There are however additional intricate issues for which the verdicts are not yet finalized [PV01].

Regardless, the expansion of Dung's framework by the subargument relation under the weakest link principle is a straightforward task. Naturally, being still within Dung's framework, we can claim that all our results so far equally hold for a Dung's framework with a subargument relation. However, what interest us is the types of substructures that the subargument relation ensues, and, how these substructures can be formulated in terms of the backings of

argument.

5.2.1.4 Strength of arguments

In general we aim to model *strength of arguments* in form of a partial or total order over arguments. The current approaches to model the *strength of arguments* are derived by the application of some preference or priority order over defeasible inference rules. The same concept is then adopted to the abstract argumentation frameworks [PS96b, AC98, AP02, Ben02].

Within the current literature it is implicitly assumed that the preference are provided by an external source. How these preferences are obtained or how reliable they are, are separate questions. Hence, it is assumed that these preferences are as given. Many of these approaches also provide certain calculus on how to resolve the tie between the conflicting arguments by means of the provided preferences. Thus, such preferences effectively determine the semantics of an argumentation system, e.g., the framework in [Ben02]. However, it is not shown that after application of the provided calculus the result shall meet our expectations, i.e., whether or not our reasoning by such calculus is sound.

We however view the strength of arguments as a measurement derived from arguments interactions within an argumentation system. This way the strength of arguments is a measurement that we obtain from an argumentation system, and not building an argumentation system from such preferences that are accompanied with its own calculus. Hence, it is the semantics of an argumentation system that provides the strength of an argument and not the other way around.

Our approach for measuring the strength of arguments is simple. Given a set of sub-argumentation frameworks, we calculate how admissible an argument is with respect to the set. Roughly speaking, an argument that is admissible in more sub-frameworks is stronger than an argument that is admissible in less number of sub-frameworks. It is the rule of survival at the face of uncertainty, the stronger survives.

In this thesis, we directly tied the admissibility of an argument to its backings. Hence, without going into details, we can see that the more positive backings an argument has, and the more these positive backings can survive different frameworks, the stronger the argument is. Hence, the findings in thesis serve as a background theory by which we can derive and explain this measure. The background theory can also account for the conditional measures and the conditional independence relation. For instance, we can reason that given availability of some arguments, one argument is always stronger than another. The underlying theory that explains this finding is that the given set of arguments safeguard more positive backings of the first argument than those of the second argument.

This modeling of strength of arguments will be more useful after the inclusion of the subargument and the independence relations. For instance, in [Lou87], Loui presents a number of heuristic rules of thumb about why to prefer one conflicting set of arguments over another. Some of his presented rules of thumb are classified under the *the directness of an argument* which in turn translates into the *size* of an argument by the number of its subargument. For example it is said that a longer argument tends to be weaker than a shorter argument. In other words, an argument that has less number of subarguments tends to be a stronger argument than an argument that has more subarguments. Using our approach along with the subargument relation, we can show the validity of these heuristic assumptions.

5.2.2   Future work for context sensitive defeasible rules

There are a number of future inquiries we can pursue in regard to the context based argumentation system. We briefly mention a few.

One line of inquiry is to present an abstract context based argumentation framework. We have already translated the current context based framework into Dung's abstract framework. Hence, it is feasible to have an abstract context based framework.

All that is required is to equip either of the frameworks in [Boc02] or [NP06] with the ability for a set of arguments to attack an attack relation. The two extended abstract frameworks in [Boc02, NP06] allow for an attack relation where a set of arguments can attack an argument. Moreover, the attack relation between an argument and an attack relation is already presented in [Mod06]. Hence, the resulting framework will be a combination of these three frameworks.

There are three other future inquiries. Currently as it stands we use a justification table and a corresponding justification function to state how the ancillary reasons influence the applicability of a rule. However, to properly follow the underlying theme of our motivation, we need to encode the interaction between the primary reasons and ancillary reasons in terms of the corroborating and conflicting reasons, instead of a justification status.

Another line of future research is to ensure that there are no redundancies in the justification table. A justification table should not be reducible to smaller justification tables.

Finally, it will be interesting to seek a connection between the presented defeasible rules and the argumentation schemes. For instance, we may investigate how the missing arguments can account for the *enthymeme* in arguments that are modeled based on the arguments in natural languages. In these cases, the implied or presumed premises act as the context to draw a conclusion or attack other arguments. The context can lead us to identify the missing or hidden arguments. The identification of missing arguments will then help us to explicate whether or not a context is used properly.

Appendices

Lottery Paradox

The lottery paradox assumes a fair lottery of 10000 tickets where there will be one winner. Let P_i denote the argument that the ith ticket will not win. We have then 10000 arguments, one argument for each ticket. By the "statistical syllogism", one can claim that all the arguments $P_1, \cdots, P_{10000}$ are prima facie justified. The question is whether or not we could accept all the arguments P_1 to P_{10000}.

Obviously, since someone has to win, the conjunction of conclusions of all the 10000 arguments is false. The overall verdict to whether or not we can accept all the 10000 arguments, is that we cannot. [1]

The rational is that if two non-contradictory arguments are accepted, then, conjunction of their conclusions should be accepted, too. In the case of lottery paradox, we are faced with a situation that no single argument attacks another. The paradox is that while we cannot not dismiss any single argument, we cannot accept them either. Pollock presents this as a form of *collective defeat*. The other forms of collective defeat relate to what is commonly referred to as the even and odd length attack sequences. It is difficult to relate the lottery paradox form of *collective defeat* to either category of the rebutting or undercutting defeat relation.

---

[1] Despite the rational put by Pollock, I am not convinced, why we cannot accept all the 10000 arguments. All the arguments appear to be good. The only argument that we cannot accept is that we can conjunct the 10000 conclusions. We are not warranted to conjunct the conclusions, on the account that it undermines the *normality assumption*, the assumption under which all arguments of any kind are constructed. Otherwise, we cannot accept any argument. The argument for the conjunction of 10000 conclusions is therefore a *self defeating* argument. This though is against the adopted consensus in the literature, that the *strict rules* have the unconditional warrant of to be applied anywhere, anytime. In this case, the strict rule is in the form of material conditional, "$Q_1, Q_2, \cdots, Q_{10000} \implies Q_1 \wedge Q_2 \wedge \cdots \wedge Q_{10000}$", where Q_i is the conclusion of the argument P_i.

Bayesian Belief Networks and Argumentations

In Bayesian belief networks, the locality of inference between two adjacent nodes is safeguarded by the postulates of the independence relation. Roughly speaking, in Bayesian belief networks, the locality is captured by the directed causal link between two adjacent nodes. In Markov network, there are no directed causal links. In Markov network, the locality is instead captured by the *blanket* of a node. A blanket for a node is the boundary nodes of a node.

There are a number of works on drawing a connection between the argumentation systems and the Bayesian belief networks. Non of the works however makes a general mapping from one reasoning system to another. The attempts generally fall within three categories. They either build some form of an argumentation framework based on the rational behind the Bayesian belief networks [FNL13], or, form the data produced by the network [NP07], or, they draw a parallel between the two systems of reasoning by heuristic analysis [Vre04].

There are still many issues to be resolved before drawing a mapping between the Bayesian networks and argumentation systems. For instance, one of the early topics that still to be conclusively discussed is the two way reading of contrapositive and abductive inference rules in a Bayesian network [Poo92, Poo93b, Poo93a, Poo97, Poo00]. If there was such a reading then relation between an inference rule and its abductive and contrapositive rule would be clear. But this is not the case.

In most argumentation systems, given an inference rule, its contrapositive, and, its abductive

form (if the abductive form applies), must explicitly be stated in the knowledge base. These restrictions are in large due to the technical reasons associated with the particular argumentation system. For instance, in the assumption based argumentation frameworks, that are largely implemented in the form of logic programming, the contrapositive of inference rules are not automatically available.

One simple example is the inductive inference that if something is red then it appears red. The contrapositive of this inference is, if something does not appear red then it is not red. Its abductive from is, if something appears red then it is red.

In the statistical syllogism, from the "most Ps are Qs", we construct the rule $P \rightarrow Q$. Rationally, assuming that our universe is not just comprised of entities with the property P, we are allowed to think that if something is not Q then it is probably not P either. For instance, let us take the inference that if something is a bird then it flies. In most frameworks, we automatically cannot have its contrapositive which is if something does not fly the it is not a bird. However, we normally expect that if something does not fly then it is not a bird. We can see this type of expectations when playing games like *Twenty Questions* or *I spy with my little eye*. In either game, if we are asked whether or not something flies, and the answer is no, we then think that the mystery object is not a bird.

The properties of $+, \circ$ and $\dot{+}, \dot{\circ}$ operators

In an argumentation framework, the admissible sets create a partial order with respect to the set inclusion. Hence, it is no surprise that the operations $\circ, +$ and $\dot{\circ}, \dot{+}$, pairwise, do possess many properties of a semiring. the following theorem presents this finding.

Theorem .0.1 (The properties of the operations $\circ, +$ and $\dot{\circ}, \dot{+}$)**.**

In the followings $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are each a set of sets of arguments in some argumentation framework $AF = \langle AR, ATT \rangle$, and, $\otimes$ stands for the operations $\circ, \dot{\circ}$ and $\oplus$ stands for $+, \dot{+}$. Moreover any combination of $\oplus, \otimes$ refers to a corresponding combination of $+, \circ$ or $\dot{+}, \dot{\circ}$.

1. *$\mathcal{A} \otimes \varnothing = \varnothing \otimes \mathcal{A} = \varnothing$.*   *(Absorbing Element)*

   If no $B \in \mathcal{B}$ is conflict free then $\mathcal{A} \otimes \mathcal{B} = \mathcal{B} \otimes \mathcal{A} = \varnothing$.

2. *$\mathcal{A} \otimes \mathcal{B} = \mathcal{B} \otimes \mathcal{A}$.*   *(Commutativity Property)*

   $\mathcal{A} \oplus \mathcal{B} = \mathcal{B} \oplus \mathcal{A}$.

3. *$(\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C} = \mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C})$.*   *(Associativity Property)*

   $(\mathcal{A} \oplus \mathcal{B}) \oplus \mathcal{C} = \mathcal{A} \oplus (\mathcal{B} \oplus \mathcal{C})$.

4. *$\mathcal{A} \otimes (\mathcal{B} \oplus \mathcal{C}) = (\mathcal{A} \otimes \mathcal{B}) \oplus (\mathcal{A} \otimes \mathcal{C})$.*   *(Distribution of $\otimes$ over $\oplus$)*

5. *The Reduction Properties*

   (a) *$\mathcal{A} \otimes \mathcal{A} \subseteq \mathcal{A}$.*

(b) $\mathcal{A} \oplus \mathcal{A} = \mathcal{A} \otimes \mathcal{A} = \mathcal{A} \oplus \varnothing = \mathcal{A} \otimes \{\varnothing\}$.

(c) *Every $A \in \mathcal{A} \oplus \mathcal{A}$ is a conflict free minimal element of $\mathcal{A}$ if and only if $\mathcal{A} \oplus \varnothing = \mathcal{A}$.*

(d) $\mathcal{A} \otimes \mathcal{A} = (\mathcal{A} \otimes \mathcal{A}) \otimes \mathcal{A}$.

(e) *If $\mathcal{A} = \mathcal{B}$ then $\mathcal{A} \oplus \mathcal{C} = \mathcal{B} \oplus \mathcal{C}$ and $\mathcal{A} \otimes \mathcal{C} = \mathcal{B} \otimes \mathcal{C}$.*

(f) $\mathcal{A} \oplus \mathcal{C} = \mathcal{B} \oplus \mathcal{C}$ *if and only if* $\mathcal{A} \oplus \varnothing = \mathcal{B} \oplus \varnothing$.

(g) $\mathcal{A} \otimes \mathcal{B} = \mathcal{A} \otimes (\mathcal{A} \otimes \mathcal{B})$.

(h) $\{A\} \otimes \{B\} = \varnothing \oplus \{A \cup B\}$.

(i) $\mathcal{A} \oplus \mathcal{B} \subseteq \mathcal{A} \cup \mathcal{B}$.

(j) $\mathcal{A} \otimes (\mathcal{B} \cup \mathcal{C}) = \mathcal{A} \otimes (\mathcal{B} \oplus \mathcal{C})$.

(k) $\mathcal{A} \oplus \varnothing = \varnothing$ *if and only if* $\mathcal{A} = \varnothing$ *or* $\forall A \in \mathcal{A}$, *A is not conflict free.*

(l) $\mathcal{A} \oplus \mathcal{B} = \varnothing$ *if and only if* $\mathcal{A} \oplus \varnothing = \varnothing$ *or* $\mathcal{B} \oplus \varnothing = \varnothing$.

(m) $\mathcal{A} \otimes \mathcal{B} = \varnothing$ *if and only if* $\mathcal{A} \oplus \varnothing = \varnothing$ *or* $\mathcal{B} \oplus \varnothing = \varnothing$ *or* $(\forall A \in \mathcal{A} \oplus \varnothing$ *and* $\forall B \in \mathcal{B} \oplus \varnothing$, $A \cup B$ *is not conflict-free).*

The proof of .0.1. We first show that the stated properties hold for $+, \circ$. Then by using the properties of $+, \circ$ as the reference we show that the properties also hold for $\dot{+}$ and $\dot{\circ}$.

The first property is the direct result of the first condition of definition 2.4.8 and the minimality condition in definition (2.4.8).

The commutativity and the associativity properties are both the direct result of the commutativity and associativity properties of operations $\cup$ and minimality in definitions 2.4.8.

To show the distribution of $\circ$ over $+$, let $D \in \mathcal{A} \circ (\mathcal{B} + \mathcal{C})$. Then $D = A \cup E$ where $E \in \mathcal{B}$ or $E \in \mathcal{C}$. Hence, $D \in \mathcal{A} \circ \mathcal{B}$ or $D \in \mathcal{A} \circ \mathcal{C}$. Conversely, if $D \in (\mathcal{A} \circ \mathcal{B}) + (\mathcal{A} \circ \mathcal{C})$ then $D = A \cup E$ where $E \in \mathcal{B}$ or $E \in \mathcal{C}$. Hence, $E \in (\mathcal{B} + \mathcal{C})$, and, therefore, $D \in \mathcal{A} \circ (\mathcal{B} + \mathcal{C})$.

The proof for reduction properties is straightforward. In regard to the last reduction property, property (k), it is easy to see that if A is a minimal, conflict free element of $\mathcal{B} \cup \mathcal{C}$ then $A \in \mathcal{B} + \mathcal{C}$ is as well. Conversely if A is a minimal conflict free element of $\mathcal{B} + \mathcal{C}$, then, $A \in \mathcal{B} \cup \mathcal{C}$. Hence, we can say that $\mathcal{A} \circ (\mathcal{B} \cup \mathcal{C}) = \mathcal{A} \circ (\mathcal{B} + \mathcal{C})$.

We observe that for any sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$, the set $\mathcal{D}$ of the minimal sets, and, the set $\mathcal{E}$ of the grounded minimal sets in $\mathcal{A} \circ \mathcal{B} \circ \mathcal{C}$, and, their respective counter parts $\mathcal{D}', \mathcal{E}'$ in $(\mathcal{A} \dot{\circ} \mathcal{B}) \dot{\circ} \mathcal{C}$ are equal, i.e. $\mathcal{D} = \mathcal{D}'$ and $\mathcal{E} = \mathcal{E}'$. $\qquad\square$

We can see that both pairs of operations stop short of forming a semiring. The reason is that, neither pair, in general, has a unique absorbing or idempotent element. For instance, any set $\mathcal{A}$ that all of its member sets are not conflict free sets, can act as an absorbing element.

We can rectify this short coming by constraining our domain $\mathcal{R}$ to $\mathcal{R} \subseteq 2^{AR}$ where $\mathcal{R}$ is closed under set union $\cup$, and, for any $\mathcal{A} \in \mathcal{R}$, any $A, B \in \mathcal{A}$, $A \not\rightsquigarrow B$ and $A \not\subset B$. Hence, it is possible to make sure that the operations $+, \circ$ will form a semi-ring over some domain $\mathcal{R}$. Furthermore, to ensure such constraint over 2^{AR}, is applicable with respect to the backings of arguments, we need to make sure that no argument of any cycle is controversial for any argument in AF.

Proofs for chapter 2

The proof of observation (2.2.2).

We observe that all the three identifying properties of a partial order, the reflexive, the anti-symmetric and the transitive property, are all directly entailed from the definitions of the sub-argumentation framework, and, the normal sub-argumentation framework. $\qquad\square$

The proof of observation (2.2.4).

We observe that if a, b are in AF', b an in/direct attacker or defender of a in AF, then, if $c \in \bar{b}$ in AF then c is in AF' as well. $\qquad\square$

The proof of theorem (2.2.5).

Let $AF' = \langle\, AR',\, ATT'\,\rangle$, for an argument a either $a \in AR'$ or $a \notin AR'$. By assumption if $a \notin AR'$ then $a \not\rightarrow S$, hence S accepts S against all $a \notin AR'$. S also accepts S against all $a \in AR'$. Hence, S accepts S against all arguments in AF. $\qquad\square$

Proof of observation (2.3.2).

The proof is trivial. If $a \in \theta(\varnothing)$ then a cannot have any attacker, because $\varnothing$ has no argument to defend a against any attacker. Conversely, if a has no attacker, then a needs no defender, that is to say it is vacuously defended by any argument. $\qquad\square$

The proof of theorem (2.3.6).

The proof of (2.3.6) (Left-to-Right). The assumption is that S is a grounded admissible set.

177

For S, we construct sets A_i, S_i as follows. Let $S_{i+1} = S_i - A_{i+1}$, $A_{i+1} = \{d \mid d \in S_i - B, d \in \widehat{\theta}(S_i - \{d\}, S)\}$ where $B = \widehat{\theta}(\varnothing, S), S_0 = S, A_0 = \varnothing$. We can see that since S is a grounded admissible set, then (1) every S_i is, by definition, a grounded admissible set, and (2) $S_i = \widehat{\theta}(S_{i+1}, S)$. The sequence $S_0, S_1, S_2, \cdots$ is a monotonically decreasing sequence with respect to set inclusion. Hence, there exists some m such that for every $n \geq m$, $S_n = S_{n+1} = B$, $A_{n+1} = \varnothing$. Following (2), we will then have $S = \widehat{\theta}^m(\varnothing, S)$.

The proof of (2.3.6) (Right-to-Left). Let $S_i = \widehat{\theta}^m(\varnothing, S)$. Since by assumption $S_m = S$, we have (1) $S_0 \neq \varnothing$, and, (2) $S_{i+1} = \widehat{\theta}^{i+1}(\varnothing, S) = \widehat{\theta}(S_i, S)$. Hence, every S_i, $0 \leq i \leq m$ is a grounded admissible set, including $S = S_m$. $\qquad\square$

The proof of lemma (2.3.7).

The proof of *(2.3.7.1).* We know that (1) $\mathbf{G} = \theta^\infty(\varnothing)$ ($\mathbf{G}$ is the least fixed point of θ, and, (2) if $R \supseteq T$ is conflict free then $\theta^i(T) \subseteq \theta^i(R)$. For any $T \subseteq \theta(\varnothing)$, we therefore have $\theta^\infty(T) \subseteq \mathbf{G}$. Hence, by theorem (2.3.6), the grounded extension $\mathbf{G}$ is the maximum grounded admissible set.

The proof of (2.3.7.2) is trivial. Let $\mathcal{A}$ be the set of $T \subseteq \mathbf{G}$ such that $S \subseteq T$ and T is a grounded admissible set. Following the first part of this lemma, since $\mathbf{G} \in \mathcal{A}$, then $\mathcal{A} \neq \varnothing$. Now since $\mathcal{A}$ is a finite, it must have a minimal element, satisfying the intended result. $\qquad\square$

The proof of lemma (2.4.2).

Let $S' = S \cap D_a$ where S be a minimal admissible set accepts a. S' accepts a because if $b \in S$ is a defender of a then $b \in S'$. S' is admissible, because, (1) for any $b \in S'$ if $c \hookrightarrow b$ then there is some $d \in S$ such that $d \hookrightarrow c$, and (2) if $b \in S$ and $d \in D_b$ then $d \in S'$. Hence, we conclude that $S = S' \subseteq D_a$.

Similarly, let S be a minimal admissible set that attacks a, and, $S' = S \cap D_{\bar{a}}$. Then, (1) there is some $b \in S'$ such that $b \hookrightarrow a$, and, (2) S' is an admissible set that accepts b (because, for

any $b \in S'$ if $c \hookrightarrow b$ then there is some $d \in S'$ such that $d \hookrightarrow c$.) Hence, we conclude that $S = S' \subseteq D_{\bar{a}}$. $\qquad\square$

The proof of lemma (2.4.4).

The proof is by contradiction. Let $S \in \langle a \rangle^+$ and assume that there is some $c \in S$ that is not a critical defender of any argument in $S \cup \{a\}$ with respect to S. Let $T = S - \{c\}$, then by assumption, if $b \hookrightarrow W$ where $W = T \cup \{a\}$ then $T \hookrightarrow c$. T is therefore an admissible set that accepts a, and, $T \subset S$. But, this contradicting the assumption that S is a minimal set that accepts a. $\qquad\square$

The proof of theorem (2.4.6).

The proof of (2.4.6.1) If AF has stable extension E, then, vy definition, for an argument a, either $a \in E$ or $E \hookrightarrow a$. If $a \in E$ then there is some S minimal subset of E that accepts a, and, if $E \hookrightarrow b$ then there is some minimal subset of E that attacks a. Hence, $\langle a \rangle^+ \neq \varnothing$ or $\langle a \rangle^- \neq \varnothing$.

The proof of (2.4.6.2) is trivial.

The proof of (2.4.6.3). Let $a \in S$ for some $S \in \langle a \rangle^-$ then by lemma (2.4.2), $a \in AT_a$. But, this contradicts the assumption that AF is a rational argumentation framework.

The proof of (2.4.6.4). Let $S \in \langle a \rangle^+$ and $b \in \bar{a}$. By definition 2.4.1, S is both admissible and attacks b. Thus, there is some minimal admissible set $S' \subseteq S$ that $S' \hookrightarrow b$. Hence, $S' \in \langle b \rangle^-$.

The proof of (2.4.6.5). Let $\langle a \rangle^- = \varnothing$ then by definition 2.4.1, there is no admissible set that attacks a. Hence, by definition a belongs to all non empty preferred extensions. On the other hand, if $a \in E$ for some preferred extension E. Then, there is some minimal subset of E with respect to set inclusion that accepts a. Hence, $\langle a \rangle^- \neq \varnothing$ and $S \in \langle a \rangle^-$. We can therefore say that (2.4.6.5a) and (2.4.6.5b) are equivalent.

Next, if $\langle a \rangle^- = \varnothing$ then for all $b \in \overline{a}$, $\langle b \rangle^+ = \varnothing$. If not, then for some $b \in \overline{a}$, $\langle b \rangle^+ \neq \varnothing$ then there is some minimal admissible set S that accepts b, and, therefore attacks a, contradicting the initial assumption. On the other hand, if for all $b \in \overline{a}$, $\langle b \rangle^+ = \varnothing$ then by definition there is no admissible set that attacks a (if a set S attacks a then S must have some attacker of a). We can therefore say that (2.4.6.5a) and (2.4.6.5c) are equivalent.

The proof of (2.4.6.6). If $\langle a \rangle^+ = \varnothing$, then there is no admissible set that accepts a. Thus, there is no preferred extension that accepts a. Conversely, if no preferred extension accepts a then $\langle a \rangle^+ = \varnothing$. Hence, we can say that (2.4.6.6a) and (2.4.6.6c) are equivalent.

Let $\langle b \rangle^- = \varnothing$ for some $b \in \overline{a}$, then by (2.4.6.5), no admissible set attacks b. Hence, no admissible set accepts a which means $\langle a \rangle^+ = \varnothing$. Next, let us assume (1) for all $b \in \overline{a}$, $\langle b \rangle^- \neq \varnothing$, and, (2) there is no conflict free set S such that if $b \in \overline{a}$ then there is $T \in \langle b \rangle^-$ such that $S_i \subseteq S$. By (2.4.6.4), if $S \in \langle a \rangle^+$ then for each $b \in \overline{a}$, there is some $T \in \langle b \rangle^-$ such that $T \subseteq S$. However, by assumption there is no such conflict free set, hence, $\langle a \rangle^+ = \varnothing$. $\qquad (\star)$

On the other hand, Let $\langle a \rangle^+ = \varnothing$, then by (2.4.6.1) $\langle a \rangle^- \neq \varnothing$. Let $S \in \langle a \rangle^-$ then there is some $b \in \overline{a}$, $b \in S$, $\langle b \rangle^+ \neq \varnothing$. Let B denote the set of all such b. It is either (1) for some $b \in B$, $\langle b \rangle^- = \varnothing$, or (2) for all $b \in B$, $\langle b \rangle^- \neq \varnothing$. Let it be the case (2) where for all $b \in B$, $\langle b \rangle^- \neq \varnothing$. Let us assume that there is a conflict free set W such that for all $b \in B$ there is some $T \in \langle b \rangle^-$, $T \subseteq W$. $\qquad (\star\star)$

Let, $R = \bigcup_{T \subseteq W} T$. R is an admissible set because any argument in R is acceptable by some $T \subseteq R$. R also accepts a, because, for any $b \in \overline{a}$, $R \hookrightarrow b$. Hence, $\langle a \rangle^+ \neq \varnothing$, but this contradicts the initial assumption that $\langle a \rangle^+ = \varnothing$. Hence, the $(\star\star)$ is incorrect, and there is no such presumed conflict free set T. $\qquad (\star\star\star)$

From the above results $(\star\star),(\star\star\star)$ we conclude that (2.4.6.6a) and (2.4.6.6b) are equivalent as well. $\qquad \square$

The Proof of theorem (2.4.9).

First we shoe that $\langle\,a\,\rangle^+ \,\dot\circ\, \langle\,a\,\rangle^- = \varnothing$. If either $\langle\,a\,\rangle^+ = \varnothing$ or $\langle\,a\,\rangle^- = \varnothing$ then $\langle\,a\,\rangle^+ \,\dot\circ\, \langle\,a\,\rangle^- = \varnothing$ by defult. So, let us assume that both are not nul. If so, then, the every $S \in \langle\,a\,\rangle^+ \,\dot\circ\, \langle\,a\,\rangle^-$ both accepts and attacks a. Hence, S is a self attackiing set. But, this means, $\langle\,a\,\rangle^+ \,\dot\circ\, \langle\,a\,\rangle^- = \varnothing$.

The proof of second part of theorem is self evident. In framework with a stable extension, at last one admissible set either accepts or attacks an argument. Hence, $\langle\,a\,\rangle^+ \,\dot+\, \langle\,a\,\rangle^- \neq \varnothing$. $\qquad\square$

The Proof of observation (2.5.2).

The proof is self evident. It directly follows from definition (2.5.1). $\qquad\square$

The Proof of theorem (2.5.3).

The proof of (2.5.3.1) where $\langle\,a\,\rangle^- = \varnothing$. From the results (.0.1.6h), (.0.1.6i) in appendix 5.2.2, we have $\sum_{b\in\bar a}^{\cdot} (\,\{\{b\}\} \,\dot\circ\, \langle\,b\,\rangle^+\,) = \varnothing$ if and only if $(\forall b \in \bar a)(\langle\,b\,\rangle^+ = \varnothing)$. Using the results in (2.4.6.5a),(2.4.6.5b), we then have $\langle\,a\,\rangle^- = \varnothing$ if and only if $\sum_{b\in\bar a}^{\cdot} (\,\{\{b\}\} \,\dot\circ\, \langle\,b\,\rangle^+\,) = \varnothing$.

The proof of (2.5.3.1) from left to right where $\langle\,a\,\rangle^- \neq \varnothing$. Let $\mathcal{Z}^- = \sum_{b\in\bar a} (\{\{b\}\} \circ \gamma(\langle\,b\,\rangle^+))$. We want to show that if $S \in \langle\,a\,\rangle^-$ then $S \in \gamma^{-1}(\mathcal{Z}^-)$. Let $S \in \langle\,a\,\rangle^-$, then by definition (2.4.1), there is some $b \in S$ such that $b \in \bar a$. Let S_b be defined as $S_b = S - \{b\}$ if b is acceptable by $S - \{b\}$, otherwise, $S_b = S$.

We claim $S_b \in \langle\,b\,\rangle^+$. If not then there must exist some $S_b' \in \langle\,b\,\rangle^+$ such that $S_b' \subset S_b$. However, this contradicts the minimality assumption in $S \in \langle\,a\,\rangle^-$, because $S' = S_b' \cup \{b\}$, $S' \hookrightarrow a$ but $S' \subset S$. Having $S_b \in \langle\,b\,\rangle^+$, then $S = S_b \cup \{b\} \in \gamma^{-1}(\mathcal{Z}^-)$. Because, if $S \notin \mathcal{Z}^-$ then by definition of $+$, there is some $c \in \bar a$ such that $\{c\} \cup S_c \subset S$ for some $S_c \in \gamma^{-1}(\langle\,c\,\rangle^+)$. Hence, the set $\{c\} \cup S_c$ is an admissible set that attacks a. But this contradicts the assumption that $S \in \gamma^{-1}(\langle\,a\,\rangle^-)$. Thus, $S \in \mathcal{Z}^-$, which means $\langle\,a\,\rangle^- \subseteq \gamma^{-1}(\mathcal{Z}^-)$.

The proof of (2.5.3.1) from right to left where $\langle\,a\,\rangle^- \neq \varnothing$. Let $S \in \mathcal{Z}^-$ then $S = \{b\} \cup S_b$ where $b \in \bar a$, $S_b \in \langle\,b\,\rangle^+$, $\langle\,b\,\rangle^+ \neq \varnothing$. By definition of $+, \circ$, S is a minimal admissible set that attacks

a. Hence, $S \in \langle a \rangle^-$ which means $\gamma^{-1}(\mathscr{Z}^-) \subseteq \langle a \rangle^-$.

The proof of (2.5.3.2) where $\langle a \rangle^+ = \varnothing$. By theorem (.0.1.6i), we have $\prod_{b \in \bar{a}} \langle b \rangle^- = \varnothing$ if and only if $(\exists b \in \bar{a})(\langle b \rangle^- = \varnothing)$ or for every selection of $A_1 \in \langle b_1 \rangle^+, A_2 \in \langle b_2 \rangle^+, \cdots, A_n \in \langle b_n \rangle^+$, $\bar{a} = \{b_1, \cdots, b_n\}$, the set $\bigcup A_i$ is not a conflict- free set. Hence, by using the results in (2.4.6.6a), (2.4.6.6b), we have $\langle a \rangle^+ = \varnothing$ if and only if $\prod_{b \in \bar{a}}^{\cdot} \langle b \rangle^- = \varnothing$.

The proof of (2.5.3.2) where $\langle a \rangle^+ \neq \varnothing$. If a is a ground argument, the claim $\langle a \rangle^+ = \{\varnothing\}$ holds by default. Hence, we assume that a is not a ground argument. In the following $\mathscr{Z}^+$ is defined as $\mathscr{Z}^+ = \prod_{b \in \bar{a}} \gamma(\langle b \rangle^-)$.

The proof of (2.5.3.2) from left to right, where $\langle a \rangle^+ \neq \varnothing$, *and,* a *is not a ground argument.* Let $S \in \langle a \rangle^+$ then S is the minimal admissible that for any $b \in \bar{b}$, there is some $S_b \in \langle b \rangle^-$ such that $S_b \subseteq S$. Hence, by definition of $\prod$, $S \in \gamma^{-1}(\mathscr{Z}^+)$, and so, $\langle a \rangle^+ \subseteq \mathscr{Z}^+$.

The proof of (2.5.3.2) from right to left, where $\langle a \rangle^+ \neq \varnothing$, *and,* a *is not a ground argument.* Let $S \in \gamma^{-1}(\mathscr{Z}^+)$ then by definition S is the minimal admissible set that attacks every $b \in \bar{a}$. Hence, S is a minimal admissible set that accepts a, and so $S \in \langle a \rangle^+$ which in turn means $\gamma^{-1}(\mathscr{Z}^+) \subseteq \langle a \rangle^+$.

The proof of (2.5.3.3) The proof directly follows from observation (2.5.2). □

The proof of theorem (2.5.4).

The proof of (2.5.4.1). Let S be a minimal element of either $\mathcal{H}$ or $\mathcal{H} \cap \mathcal{G}$. Then S should be minimal element that for all $a \in A$, there is some $S_a \in \langle a \rangle^+$ such that $S_a \subseteq S$. If so, then, $S \in \gamma^{-1}(\prod_{a \in A} \gamma(\langle a \rangle^+))$. Conversely, let $S \in \gamma^{-1}(\prod_{a \in A} \gamma(\langle a \rangle^+))$ then by definition, there is no proper subset of S that accepts every $a \in A$. Hence, S is either a minimal element of $\mathcal{H}$ or a minimal element of $\mathcal{H} \cap \mathcal{G}$ or both.

The proof of (2.5.4.3). The results (2.5.4.3) and (2.5.3.2) are essentially the same result that are said in different context. Hence, the proof for (2.5.4.3) is the same as the proof for (2.5.3.2).

The proof of (2.5.4.2) and (2.5.4.4). The proof directly follows from property of operation $\sum$ that $\sum_{\mathcal{A}_i \in \mathbf{A}} \mathcal{A}_i = \varnothing$ if and only if every $\mathcal{A}_i = \varnothing$. $\qquad\square$

The proof of theorem (2.5.6).

To show (2.5.6), we first introduce a number of notations, and terms. In relation to definition 2.5.5, we represent the members in $\mathbb{D}$ and $2^{\mathbb{D}}$ by $\vec{d} \in \mathbb{D}$, and $\vec{D} \in 2^{\mathbb{D}}$, and, denote the elements of $\vec{d} = (d, T, j)$ by

$$\lfloor \vec{d}{:}a \rfloor = d, \ \lfloor \vec{d}{:}S \rfloor = T, \ \lfloor \vec{d}{:}j \rfloor = j.$$

In the same manner, we denote the corresponding elements for a set $\vec{D} \subseteq \mathbb{D}$ by

$$\lfloor \vec{D}{:}a \rfloor = \{ d \mid d = \lfloor \vec{d}{:}a \rfloor, \vec{d} \in \vec{D} \}, \ \text{ and etc.}$$

We define $\bar{j}$ as the complement of j in $\{0, 1\}$ as

$$\bar{j} = 1 \quad \text{if and only if} \quad j = 0.$$

From the definition (2.5.5), we see that the recursion in backing function traverses form each $ß(\vec{b})$ to the next $ß(\vec{c})$ such that $\vec{b} = (b, T_b, j)$, $\vec{c} = (c, T_b \cup \{b\}, \bar{j})$, $c \in \bar{b}$. To capture this relation between $\vec{b}$, $\vec{c}$, we extend the notation $\overline{d}$, defined for arguments $d \in AR$, to the members $\vec{d} = (d, T, j) \in \mathbb{D}$.

$$\lfloor \overline{\vec{d}} \rfloor = \{ (b, T', j') \mid b \in \bar{d}, d \notin T, T' = T \cup \{d\}, j' = \bar{j} \}.$$

Next, to map the recursion path, for a given $\vec{d} \in \mathbb{D}$, we construct tree-like structures called *γ-structures*. A $γ$-structure, $γ_{\vec{d}}$ for a given $\vec{d} \in \mathbb{D}$ is a directed graph such that

1. the nodes of $γ_{\vec{d}}$ belong to $\mathbb{D}$ where $\vec{d}$ is the root node,

2. if $\vec{c}$, $\lfloor \vec{c}{:}j \rfloor = 1$ is a node of $γ_{\vec{d}}$ then $\vec{b} \in \lfloor \overline{\vec{c}} \rfloor$ is the child node of $\vec{c}$,

3. if $\vec{b}$, $\lfloor \vec{b}{:}j \rfloor = 0$ is a node of $\gamma_{\vec{d}}$ where $\lfloor \overline{\vec{b}} \rfloor \neq \varnothing$, then, $\vec{b}$ has one and only one child node $\vec{c} \in \lfloor \overline{\vec{b}} \rfloor$ in $\gamma_{\vec{d}}$.

The γ-structure is almost a tree structure, except maybe at the leaf nodes. The reason is that the child nodes of distinct parents have distinct tracking set $\lfloor \vec{c}{:}S \rfloor$. Hence, no node has more than one parent except maybe at the leaf nodes by which a cycle is denoted.

For the reference purposes, we call the descendant nodes $\vec{b}$ of a node $\vec{c}$, the *in/direct defender* of $\vec{c}$ in $\gamma_{\vec{d}}$ if $\lfloor \vec{b}{:}j \rfloor = \lfloor \vec{c}{:}j \rfloor$ and the *in/direct attackers* of $\vec{c}$, otherwise. (note: for a node $\vec{c}$ to regarded as an in/direct defender or attacker of some node $\vec{b}$ in $\gamma_{\vec{d}}$, $\vec{c}$ needs to be a descendant node of $\vec{b}$ in $\gamma_{\vec{d}}$.)

We observe that for each $\vec{d} \in \mathbb{D}$, we can reconstruct the $ß(\vec{d})$ from its γ-structures $\gamma_{\vec{d}}$. To do that, for each γ-structure we define a function ψ_γ as follows, where $\vec{B}$ denotes the set of child nodes of $\vec{c}$.

1. If $\vec{c}$ is a leaf-node of γ then $\psi_\gamma(\vec{c}) = ß(\vec{c})$,

2. otherwise

   (a) if $\lfloor \vec{c}{:}j \rfloor = 1$ then $\psi_\gamma(\vec{c}) = \prod\limits_{\vec{b} \in \vec{B}} \psi_\gamma(\vec{b})$,

   (b) if $\lfloor \vec{c}{:}j \rfloor = 0$ then $\psi_\gamma(\vec{c}) = \sum\limits_{\vec{b} \in \vec{B}} \{\lfloor \vec{b}{:}a \rfloor\} \circ \psi_\gamma(\vec{b})$.

Let $\vec{d}$ be of some $\vec{d} \in \mathbb{D}$ and $\Gamma_{\vec{d}}$ denote the set of all γ-structures, $\gamma_{\vec{d}}$, of $\vec{d}$. Then, by a simple application of induction we can show that

$$ß(\vec{d}) = \sum_{\gamma_{\vec{d}} \in \Gamma_{\vec{d}}} \psi_{\gamma_{\vec{d}}} \tag{1}$$

We observe that the equation (1) holds for $\vec{c}$ that are the leaf nodes the γ-structures γ_d. Next, we suppose that the equation (1) holds for all $\vec{b} \in \overline{\vec{d}}$. If $\lfloor \vec{d}{:}j \rfloor = 0$ then (1) holds, straight from the definition of backing function. Next, to draw $ß\vec{d}$ where $\lfloor \vec{d}{:}j \rfloor = 1$, we substitute every $ß(\vec{b})$,

$\vec{b} \in \vec{d}$ with its equivalent from the equation (1).

$$\text{\ss}(\vec{d}) = \prod_{\vec{b} \in \vec{d}} \text{\ss}(\vec{b}) = \prod_{\vec{b} \in \vec{d}} \left(\sum_{\gamma_{\vec{b}} \in \Gamma_{\vec{b}}} \psi_{\gamma_{\vec{b}}} \right)$$

By the distributive property of $\circ$ over $+$ with then have,

$$\text{\ss}(\vec{d}) = \sum_{\Lambda \in \Lambda_{\vec{d}}} \left(\prod_{\gamma_{\vec{b}} \in \Lambda} \psi_{\gamma_{\vec{b}}} \right)$$

where $\Lambda_{\vec{d}}$ is the set of collections Λ such that for each $\vec{b} \in \vec{d}$ there is one and only one $\gamma_{\vec{b}}$ in Γ. On the other hand, by definition of the γ-structure, any $\gamma_{\vec{d}}$ is simply a collection of $\gamma_{\vec{b}} \in \Lambda$ with the root node $\vec{d}$. Hence, we can say,

$$\text{\ss}(\vec{d}) = \sum_{\gamma_{\vec{d}} \in \Gamma_{\vec{d}}} \psi_{\gamma_{\vec{d}}}$$

For a leaf node $\vec{c}$, the $\psi(\vec{c})$ is a singleton set. Moreover, for a node $\vec{c}$, if $\psi_{\gamma}(\vec{b})$ of all its child nodes $\vec{b}$ are singleton sets, then, $\psi_{\gamma}(\vec{c})$ is a singleton set, too. Hence, by a simple application of induction, we can say,

$$\psi_{\gamma_{\vec{d}}}(\vec{d}) = \varnothing \text{ or } \psi_{\gamma_{\vec{d}}}(\vec{d}) = \{\{W\}\}. \tag{2}$$

From the definition of ψ-function, We observe that,

$$\text{if } \lfloor \vec{d} : j \rfloor = 1 \text{ then } \psi_{\gamma_{\vec{d}}} = \varnothing \text{ iff for some leaf-node } \vec{c}, \ \psi_{\gamma_{\vec{d}}}(\vec{c}) = \varnothing. \tag{3}$$

Next, to find the set W, let $\vec{W}^1$, $\vec{W}^0$ each denote the set of in/direct defending and in/direct attacking nodes of $\vec{d}$ in $\gamma_{\vec{d}}$, then,

$$\text{if } \psi_{\gamma}(\vec{d}) \neq \varnothing \text{ then } W - \{\star\} = \{\lfloor \vec{W} : a \rfloor\}, \text{ where}$$

$$\vec{W} = \vec{W}^1 \text{ if } \lfloor \vec{d} : j \rfloor = 1, \text{ and, } \vec{W} = \vec{W}^0 \text{ otherwise.} \tag{4}$$

Next, in order to find the minimal $\psi_{\gamma_{\vec{d}}}(\vec{d})$, we define the class of $\Gamma_{\vec{d}}^{M}$ in $\Gamma_{\vec{d}}$. The class $\Gamma_{\vec{d}}^{M}$ is the

set of all $\gamma_{\vec{d}}^{M}$ such that every node $\vec{c}$, $\lfloor \vec{c}:j \rfloor = 1$ in $\gamma_{\vec{d}}^{M}$ is the unique child of some parent in $\gamma_{\vec{d}}^{M}$,

unless it is $\vec{d}$. (To characterize $\gamma_{\vec{d}}^{M}$ we employed a variation of the notion *critical defender*.) We

observe that, for each $\gamma_{\vec{d}}$ there is some $\gamma_{\vec{d}}^{M}$ in $\Gamma_{\vec{d}}$ such that $\gamma_{\vec{d}}^{M}$ is a subgraph of $\gamma_{\vec{d}}$. Moreover, for

each $\gamma_{\vec{d}}$ and its corresponding $\gamma_{\vec{d}}^{M}$, if $\gamma_{\vec{d}} \neq \varnothing$ then $\psi_{\gamma_{\vec{d}}^{M}}(\vec{d}) \subseteq \psi_{\gamma_{\vec{d}}}(\vec{d})$. Hence, following theorem

(.0.1), the reduction properties for $\sum$, we can say,

$$\text{ß}(\vec{d}) = \sum_{\gamma_{\vec{d}} \in \Gamma_{\vec{d}}^{M}} \psi_{\gamma_{\vec{d}}} \tag{5}$$

The proof of (2.5.6.1) directly follows from the fact that the tracking set T in $\text{ß}(d, T, j)$ is

monotonically increasing with the the supremum AR.

To prove (2.5.6.2.1), we need to show,

$$\langle a \rangle^{+} = \{ S - \{\star\} \mid S \in \text{ß}(a, \varnothing, 1) \}. \tag{6}$$

$$\langle a \rangle^{-} = \{ S - \{\star\} \mid S \in \text{ß}(a, \varnothing, 0) \}. \tag{7}$$

The proof of (7) directly follows from (6). We first show (6). The trivial case is when a is a

ground argument for which by definition $\text{ß}(a, \varnothing, 0) = \{\varnothing\}$. Hence, $S \in \text{ß}(a, \varnothing, 0)$ if and only

if $S = \varnothing$, where S is the minimal set that accepts a.

Next, we proceed with the non-trivial case where $\overline{a} \neq \varnothing$. Let $\vec{a} = (a, \varnothing, 1)$. To show, if

$S \in \langle a \rangle^{+}$ then $S \in \text{ß}(\vec{a})$, for every $S \in \langle a \rangle^{+}$ we construct a γ-structure for $\vec{a}$, denoted by γ_{S},

such that

$$\text{if } S \in \langle a \rangle^{+} \text{ then } S = W - \{\star\}, \; W = \psi_{\gamma_{S}}(\vec{a}) \in \text{ß}(\vec{a}). \tag{8}$$

In order to construct γ_{S}, we define a function μ from $Z = \bigcup_{c \in R} \overline{c}$ to S where $R = S \cup \{a\}$.

Since, $S \in \langle a \rangle^{+}$, there exists such function μ where by definition for every $b \in Z$ there is only

one $\mu(b) \in S$. We construct γ_{S} as follows. Let $\vec{a}$ be the root node. If $\vec{c} = (c, T_{c}, 1)$ is a node of

γ_S then $\vec{b} = (b, T_b \cup \{c\}, 0)$, $b \in \overline{(c)}$ is a child node of $\vec{c}$ in γ_S. If $(b, T_b, 0)$ is a node of γ_S then $(\mu(b), T_b \cup \{b\}, 1$ is the child node of $\vec{b}$ in γ_S. It is easy to check that γ_S is a minimal γ-structure for $\vec{a}$. Hence, by (17), (18) we have $S = \mu(R) = W - \{\star\} = \psi_{\gamma_S}(\vec{a}) \in \text{ß}(\vec{a})$.

We next show that the converse of (19) holds as well, that is,

$$\text{if } S \in \text{ß}(\vec{a}) \text{ then } S - \{\star\} \in \langle\, a\,\rangle^+. \tag{9}$$

If $S \in \text{ß}(\vec{a})$ then we know that there is a γ-structure, $\gamma_{\vec{a}}^M$, where $S = \psi_{\gamma_{\vec{a}}^M}(\vec{a})$. By result (3), if $W \in \psi_{\gamma_{\vec{a}}}$, then, for all leaf-nodes $\vec{c}$ of $\gamma_{\vec{a}}^M$, $\text{ß}(\vec{c}) \neq \varnothing$. Thus, for every in/direct defending node $\vec{d}$ of $\vec{a}$ in $\gamma_{\vec{a}}^M$, $d = \lfloor \vec{d} : a \rfloor$ is accepted by $\lfloor \vec{W} : a \rfloor$ where $\lfloor \vec{W} : a \rfloor$ is the set of in/direct defending nodes of $\vec{a}$ in $\gamma_{\vec{a}}^M$. Hence, $S \in \text{ß}(\vec{a})$ is an admissible set that accepts a. Moreover, since any $S \in \text{ß}(\vec{a})$ is a minimal element of $\text{ß}(\vec{a})$, S is a minimal admissible set that accepts a. If S is not minimal then there is some $S' \in \langle\, a\,\rangle^+$, $S' \subset S$ where by (19) $S' \in \text{ß}(\vec{a})$, contradicting that S is the minimal element of $\text{ß}(\vec{a})$. We can therefore conclude that the result (21) holds. The results (19), (21) conclude the proof of (2.5.6.2.1).

We next show that the second claim of (2.5.6.2) holds, that is if $S \in \text{ß}(\vec{a})$ and $\star \in S$ then $W = S - \{\star\}$ is not a grounded backing of a.

The proof is by contradiction. Let $\star \in S$, $S \in \text{ß}(\vec{a})$, $W = S - \{\star\}$ and W be a grounded admissible set. For W we construct an argumentation framework $AF_W = \langle\, AR_W\,,\, ATT_W\,\rangle$ where $AR_W = W \cup \overline{W}$ and ATT_W is selected as follows.

If W is a grounded admissible set then by lemma (2.3.6) we can define a strict partial order $\prec$ on W such that $d \prec c$ if $d \in \widehat{\theta}^{\,i}(\varnothing, S)$, $c \notin \widehat{\theta}^{\,i}$ and d is an in/direct defender of c.

Next, we define ATT_W such that $(a, b) \in ATT_W$ if and only if (1) $a \in AR_W, b \in AR_W, (a, b) \in ATT$, (2) if there is some $(c, b) \in ATT$, $c \in AR_W$ then $a \prec b$, and, (3) for every in/direct attacker b of a there is only one $c \in AR_W$ such that $(c, b) \in ATT_W$.

From the definition of AF_W, We observe that (1) W is the backing of a in AF_W, (2) there are no attack cycles in AF_W, and, (3) we can build a γ-structure, γ_W, for $\vec{a} = (a, \varnothing, 1, a)$ such that for every parent-child $(\vec{b}, \vec{c})$ in γ_W, $(c, b) \in ATT_W$ where $c = \lfloor \vec{c} : a \rfloor$, $b = \lfloor \vec{b} : a \rfloor$, and, vice versa.

Since AF_W contains no attack cycles, we have, $\psi_{\gamma_W} = \{W\}$, $W \in \text{ß}(\vec{a})$. This, however, leads to a contradiction, because, both $W, S \in \text{ß}(\vec{a})$ and $W \subset S$. Hence, W cannot be a grounded admissible set.

Thus far, we proved (2.5.6.1), (2.5.6.2.2) and $\langle a \rangle^+ = \{S - \{\star\} \mid S \in \text{ß}(a, \varnothing, 1)\}$, the first part of (2.5.6.2.1). To complete the proof of (2.5.6), we are left to show the second part of (2.5.6.2.1), which is $\langle a \rangle^- = \{S - \{\star\} \mid S \in \text{ß}(a, \varnothing, 0)\}$.

By theorem (2.5.3), we know that,

$$S \in \langle a \rangle^- \text{ if and only if } S \in \sum_{b \in \bar{a}} \{\{b\}\} \circ \langle b \rangle^+.$$

By the results shown so far, we also know that,

$$\sum_{b \in \bar{a}} \{\{b\}\} \circ \langle b \rangle^+ = \{ S - \{\star\} \mid S \in \sum_{b \in \bar{a}} \{\{b\}\} \circ \text{ß}(b, \varnothing, 1) \}.$$

Hence, if we show that,

$$\sum_{b \in \bar{a}} \{\{b\}\} \circ \text{ß}(b, \varnothing, 1) = \sum_{b \in \bar{a}} \{\{b\}\} \circ \text{ß}(b, \{a\}, 1), \tag{10}$$

then, we can claim the result (7),

$$\langle a \rangle^- = \{ S - \{\star\} \mid S \in \sum_{b \in \bar{a}} \{\{b\}\} \circ \text{ß}(b, \varnothing, 1) \}$$
$$= \{ S - \{\star\} \mid S \in \sum_{b \in \bar{a}} \{\{b\}\} \circ \text{ß}(b, \{a\}, 1) \}$$
$$= \{ S - \{\star\} \mid S \in \text{ß}(a, \varnothing, 0) \}.$$

We observe that every γ-structure, $\gamma_{\vec{b}*}$, for $\vec{b}* = (b, \{a\}, 1)$ is a subgraph of some γ-structure, $\gamma_{\vec{b}}$, for $\vec{b} = (b, \varnothing, 1)$. Let $\gamma_{\vec{b}*}$, be a γ-structure with a leaf node $\vec{a'}$, $\lfloor \vec{a'} : a \rfloor = a$. Since $b \in \bar{a}$, there is some γ-structure, $\gamma_{\vec{b}}$, that is identical to $\gamma_{\vec{b}*}$ except that $\vec{a'}$ in $\gamma_{\vec{b}}$, now has a leaf child node $\vec{b'}$, $\lfloor \vec{b'} : a \rfloor = b$. It then follows that,

$$\{\{b\}\} \circ \psi_{\gamma_{\vec{b}*}} = \{\{b\}\} \circ \psi_{\gamma_{\vec{b}*}}.$$

Hence, we can say,

$\forall b \in \bar{a}$,

1. $\{\{b\}\} \circ \text{ß}(b, \{a\}, 1) \subseteq \{\{b\}\} \circ \text{ß}(b, \varnothing, 1)$,

2. if $W \in (\{\{b\}\} \circ \text{ß}(b, \varnothing, 1) - \{\{b\}\} \circ \text{ß}(b, \{a\}, 1))$

   then $\exists c, \exists R, c \in \bar{a}, \ R \in \{\{c\}\} \circ \text{ß}(c, \{a\}, 1)$

   such that $R \in \mathcal{G}, R \subseteq W$. $\hspace{4cm}$ (11)

Following (11), we can then conclude that,

$$\sum_{b \in \bar{a}} \{\{b\}\} \circ \text{ß}(b, \varnothing, 1) = \sum_{b \in \bar{a}} \{\{b\}\} \circ \text{ß}(b, \{a\}, 1). \tag{12}$$

and so, we have completed the proof of theorem (2.5.6). $\hspace{2cm}$ $\square$

The proof of observation (2.6.1).

The proof for (2.6.1) directly follows from theorm (2.2.5). Let AF', AF'' be as defined in (2.2.5). From one hand, AF does not contain any members of S. On the other hand, $\text{ß}(a, \varnothing, j)$ remains unchanged in all $AF'' \sqsubseteq AF'$. Hence, we can conclude that any $S' \subseteq S$ has no impact on $\text{ß}(a, \varnothing, j)$, and therefore, $\text{ß}(a, S, j) = \text{ß}(a, \varnothing, j)$. $\hspace{2cm}$ $\square$

The proof of theorem (2.6.2).

The proof for (2.6.2) and (2.6.1.2), both have the same underlying principle which closely

resembles the proof of (2.6.1). However, they instead, directly follows from the theorem (2.6.3).

Let $\mathcal{AF}_a$ be the class of all minimal argumentation frameworks AF_S in (2.6.2.1) for all backings S of the argument a. Let $AF_a = \langle AR_a, ATT_a \rangle$ be the least upper bound sub-argumentation framework in AF. We observe that AF_a is the normal sub-argumentation framework in AF such that $AR_a = W \cup \overline{W}$ where $W = \bigcup_{R \in \mathcal{S}} S$ and $\mathcal{S} = \langle a \rangle^+ \cup \langle a \rangle^-$. Hence, AF contains only the members of backings of a or any argument that attacks those members.

By assumption, for sets S, T, we have $S \cap AR_a = \varnothing$ and $T \cap AR_a = \varnothing$. Let $AF' \sqsupseteq AF_a$ and $S' \subseteq S$, $T' \subseteq T$. From one hand, AF_a does not contain any members of S or T. On the other hand, by theorem (2.6.3), $ß(a, \varnothing, j)$ remains unchanged in all $AF' \sqsupseteq AF_a$. Hence, we can conclude that any $S' \subseteq S$ has no impact on $ß(a, \varnothing, j)$, and therefore, $ß(a, S, j) = ß(a, \varnothing, j)$. For the same reason, if both $ß(a, S, j) = ß(a, \varnothing, j)$ and $ß(a, T, j) = ß(a, \varnothing, j)$ then $ß(a, S \cup T, j) = ß(a, \varnothing, j)$. $\square$

The proof of theorem (2.6.3).

The proof of (2.6.3.1). Let S be some backing of a, and, $\mu \colon \overline{S} \longrightarrow S$ some function for S such that if $c = \mu(b)$ then $(c, b) \in ATT$. It can be seen that since S is a backing for a, for every backing S, there is some function μ, where μ, is an onto function that selects exactly one attacker in S, for every $b \in \overline{S}$.

We then construct the argumentation framework $AF_S = \langle AR_S, ATT_S \rangle \sqsubseteq AF$ for S, such that $AR_S = S \cup \overline{S}$ and $(c, b) \in ATT_S$ if and only if $(c, b) \in ATT$, $c \in AR_S, b \in AR_S$ and if $c \in S$, $b \in \overline{S}$ then $c = \mu(b)$.

The claim is that this AF_S is the intended minimal $AF_S \sqsubseteq AF$. We observe that for any such $AF_i \sqsupseteq AF_S$, for all b in AF_i, if b in AF_i attacks some $d \in AR_S$ then by the initial assumption that S is admissible in AF, there is some $c \in AR_S$ that $c \hookrightarrow b$. Hence, S is admissible in all such AF_i.

To show that AF_S is such minimal sub-argumentation framework, we show that for any $AF_j = \langle AR_j, ATT_j \rangle$ where $AR_j \subset AR_i$ or $ATT_j \subset ATT_S$, S is not admissible in AF_j. By definition of the function μ, for every attacker b of S, there is only one $(c, b) \in ATT_S$. Hence, if S is admissible in AF_j then $ATT_j \not\subset ATT_S$.

$AR_j \subset AR_S$ is not possible either. First, since S is a backing of a in AF, then $S \subseteq AR_j$. Consequently, if $AR_j \subset AR_S$ then there is some $b \in AR_S, b \notin AR_j, b \hookrightarrow S$. However, if b is not a member of AF_j then by the definition of sub- argumentation frameworks and the definition of function μ_S, no argument in AF_j attacks b. But if so, we can then construct an argumentation framework $AF_j \sqsubset AF_k$ where AF_k is the same as AF_j, except for b. However, S is no longer admissible in AF_k as there is no $c \in S$ that attacks b in AF_k. Hence, $AR_j \not\subset AR_S$.

Moreover, since S is the set of all in/direct defenders (respectively in/direct attackers) of a in AF_S, S must be the only preferred extension in AF_S that accepts (respectively rejects a.

Next, we show that AF_S contains no controversial argument. From one hand, any controversial argument in AF_S must be a member of S. On the other hand, no member of S can be controversial for a, otherwise, S will not be a conflict-free set. Hence, AF_S contains no controversial arguments.

The proof of (2.6.3.2). Let AF_S be such minimal argumentation framework for which the condition in (2.6.3.2) holds. If S is not a backing for a then there is some backing W of a where $W \subset S$. For such set W, by (2.6.3.1), we can construct $AF_W \sqsubset AF_S$ that satisfies the premise of (2.6.3.2). This, however, is in contradiction with the initial assumption that AF_S is such minimal sub-argumentation framework. Hence, there is no backing W of a where $W \subset S$.

The proof of (2.6.3.3). Let $\mathcal{AF}_S$ be the set of such minimal sub-argumentation frameworks for S. It is then easy to see that the normal sub-argumentation framework AF_S is the least upper bound normal sub-argumentation framework in AF for $\mathcal{AF}_S$. $\qquad\square$

The proof of observation (2.7.2).

The proof directly follows from the definition of a rational argumentation framework. □

The proof of observation (2.7.4).

The proof directly follows from the definition of strongly stable argumentation frameworks, and the fact that the sub-argumentation framework relation is a partial order relations. □

The proof of theorem (2.7.5).

The proof of (2.7.5) is straight forward. By theorem (1.3.16) every limited-controversial framework is coherent. On the other hand, if AF is not limited-controversial then we can isolate a sub-argumentation framework of AF that consists of only attack cycles of odd-length. Hence, limited-controversial argumentation frameworks identify the class of strongly stable argumentation frameworks. □

The proof of observation (2.7.7).

The proof directly follows from the definitions of rational, strongly and normally stable frameworks, and, the fact that the sub-argumentation framework relation and the normal sub-argumentation framework relation, both, are partial order relations. □

The proof of theorem (2.7.8).

The proof of (2.7.8), from left to right, is straight forward. Let us assume that the framework in question is normally stable, but, there is some minimal attack cycle, L, of odd length, that does not contains an attack cycle of even length. Let, $AF_L = \langle\, L\,,\, ATT_L\,\rangle \sqsubseteq^N AF$. Since, AF_L contains only an attack cycle of odd length, AF_L cannot be a stable framework. This, however, contradicts the assumption that AF is a normally stable framework.

The proof of (2.7.8), from right to left, is by contradiction. Let AF be some argumentation framework such that, while AF satisfies the condition in (2.7.8), AF has no stable extensions.

Let A be the set of all arguments in AF that are neither accepted nor attacked by some admissible set, and, the relation $\preceq$ be some order on U such that for any $a, b \in U$, $a \preceq b$ if and only if a is an in/direct defender or an in/direct attacker of b. Since, U is countable, it must have some minimal element under $\preceq$.

Let V be the set of minimal elements in U, and, $\mathcal{M}$ the set of all admissible sets in AF. By assumption, for any $a \in V$, a is neither accepted or attacked by any $S \in \mathcal{M}$. Next, for any $a \in V$, there must be some $b \in V$ such that a, b belong to some attack cycle of odd length. (note: a, b belong to an attack cycle if there are attack sequences, one from a to b, and, the other from b to a.) Otherwise, either $\overline{a} \subseteq AR - U$ or there is an even length cycle L such that $\overline{L} \subseteq (AR - U) \cup L$. In either case, a is either accepted or attacked by some $S \in \mathcal{M}$, contradicting the assumption that $a \in V$. Hence, all $a \in V$ belong to some attack cycle of odd length.

Let $AF_V = \langle AR_V, ATT_V \rangle \sqsubseteq^N AF$ where $AR_V = V$. We show that AF_V has a stable extension by appealing to the induction principle.

Let $a_1, a_2, \cdots, a_n$ be some arbitrary enumeration of $a \in V$, and, $AF_i \sqsubseteq^N AF_V$ for $1 \leq i \leq n$ where $AF_i = \langle AR_i, ATT_i \rangle$, $AR_1 = \{a_1\}$, $AR_i = AR_{i-1} \cup \{a_i\}$. We want to show that if AF_{i-1} is such that for any AF'_{i-1}, $AF'_{i-1} \sqsubseteq^N AF_{i-1}$, AF'_{i-1} has some stable extension, then, any AF'_i, where $AF'_i \sqsubseteq^N AF_i$, has some stable extension as well.

Let $\Delta = \{E^+, E^-, a_i\}$ where E^+ denotes some stable extension in AF_{i-1}, and, $E^- = AR_{i-1} - E^+$. There are eight possible scenarios, δ, of attack relation between members of Δ in AF_i, where $\delta \subseteq \Delta \times \Delta$. The scenarios δ that do not entail $(E^+, E^-), (E^-, E^+)$ are not possible. Of the remaining scenarios,

$$\text{if } a_i \not\hookrightarrow E^+ \text{ or } E^+ \hookrightarrow a_i, \text{ then, } E^+ \text{ is a stable extension of } AF_i, \tag{13}$$

$$\text{if } a_i \hookrightarrow E^+ \text{ and } AR_{i-1} \not\hookrightarrow a_i, \text{ then, } AF_i \text{ has some stable extension.} \tag{14}$$

The result (13) is self evident. To show (14), let $W = AR_{i-1} - \{a \mid a_i \not\hookrightarrow a\}$, and, $AF_W \sqsubseteq^N$ AF_{i-1}, where $AF_W = \langle AR_W, ATT_W \rangle$, $AR_W = W$. By assumption, AF_W has some stable extension E_W^+, by which, $E_W^+ \cup \{a_i\}$ is a stable extension in AF_i. Only two possible scenarios remain.

$$E^+ \not\hookrightarrow a_i, \ E^- \hookrightarrow a_i, \ a_i \hookrightarrow E^+, \ a_i \not\hookrightarrow E^-. \tag{15}$$

$$E^+ \not\hookrightarrow a_i, \ E^- \hookrightarrow a_i, \ a_i \hookrightarrow E^+, \ a_i \hookrightarrow E^-. \tag{16}$$

The scenario (16) is similar to the scenario (15). Hence, we proceed with the scenario (15), showing that, AF_i in (15) has some stable extension.

Let c, b denote the elements $c \in E^+, b \in E^-$ for which $b \hookrightarrow a_i$, $a_i \hookrightarrow c$. If b belongs to some stable extension of AF_{i-1} or $c \hookrightarrow a_i$, then, by (13), AF_i has some stable extension, and, our intended objective is reached. If that is not the case, then,

$$\text{neither any } b \text{ belongs to some stable extension of } AF_{i-1}, \text{ nor, any } c \hookrightarrow a_i. \tag{17}$$

Next, let $\mathcal{A}$ denote the set of all admissible sets in AF_{i-1}, and, T be the maximum element of $\mathcal{A}$ for which $c \notin T$. We observe that the admissibility of T is independent of c. Hence, by lemma (.0.2),

$$T \text{ is admissible in } AF_i. \tag{18}$$

Next, let $R = \{a \mid a \in AR_i, a \notin T, T \not\hookrightarrow a\} \cup \{a_i\}$, and, $\mathcal{C}$ denote the class of attack cycles, C, of odd length in R, characterized by some attack sequence, π, $\pi = (c, x_1, \cdots, x_m, a_i, c)$, $x_j \in R$, $1 \leq j \leq m$, for which π has no repeating member except c. Next, let $AF_R = \langle AR_R, ATT_R \rangle$, $AR_R = R$, and, $AF_W = \langle AR_W, ATT_W \rangle$, $AR_W = \bigcup_{C \in \mathcal{C}} C$ be the normal sub-argumentation frameworks of AF_i. Either, $AF_{\mathcal{C}}$ has some stable extension, or, not.

Let E_W be the stable extension of AF_W. From E_W, we then construct W^* and $AF_{W^*} =$

194

$\langle AR_{W^*}, ATT_{W^*} \rangle$, where, $AF_{W^*} \sqsubseteq^N AF_R$, $W^* = R - (W \cup W^+)$, $W^+ = \{a \mid a \in R, E_W \hookrightarrow a\}$. By assumption, AF_{W^*} has some stable extension, E_{W^*}, where, by construction, neither $E_W \hookrightarrow E_{W^*}$, nor, $E_{W^*} \hookrightarrow E_W$. Hence, we can say that $E_R = E_W \cup E_{W^*}$ is a stable extension in AF_R. Moreover, since $T \not\hookrightarrow E_R$, we can also say that $T \cup E_R$ is a stable extension in AF_i. Hence, we can say that,

$$\text{if } AF_W \text{ has a stable extension then } AF_i \text{ has a stable extension.} \tag{19}$$

Hence, to proceed with the proof by contradiction, we assume that AF_W has no stable extension. Next, to simplify the matters, for the moment,

$$\text{we assume that there are only one such } b, c, \text{ where, } c \in E^+, b \in E^-. \tag{20}$$

By assumption, for every $C \in \mathcal{C}$, there is some attack cycle of even length L_C such that $L_C \subset C$ and $b, a_i, c \in L$. Any even length attack cycle L_C can be partitioned into two (maximal) conflict free sets L_C^1, L_C^2. Without loss of generality, let $a_i, x_m \in L_1$. We observe that, not only, L_C^1 defends itself against all its attackers in L_C, but also,

$$L_C^1 \text{ defends itself against all its attackers in } C, \text{ and, attacks any } a \in (C - L_C^1). \tag{21}$$

Next, for all members of $\mathcal{C}$, no two L_C^1 attack another. Because, if there are some $C, C' \in \mathcal{C}$ such that $L_C^1 \hookrightarrow L_{C'}^1$, then, there is some $C'' \in \mathcal{C}$ where $C'' \subseteq C \cup C'$, $C'' \neq C$, $C'' \neq C'$. Hence, there is some $L_{C''}$ for which three is some $C''' \in \mathcal{C}$, such that, $C''' \subset C$ and $L_{C'}^1 \hookrightarrow L_{C'''}^1$, or, $C''' \subset C'$ and $L_{C'''}^1 \hookrightarrow L_C^1$. In either case, the process shall continue, and since C''' is a monotonically decreasing set, each time, C'', C''' result in the existence of a new smaller attack cycle of odd length. The process halts when either $a_i \hookrightarrow b$ or $c \hookrightarrow a_i$. However, this contradicts the assumption in (15).

Let L_W^1 be the set union of all such L_C^1. L_W^1, by (21), defends itself against all its attackers in

W and attacks any argument in $W - L_W^1$. Hence, we can say that,

$$L_W^1 \text{ is a stable extension of } AF_W. \tag{22}$$

This however contradicts the assumption that AF_W has no stable extensions.

Next, we remove the restriction set by the assumption in (20), ad, allow for multiple b, c, where, $c \in E^+, b \in E^-$. This however has no impact on the arguments made so fas, and therefore, has no impact on the result (22). One way to visualize this is by extending AF_W by arguments c', c'', b', b'' where $b' \hookrightarrow b''$, $c' \hookrightarrow c''$ and for all such b, c, $b \hookrightarrow b'$, $c \hookrightarrow c'$. Our new b, c now are b'', c'' which are the only b, c with respect to the new extended AF_W.

Finally, we show that any $AF_i' \sqsubseteq^N AF_i$ has some stable extension. This result directly follows from the fact that the enumeration of AR_V is done arbitrarily, and, our results are independent of any particular enumeration of AR_V. $\qquad \square$

The proof of theorem (2.7.9).

Let $AF = \langle AR, ATT \rangle$ be a normally stable argumentation framework and E a preferred extension in AF. We need to show that E is a stable extension in AF. Let $F = \{a \mid a \in AR, E \hookrightarrow a\}$, $W = E \cup F$, and, $V = AR - W$. If E is not a stable extension in AF, then, $V \neq \varnothing$. Next, let AF_E, AF_V be the normal sub-argumentation frameworks of AF, constructed from E, V, i.e., $AF_E \sqsubseteq^N AF$, $AF_V \sqsubseteq^N AF$, $AF_E = \langle AR_E, ATT_E \rangle$, $AF_E = \langle AR_E, ATT_E \rangle$, $AR_E = E$, $AR_V = V$. By assumption, AF_V has some stable extension E_V where $E_V \neq \varnothing$. By assumption, $E \not\hookrightarrow E_V$, hence, $E \cup E_V$ should be admissible in AF. However, this contradicts the assumption that E is a preferred extension in AF, because, $E \subset E_V$. $\qquad \square$

The proof of observation (2.7.11).

The proof directly follows from the definition of a compact argumentation framework. $\qquad \square$

Lemma .0.2. *Let S be an admissible set in some $AF = \langle AR, ATT \rangle$. Then, S remains admissible in AF^*, $AF^* \sqsupseteq AF$, $AF^* = \langle AR^*, ATT^* \rangle$ if for all $(a, b) \in ATT^* - ATT$, either $b \notin S$ or $S \cap \overline{a} \neq \varnothing$ in AF^*.*

The proof of lemma .0.2.

The proof is self evident. We observe that for all $a \in AR^*$, if $a \hookrightarrow S$ then $S \hookrightarrow a$. Hence, S remains admissible in AF^*. $\square$

Proofs for chapter 3

The proof of observation (3.2.6).

The reason is obvious. The minimality condition of backings, require that T cannot dispense with any of its arguments including b. Hence, it must contain some minimal subset that accepts c. This minimal subset is in turn a backing of b. □

The proof of observation (3.2.7).

We observe that if a is not an active argument for any $c \in \bar{b}$ then $a \notin S$, $a \notin T$, for any $c \in \bar{b}$, any $S \in \langle c \rangle^+$ and any $T \in \langle c \rangle^-$. Hence, by definition of operation $\circ$, $+$, $a \notin S$, $a \notin T$, for any $S \in \langle b \rangle^+$ and any $T \in \langle b \rangle^-$. □

The proof of lemma (3.2.9).

We observe that if the condition does not hold, then, either for all odd numbers i, or, for all even numbers i, $\langle a_i \rangle^+ \neq \varnothing$. However, this contradicts the original assumption that the attack sequence is not active. □

The proof of lemma (3.2.10).

If a is an active argument for b, then a belongs to some backing of b. The, by definition, there is some active attack sequence from b to a. □

The proof of observation (3.3.2).

The proof directly follows from definition of intercepts, definition (3.3.1). □

The proof of lemma (3.3.3).

The proof from right to left. By assumption, if $\langle\, S\, \rangle^- = \varnothing$ then $\langle\, b\, \rangle^+ = \varnothing$. Hence, there is some non trivial attack sequence from b to a, namely $\pi_0 = (b, a)$ while all the attack sequences π from b to a are intercepted. Hence, we can say that a is intercepted for b.

The proof from left to right. If $\langle\, a\, \rangle^+ \neq \varnothing$ then π_0 is not intercepted, and, so a is not intercepted for b, contradicting the original assumption. Hence, let $\langle\, a\, \rangle^+ = \varnothing$. Next if there is no such $S \subseteq \bar{b} - \{a\}$ where $\langle\, S\, \rangle^- = \varnothing$ then there is some admissible set that attacks all the attackers of b. Hence, $\langle\, b\, \rangle^+ \neq \varnothing$ which contradicts the original assumption that $\langle\, b\, \rangle^+ = \varnothing$. $\qquad\square$

The proof of lemma (3.3.4).

The proof is trivial. We observe that for either case (3.3.4.1) or (3.3.4.2), there is no active attack sequence from b to a. Hence, a is intercepted for b. $\qquad\square$

The proof of theorem (3.3.5).

The proof from left to right. Let $a \in \bar{b}$ then by lemma 3.3.3 there must be some $S \subseteq \bar{b} - \{a\}$ such that $\langle\, S\, \rangle^- = \varnothing$. Next, let $a \notin \bar{b}$. Now, if there is some $d \in D$ such that neither d is intercepted for b nor a is intercepted for d, then, there is some attack sequence from b to a which is active. This however contradicts the original assumption that a is intercepted for b.

The proof from right to left is directly followed from lemmas 3.3.3 and 3.3.4. $\qquad\square$

The proof of observation (3.4.2).

The proof is straight forward. Every admissible set that accepts a, must attack all $c \in \bar{a}$. Hence, by the definition of critical argument, defender or attacker, every admissible set that accepts, or respectively attacks, b, must also attack c. $\qquad\square$

The proof of observation (3.4.3).

The proof is trivial following from the definition of backings of an argument that for every

admissible set T that accepts a there is some admissible set $S \in \langle a \rangle^+$ that $S \subseteq T$. Similarly, for every admissible set T that attacks a there is some admissible set $S \in \langle a \rangle^-$ that $S \subseteq T$. $\square$

The proof of observation (3.4.4).

The proof is self evident. It follows directly from the definition of $\prod$ and critical arguments. $\square$

The proof of theorem (3.4.6).

We show the proof of 3.4.6 by means of contradiction. A critical defender (resp. attacker) belong to all positive (resp. negative) backings of an argument. Hence, there is some minimal set B of negatively critical arguments for the acceptance of a such that for any $c \in C$ there is some $T \in \langle B \rangle^+$ where $c \in T$. Let $\mathcal{B}$ be the set of such minimal sets B.

Now, if the claim of 3.4.6 then for all $B \in \mathcal{B}$ there is some $W \in \langle B \rangle^-$ where $W \cap C = \varnothing$. However, if this is the case then there is some $S \in \langle a \rangle^+$ such that $W \subseteq S, S \cap C = \varnothing$, contradicting the initial assumption that for all $S \in \langle a \rangle^+, S \cap C \neq \varnothing$. $\square$

The proof of theorem (3.4.8).

The proof of (3.4.8.1), (3.4.8.2) is trivial and directly follows for the definitions of critical attacker and defender of an argument, and, the partial order over subset relations. For the proof of (3.4.8.3), (3.4.8.4), if $d \in \overline{b}$ then all positive backing of b must attack d, and therefore, include a. Hence, a must respectively belong to all admissible sets that accept or respectively attack c. $\square$

The proof of (3.5.2).

The proof is straightforward, resulting from the fact that if the union of two admissible sets is not a conflict free set, then, the two set must symmetrically attack each other. $\square$

The proof of theorem (3.5.3).

The proof is trivial. It directly follows from the definitions of the incompatible arguments and

the backings of an argument. In addition, for any $\mathcal{A} \in 2^{2^{AR}}$, if S is not conflict free, then,

$$\{S\} \mathbin{\dot{\circ}} \mathcal{A} = \varnothing.$$ □

The proof of observation (3.5.5).

The proof is straightforward. We only show the proof for first part of the observation. The proof for the second part is identical to the proof for the first part, with a minor adjustment that we need to replace all occurrences of $\langle\, x \,\rangle^+$ with $\langle\, x \,\rangle^-$, and, vice versa.

If $\langle\, a \,\rangle^+ \neq \varnothing, \langle\, b \,\rangle^+ \neq \varnothing$ and $\langle\, a \,\rangle^+ \mathbin{\dot{\circ}} \langle\, b \,\rangle^+ = \varnothing$, then, by theorem (3.5.3), a, b are positively incompatible. Moreover, for any $S \in \langle\, a \,\rangle^+$, and, any $T \in \langle\, b \,\rangle^+$, the sets $S \cup \{a\}$ and $T \cup \{b\}$ are admissible. Then, by observation (3.5.2), either a or S is attacked by some admissible set. Hence, for both a, b, $\langle\, a \,\rangle^- \neq \varnothing, \langle\, b \,\rangle^- \neq \varnothing$. □

The proof of observation (3.5.6).

The proof is trivial. We observe that if the conditions of definition (3.5.1) hold for the sets M, N of definition (3.5.1), those conditions will hold for any superset of M, N as well. □

The proof of Observation (3.5.7).

Obviously, if for some $b \in \overline{a}$, $\langle\, b \,\rangle^- = \varnothing$, then, $\langle\, a \,\rangle^+ = \dot{\prod}_{b \in \overline{a}}\langle\, b \,\rangle^- = \varnothing$. Similarly, if there is some subset of $\overline{a}$ that is negatively incompatible, then, by theorem (3.5.3), again, $\langle\, a \,\rangle^+ = \dot{\prod}_{b \in \overline{a}}\langle\, b \,\rangle^- = \varnothing$. □

The proof of lemma (3.5.8).

The proof of (3.5.8.1). If two admissible sets T_1, T_2 attack each other, then there are some admissible subsets of each admissible set, $S_1 \subseteq T_1, S_2 \subseteq T_2$ that attack ech other, i.e., $S_1 \hookrightarrow S_2$ and $S_2 \hookrightarrow S_1$. Hence, for every $a_1 \in S_1, a_2 \in S_2$, there is some attack sequence π_1 from a_1 to a_2 and some attack sequence π_2 from a_2 to a_1. By the virtue of the assumption that every

202

argument on both π_1, π_2 is admisible, every argument in either attack sequence has a positive backings. As a result, both attack sequences are active.

The proof of (3.5.8.2). If two arguments a, b are incompatible then for any pair of their backings, $S_1 \in \langle\, a\, \rangle^+$ and $S_2 \in \langle\, b\, \rangle^+$, we have $S_1 \hookrightarrow S_2$, $S_2 \hookrightarrow S_1$. Following the first part of this lemma, we can the say, there are active arguments $c \in S_1$ for a and $d \in S_2$ for b such that the path between c and d is not intercepted. $\square$

The proof of observation (3.5.9).

The proof for the first part of observation. Let $S = \{a, c\}$. By definition, if b is a critical defender of c, then, by assumption, $\langle\, c\, \rangle^+ \neq \varnothing$, and, $b \in R$, for every $R \in \langle\, c\, \rangle^+$. Therefore, $a, b \in T$, for every $T \in \{S\}\, \dot{\circ}\, \langle\, c\, \rangle^+$, and, since $a \hookrightarrow b$, no T will be conflict free. The proof for the second part of observation is identical to the proof for the first part, except we need to replace $\langle\, c\, \rangle^+$ with $\langle\, c\, \rangle^-$. $\square$

The proof of lemma (3.5.10).

The proof is similar to the proof of lemma (3.5.9). Let $S = \{a, c\}$. For the first part of lemma, by definition, if b is a critical defender of c, then, by assumption, $\langle\, c\, \rangle^+ \neq \varnothing$, and, $b \in R$, for every $R \in \langle\, c\, \rangle^+$. Therefore, $a, b \in T$, for every $T \in \{S\}\, \dot{\circ}\, \langle\, c\, \rangle^+$. Next, since a and b are positively incompatible, no admissible set T can accept both a and b, and, therefore, $\{S\}\, \dot{\circ}\, \langle\, c\, \rangle^+ = \varnothing$. The proof for the second part of lemma is identical to the proof for the first part, except we need to replace $\langle\, c\, \rangle^+$ with $\langle\, c\, \rangle^-$. $\square$

The proof of theorem (3.5.11).

Since, b_2 is positively incompatible with b_1, a critical defender of a_1, then, by lemma (3.5.10), a_1, b_2 are positively incompatible arguments. Next, since a_1, b_2 are positively incompatible, and, b_2 is a critical defender of a_2, by lemma (3.5.10), a_1, a_2 are positively incompatible. The proof

of the second part of theorem is identical to the proof for the first half, with the difference that we need to replace all the instances *positively incompatible* with *negatively incompatible*. ☐

The proof of lemma (3.5.12).

The proof is straightforward. If an argument is positively incompatible with an argument, it then cannot be in any admissible set that accepts the argument which also includes the positive backings of the argument. ☐

The proof of observation (3.6.2).

The proof is simple. By assumption, c is a critical for b. Thus, for any admissible set T, if $b \in T$ then there is some admissible set $W \subset T$, $W \hookrightarrow a$ where $W = R \cup \{c\}$, $R \in \langle c \rangle^+$. Hence, T cannot be a negative backing for a. ☐

The proof of lemma (3.6.3).

The proof is similar to the proof for (3.6.2). The proof from right to left. By assumption, for every $T \in \langle b \rangle^+$, $S \cap T \neq \varnothing$, so, for every T there is some $d \in \bar{a}$, and, some $R \in \langle d \rangle^+$ such that $R \cup \{d\} \subseteq T$. Hence, T cannot be a negative backing for a.

The proof from left to right. Suppose, b is negatively redundant for a, and, there is no such set $S \subset \bar{a}$ such that for every $T \in \langle b \rangle^+$, $S \cap T \neq \varnothing$. If so, then, there is some $R \in \langle b \rangle^+$ where, for every $c \in \bar{a} - \{b\}$ and for every $W \in \langle c \rangle^+$, $R \not\subset W$. Hence, $R \cup \{b\}$ is a negative backing for a which contradicts the original premise. ☐

The proof of observation (3.6.4).

The proof is by contradiction. By assumption, c is a critical for a. Thus, for any $S \in \langle a \rangle^+$, if $b \in S$ then there is some admissible set $T = S - \{b\}$ that accepts a. Because, since $c \in S$, then T attacks all the arguments that S attacks, and so, accepts all the arguments that S accepts. Hence, S cannot be a positive backing for a. ☐

The proof of lemma (3.6.5).

Let, P denote the set of all arguments c where for any $\pi \in \Pi$, there is some c on π such that $D \hookrightarrow c$, or, $d \hookrightarrow c$ for every $d \in C$.

The proof from right to left is by contradiction. For any $T \in \langle\, Y\, \rangle^+$, let $b \in T$ and $Q = P \cap T$, then, T can be partitioned into T_b, T_Q where T_b accepts Q and $b \in T_b$. By assumption in the antecedent, for any such Q, for any arbitrary $d \in C$, there is some $V \subseteq D \cup \{d\}$ such that V accepts Q.

Next, since Q is on every $\pi \in \Pi$, AF^*, there is a subset T_Q of T such that if an admissible set R accepts Q then $R \cup T_Q$ is admissible. Hence, there is some $R \in \langle\, V\, \rangle^+$ such that $Z \cup T_Q \in \langle\, Y\, \rangle^-$ where $Z = R \cup V$.

Thus, we can conclude, for any $T \in \langle\, Y\, \rangle^-$, where $T = T_b \cup T_Q$, $b \in T_b$, there is some Z, independent of the choice of $d \in C$, such that $Z \cup T_Q \in \langle\, Y\, \rangle^-$.

Now, if there is some $S_b \in \langle\, a\, \rangle^+$ where $b \in S_b$, S_b must be in the form of $S_b = T_b \cup T_Q \cup T_W$ where $T_W \in \langle\, W\, \rangle^-$. However, by the results so far, for any such S_b, there is some $S_Z \in \langle\, a\, \rangle^+$ where $S_Z = Z \cup T_Q \cup T_W$.

Next, since $d \in C$ is chosen arbitrarily, and, by assumption C is a positive critical set and D is the set of critical defenders of b, we then have $Z \subseteq T_W$ for which $S_Z = T_Q \cup T_W$.

Thus, we can say, for any S_b there is some S_Z such that $S_Z \subset S_b$. This, however, leads to a contradiction.

The proof from left to right is also by contradiction. Let us suppose that b is a positively redundant argument for a while the conditions in the consequent of lemma are not fully met. That is, there is some positively active attack sequence π for which no argument c on π is either attacked by D, or, is attacked by all $d \in C$. Let $y \in Y$, be the corresponding y for this attack sequence π.

If we follow the construction above, for this $y \in Y$, there is some $T \in \langle\, y\,\rangle^-$, for which we can make sure that there is some S_Z such that $S_Z \subset S_T$. Hence, S_T accepts a while no subset of S_T accepts a. Hence, S_T must be a positive backing for a. However, since $b \in S_T$, w then have a contradiction where by assumption b is positively redundant for a. $\qquad\square$

The proof of theorem (3.6.7).

The proof for (3.6.7.1a). If for all $c \in C$, b is an active defender for c and c is positively redundant for a, then, there is some active attack sequence from a to b. Next, since, for all $c \in C$ and all $T \in \langle\, a\,\rangle^+$, $c \notin T$, and, there is no other active attack sequence from a to b that does not pass through some $c \in C$, we can conclude that $b \notin T$ for all $T \in \langle\, a\,\rangle^+$. Otherwise, we will have some $c \in C$, for which $c \in T$, for some $T \in \langle\, a\,\rangle^+$ which contradicts the original premise. The proof for (3.6.7.1b) is the same as the proof for (3.6.7.1a) with the difference that every instance of *positively redundant* should be changed with *negatively redundant*.

The proof for (3.6.7.2a). If for all $c \in C$, c is an active defender for a and b is positively redundant for c, then, there is some active attack sequence from a to b. Next, since, for all $c \in C$ and all $T \in \langle\, c\,\rangle^+$, $b \notin T$, and, there is no other active attack sequence from a to b that does not pass through some $c \in C$, we can conclude that $b \notin T$ for all $T \in \langle\, a\,\rangle^+$. Otherwise, we will have some $c \in C$, for which $b \in T$, for some $T \in \langle\, c\,\rangle^+$, and, since there is some $R \in \langle\, a\,\rangle^+$ where $c \in R$, then, there is some $W \in \langle\, a\,\rangle^+$ that $b \in W$. This, however, contradicts the original premise. The proof for (3.6.7.2b) is almost identical to the proof for The proof for (3.6.7.2a). $\qquad\square$

The proof of theorem (3.6.8).

The proof of (3.6.8.1).

The proof of (3.6.8.2). If $Y \neq$ then by lemma (3.6.5) b is positively redundant for a. Otherwise, let $Y = \varnothing$.

The proof from right to left. We observe that $W \neq \varnothing$. Because, otherwise, there is no positive attack sequence from a to be, and, so b cannot be positively redundant for a. Next, if there is some $w \in W$ for which b is not negatively redundant, then, $Y \neq \varnothing$ which contradicts the original premise.

The proof from left to right. By definition if b is negatively redundant for all $w \in W$, then, for all $w \in W$, there is no $T \in \langle w \rangle^-$ where $b \in T$. Hence, $b \notin S$ for all $S \in \langle a \rangle^+$ which means b is positively redundant for a. $\square$

The proof of lemma (3.6.9).

The proof of (3.6.9.1). Let π be a positively active attack sequence. Then, by definition b is neither intercepted nor is positively incompatible with a. If b is not an active defender for a, then, b is by definition positively redundant for a. Conversely, if b is an active defender or positively redundant for a, then, by definition, π is a positively active attack sequence.

The proof of (3.6.9.2) and (3.6.9.3) are similar to the proof of (3.6.9.1). $\square$

The proof of lemma (3.7.2).

The proof from right to left is self evident. Hence, we only present the proof from left to right. If $\mathcal{A}_3 \neq \mathcal{A}_1 \circ \mathcal{A}_2$ then there is some $S_3 \in \mathcal{A}_3$ such that for all $S_1 \in \mathcal{A}_1, S_2 \in \mathcal{A}_2$, $S_3 - (S_1 \cup S_2) \neq \varnothing$. However, this contradicts the original assumption about $\mathcal{A}_{31}, \mathcal{A}_{32}$. $\square$

The proof of lemma (3.7.5).

The proof of (3.7.5.1). Let AF_1, AF_2 and etc., be generally referred by $AF_i = \langle AR_i, ATT_i \rangle$.

If no $a_1 \in AR_1$ is an active argument for any $a_2 \in AR_2$ in AF, then, for any arguments $a_2 \in AR_2$, all the backings S of a_2 are $S \subseteq AR_{23}$. Consequently, for all arguments in AR_2, they have the same backings in AF_{23} as they have in AF. Hence, w can say that $\mathcal{A}_2^* = \mathcal{A}_{23}^*$. The same can be said with respect to the arguments in AF_1, for which we then have $\mathcal{A}_1^* = \mathcal{A}_{13}^*$.

The proof of (3.7.5.2). The proof is by contradiction. If the claim of does not hold then, without loss of generality, there is some $a_1 \in AR_1$ that is an active argument for some $a_2 \in AR_2$. If that is the case then there are some AF_1', AF_2' in which a_1 is a critical argument argument for a_2 such that the admissibility of a_2 changes depending on whether or not a_1 is in $AF' = AF_1' +^N AF_2' +^N AF_3$. Under this condition, it is then easy to see that the original premise that $\mathcal{A}_2^{*'} = \mathcal{A}_{23}^{*'}$ is violated where $\mathcal{A}_2^{*'} = \{S \mid S = T \cap AR_2', T \in \mathcal{A}'\}$, $\mathcal{A}_{23}^{*'} = \{S \mid S = T \cap AR_2', T \in \mathcal{A}'_{23}\}$. Hence, If no $a_1 \in AR_1$ can be an active argument for any $a_2 \in AR_2$ in AF, and vice versa. $\square$

The proof of observation (3.7.7).

The proof directly follows from observation (3.3.2) and the definition (3.7.6). $\square$

The proof of lemma (3.7.8).

The proof follows naturally from lemma (3.7.5) and the definition (3.7.6) of "disjoint by intercept". $\square$

The proof of lemma (3.7.11).

The proof is trivial. If any two sub-frameworks are intersecting, then, they cannot be disjointed by intercept. Hence, they must be non-intersecting. $\square$

The proof of theorem(3.7.12).

If two sub-framework are disjointed by intercept then no argument in one sub-framework is an active argument for any argument another sub-framework. Hence, all the backings of all the arguments in a sub-framework is decided by the arguments in that framework, and, so the admissible sets in all the sub-frameworks stand on their own. Therefore, $\mathcal{A} = \prod_{AF' \in \mathcal{AF}} \mathcal{A}'$. $\square$

The proof of theorem(3.7.14).

The proof of this theorem directly follows from the proof of theorem (3.7.12). We observe that

in all sub-frameworks AF°, all sub-frameworks in $\widehat{\mathcal{AF}} \cup \{AF^{*'}\}$ are disjointed by intercept.

By theorem (3.7.12), we then have, $\quad \mathcal{A}^\circ = \prod_{AF' \in \widehat{\mathcal{AF}}} \mathcal{A}' \dot{\circ} \mathcal{A}^{*'} = \widehat{\mathcal{A}} \dot{\circ} \mathcal{A}^{*'}.$ $\hfill \square$

The proof of theorem(3.7.15).

The proof is by the induction principle. Let $\mathcal{AF}' = \{AF_1, AF_2, \cdots\}$ be some enumeration of $\mathcal{AF}'$, for which the corresponding sub-frameworks AF_i^* are constructed such that $AF_1^* = AF_1 +^N \widehat{AF}$, and, $AF_{i+1}^* = AF_i^* +^N AF_{i+1}$.

We need to show that for each AF_{i+1}^*, the set of admissible sets $\mathcal{A}_{i+1}^*, \mathcal{A}_i^*, \mathcal{A}_i$ in their respective sub-frameworks AF_{i+1}^*, AF_i^*, AF_i form $\mathcal{A}_{i+1}^* \subseteq \mathcal{A}_i^* \dot{\circ} \mathcal{A}_{i+1}$.

Following the first original premises that any two distinct AF', AF'' in $\mathcal{AF}$ are disjointed by intercept, and, the definitions of *disjointed by intercept frameworks* and *intercepted attack sequences*, we can conclude that, all $S \in \widehat{\mathcal{A}}$ remain admissible in any AF_i^*. The reason is, all the attacks from arguments in AF_i^* to arguments in $\widehat{AF}$ are to the arguments a for which $\langle a \rangle^+ = \varnothing$, $\langle a \rangle^- \neq \varnothing$. Hence, the admissibility of no $\S \in \widehat{\mathcal{A}}$ is changed. $\hfill$ *(res 1)*

Next, following the second original premise that any two distinct AF', AF'' in $\mathcal{AF}^*$ are disjointed by intercept by some set of arguments in $\widehat{AF}$, and, the definitions of *disjointed by intercept frameworks* and *intercepted attack sequences*, we can see that, all the attacks from arguments in AF_i^* to arguments in AF_{i+1} are by the arguments a in $\widehat{AF}$ for which $\langle a \rangle^+ = \varnothing$, $\langle a \rangle^- \neq \varnothing$. Hence, an argument b in AF_{i+1} that is defended against all its attackers by some admissible set in AF_{i+1}, is still defended against all its attackers in AF_{i+1}^*. $\hfill$ *(res 2)*

Following the results *(res 1), (res 2)* above, we can then conclude, for every AF_i^*, AF_{i+1}^*, the relation $\mathcal{A}_{i+1}^* \subseteq \mathcal{A}_i^* \dot{\circ} \mathcal{A}_{i+1}$ holds. $\hfill \square$

The proof of lemma (3.7.18).

The proof of (3.7.18.1). The proof is trivial, if $AF' \in \mathcal{AF}$ is not polarized then some attack

relation can be removed where the resulting $\mathcal{AF}'$ is still in $\mathbb{AF}$. This however contradicts the original assumption that $\mathcal{AF}$ is the minimal element of $\mathbb{AF}$.

The proof of (3.7.18.2) from right to left. Let $AF'' \sqsubset AF'$, and, $AF''' = AF' -^N AF''$. By assumption that $AF' \in \mathcal{AF}$ is polarized, AF''' cannot be disjointed by intercept for AF''. Hence, AF'' must only be missing some attack relations from AF'. If so, then $\mathcal{AF}'$ cannot be the element of $\mathbb{AF}'$. Because, since, $AF' \in \mathcal{AF}$ is polarized, the removal of any attack relation may affect the admissibility of argument in some sub-framework of AF'. Hence, $\mathcal{AF}$ must be the minimum element of $\mathbb{AF}'$.

The proof of (3.7.18) from left to right. If $\mathcal{AF}$ is the minimum element of $\mathbb{AF}'$ then no argument or attack relation can be removed from any $AF' \in \mathcal{AF}$. Hence, AF' □

The proof of theorem (3.7.19).

The proof is simple. By definition of $\mathcal{AF}$, all $AF' \in \mathcal{AF}$ are disjointed by intercept where their normal sum, $\sum^N$ is AF. Again, by the definition, all $AF' \in \mathcal{AF}$ contain no intercepted attack sequences, and thus, they will be biased sub-frameworks of AF.   □

Proofs for chapter 4

The proof of theorem (4.3.3).

Prooof of (4.3.3.1). $\mathbf{T}_d, \mathbf{U}_d, \mathbf{F}_d$ are mutually exclusive, therefore $\mathbf{H}_d(A)$ has one and only one value.

Proof of (4.3.3.2) from left to right. If $\exists A \in \mathbf{F}_d$ that A is a subset of $B \in \mathbf{T}_d$ then not all $A \in \mathbf{F}_d$ are conclusive defeaters of d which is contradictory to the assumption.

Proof of (4.3.3.2) from right to left. If no $A \in \mathbf{F}_d$ is a subset of $B \in \mathbf{T}_d$ then for $\forall A \in \mathbf{F}_d$, A is a conclusive defeater of d making d a default rule.

Proof of (4.3.3.3.a). The members in $\mathbf{F}_d$ create a partial order w.r.t set inclusion. Since every $A \in \mathbf{F}_d^k$ is a minimal set in $\mathbf{F}_d$ then no proper subset of A is in $\mathbf{F}_d$. Hence, if A is not a singleton then all its members should belong to some $B \in \mathbf{T}_d$. Otherwise, it contradicts the assumption.

Proof of (4.3.3.3.b). Accordingly, if $\mathbf{H}_d(X) = 0$ then $\exists C \in \mathcal{F}_d$ such that $C \subseteq X$ otherwise $\mathbf{H}_d(X) \neq 0$. So, there is $A \in \mathbf{F}_d^k$ ssuch that $A \subseteq C$. Conversely, if there is $A \in \mathbf{F}_d^k$ such that $A \subseteq X$ then by definition of the conclusive defeater $\mathbf{H}_d(X) = 0$. $\square$

The proof of observation (4.3.10).

The function $\min_{d \in D}(\mathbf{H}_d(X))$ has one and only one value for a given argument. Next, the definition of conclusive defeat implies that defeat is context independent so $\min_{d \in D}(\mathbf{H}_d(X)) = 0$ is context independent. $\square$

The proof of theorem (4.4.2).

Proof of (4.4.2.1). By definition of defeat relationship, $\exists \mathcal{A}_1 \subseteq \mathcal{A}$ such that given defeat conditions are satisfied (for outright or provisional defeat). Now, since $\mathcal{A}_1$ is also subset of $\mathcal{A}'$, the same defeat conditions are still satisfied in relation to $\mathcal{A}'$.

Proof of (4.4.2.2). By definition of conclusive defeat scenario, $\forall \mathcal{A}''$ and $\mathcal{A}_c \subseteq \mathcal{A}'' \subseteq \mathcal{A}$ then $\exists d \in D_{\AA}$ such that $\mathbf{H}_d(Cn(\mathbf{A}'')) = 0$. Moreover, since $\AA \sqsubset \AA'$ we have $d \in D_{\AA'}$. Hence, $\mathbf{H}_{\AA'}(\mathcal{A}'') = 0$ or $\mathcal{A}_c$ is a conclusive defeat scenario for $\AA'$ in $\mathcal{A}$.

Proof of (4.4.2.3). Let $\mathcal{A}_c$ be a defeater of $\AA$ in $\mathcal{A}$ then by definition (4.4.1), $Cn(\mathcal{A}_c)$ are defeaters of a rule $d \in D_{\AA}$ and by the second part of this theorem, $F \subseteq Cn(\mathcal{A}_c)$ where $F \in \mathbf{F}_d^k$. Therefore, $\mathcal{A}_c$ is a TConclusive defeat scenario (observation (4.3.10) and definition TConclusive defeat scenario). Next, if $\mathcal{A}_c$ is not the minimal set where $F \subseteq Cn(\mathcal{A}_c)$ then there is $\mathcal{A}'_c \subset \mathcal{A}_c$ such that $F \subseteq Cn(\mathcal{A}'_c)$. This means $\mathbf{H}_{\AA}(\mathcal{A}' \setminus \{\AA_j\}) = 0$ contradicting the defeat condition in definiton (4.4.1.1) ($\mathcal{A}'$ is the $\mathcal{A}'$ in definition (4.4.1.1) and $\AA_j \in (\mathcal{A}_c \setminus \mathcal{A}'_c)$. For the same reason (i.e. $\mathcal{A}_c$ being the maximal set in definition (4.4.1.1)) the context of defeat is $\mathcal{A}' \setminus \mathcal{A}_c = \varnothing$, and, if $F \in \mathbf{F}_d^k$ is singleton then $\mathcal{A}_c$ will be singleton.

Proof of (4.4.2.4). $Cn(\mathcal{A}_c)$ defeating a rule d implies $\mathbf{H}_{\AA}(\mathcal{A}'') = 0$ for any $\mathcal{A}''$ such that $\mathcal{A}_c \subseteq \mathcal{A}''$. $\square$

The proof of theorem (4.4.5).

If D is only comprised of indefeasible or default rules then either arguments have no defeaters or any argument set $\mathcal{A}_c$ that defeats an argument in $\mathcal{A}$ is its conclusive defeater (theorem (4.4.2.3)). Furthermore, any attack by a conclusive defeater cannot be reinstated by context ($\mathbf{H}_{\AA}(\mathcal{A}')$ where $\mathcal{A}_c \subseteq \mathcal{A}'$ is always zero). $\square$

The proof of theorem (4.5.3).

The Status function $\mathbf{E}$ is sum of three partial functions with exclusive domains. Moreover, from

definition of Translation, every $\{\text{Å}\}$ is mapped to one and only one argument $\alpha \in AR$ of AF. Hence, every argument has one and only one status. The status function $\mathbf{S}$ is a "max function", therefore, every literal also has one and only one status. $\square$

The proof of theorem (4.5.6).

Proof of (4.5.6.1). If $\mathbf{S}(a) = 1$ and $\mathbf{S}(\neg a) = 1$ then there are two justified arguments $\text{Å}_1, \text{Å}_2 \in \mathcal{A}$ where $a \in Cn(\mathcal{A}_1)$ and $\neg a \in Cn(\mathcal{A}_2)$. Thus, their mapped arguments α_1, α_2 in Dung's AF belong to all preferred extensions contradicting the initial assumption.

Proof of (4.5.6.2.a). If AT is not consistent then there are two arguments $\text{Å}_1, \text{Å}_2 \in \mathcal{A}$ such that $a \in Cn(\text{Å}_1)$ and $\neg a \in Cn(\text{Å}_2)$. Hence, there are two rules $d_1, d_2 \in D$ such that $hd(d_1) = a$ and $hd(d_2) = \neg a$. Now, if one of the d_1 or d_2 is a normal rule then the corresponding argument should have been conclusively defeated by the other argument and so not to be accepted. Therefore, d_1 and d_2 should both be indefeasible rules. But, if d_1 and d_2 are both indefeasible rules then the induced argumentation theory AT' from $(bd(d_1) \cup bd(d_2), \{d_1, d_2\})$ would be inconsistent which is contradictory to the initial assumption.

Proof of (4.5.6.2.b). According to theorem (4.4.5), $\mathbb{AT}$ is context insensitive, therefore, rules 1(c), 2(a)(ii), 2(a)(iii), 2(b) of translation are not applied. Hence, the mapping $\mathbf{M1}$ and $\mathbf{M2}$ are one-to-one and consequently bijective mappings. Now, if all members of $\mathbf{F}_d^k$ are singletons then all $\mathcal{A}_c$ in $\mathbb{AT}$ are singletons (theorem (4.4.2.3)). Thus, by rule 1(b) of translation, arguments in $AF = \langle AR, ATT \rangle$ are bijective mappings of singleton argument sets in $\mathbb{AT}$. Hence, $\mathbf{M1}$ acts like a bijective function from $\mathcal{A}$ to AR. Moreover, since all attack relations in ATT are bijective mappings from the domain $\{Z \mid Z \text{ is a singleton set}\} \times \{Z \mid Z \text{ is a singleton set}\}$ (all $\mathcal{A}_c$ are singletons), $\mathbf{M2}$ acts like a bijective function from R^* to ATT. Hence, $\langle \mathcal{A}, R^* \rangle$ and $\langle AR, ATT \rangle$ are two isomorphic structures. $\square$

Bibliography

[AC98]   Leila Amgoud and Claudette Cayrol. On the acceptability of arguments in preference-based argumentation. In *UAI '98: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, University of Wisconsin Business School, Madison, Wisconsin, USA, July 24-26, 1998*, pages 1–7, 1998.

[AC02]   Leila Amgoud and Claudette Cayrol. A reasoning model based on the production of acceptable arguments. *Ann. Math. Artif. Intell.*, 34(1-3):197–215, 2002.

[AP97]   Nicholas Asher and Francis Jeffry Pelletier. Generics and defaults. *Handbook of Logic and Language, John Van Bentham and Alice ter Meulen (eds.)*, 1997.

[AP02]   Leila Amgoud and Simon Parsons. An argumentation framework for merging conflicting knowledge bases. In *Logics in Artificial Intelligence, European Conference, JELIA 2002*, pages 27–37, 2002.

[Bau11]   Ringo Baumann. Splitting an argumentation framework. In *Logic Programming and Nonmonotonic Reasoning - 11th International Conference, LPNMR 2011, Vancouver, Canada, May 16-19, 2011. Proceedings*, pages 40–53, 2011.

[BDKT97]   Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artif. Intell.*, 93:63–101, 1997.

[Ben02]   Trevor J. M. Bench-Capon. Value based argumentation frameworks. In *9th In-*

ternational Workshop on Non-Monotonic Reasoning (NMR 2002), April 19-21, Toulouse, France, Proceedings, pages 443–454, 2002.

[BG04]   Pietro Baroni and Massimiliano Giacomin. A general recursive schema for argumentation semantics. In *Proceedings of the 16th Eureopean Conference on Artificial Intelligence, ECAI'2004, including Prestigious Applicants of Intelligent Systems, PAIS 2004, Valencia, Spain, August 22-27, 2004*, pages 783–787, 2004.

[BGG05]   Pietro Baroni, Massimiliano Giacomin, and Giovanni Guida. Scc-recursiveness: a general schema for argumentation semantics. *Artif. Intell.*, 168(1-2):162–210, 2005.

[BH01]   Philippe Besnard and Anthony Hunter. A logic-based theory of deductive arguments. *Artif. Intell.*, 128(1-2):203–235, 2001.

[Boc02]   Alexander Bochman. Collective argumentation. In *9th International Workshop on Non-Monotonic Reasoning (NMR 2002), April 19-21, Toulouse, France, Proceedings*, pages 436–443, 2002.

[Bod02]   Gustavo Adrian Bodanza. Disjunctions and specificity in suppositional defeasible argumentation. *Logic Journal of the IGPL*, 10(1):23–49, 2002.

[BPWA13]   Trevor J. M. Bench-Capon, Henry Prakken, Adam Wyner, and Katie Atkinson. Argument schemes for reasoning with legal cases using values. In *International Conference on Artificial Intelligence and Law, ICAIL '13, Rome, Italy, June 10-14, 2013*, pages 13–22, 2013.

[BTK93]   Andrei Bondarenko, Francesca Toni, and Robert A. Kowalski. An assumption-based framework for non-monotonic reasoning. In *Logic Programming and Non-*

monotonic Reasoning LPNMR, Proceedings of the Second International Workshop, Lisbon, Portugal, June 1993, pages 171–189, 1993.

[CA05]   Martin Caminada and Leila Amgoud. An axiomatic account of formal argumentation. In *Proceedings, The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference, July 9-13, 2005, Pittsburgh, Pennsylvania, USA*, pages 608–613, 2005.

[CA07]   Martin Caminada and Leila Amgoud. On the evaluation of argumentation formalisms. *Artif. Intell.*, 171(5-6):286–310, 2007.

[Cam06a]   Martin Caminada. On the issue of reinstatement in argumentation. In *Logics in Artificial Intelligence, 10th European Conference, JELIA 2006, Liverpool, UK, September 13-15, 2006, Proceedings*, pages 111–123, 2006.

[Cam06b]   Martin Caminada. Semi-stable semantics. In *Computational Models of Argument: Proceedings of COMMA 2006, September 11-12, 2006, Liverpool, UK*, pages 121–130, 2006.

[Chi66]   Roderick M. Chisholm. *Theory of Knowledge*. Prentice Hall, 1966.

[CMDK^{+}07]   Sylvie Coste-Marquis, Caroline Devred, Sébastien Konieczny, Marie-Christine Lagasquie-Schiex, and Pierre Marquis. On the merging of dung's argumentation systems. *Artif. Intell.*, 171(10-15):730–753, 2007.

[CML00]   Carlos Iván Chesñevar, Ana Gabriela Maguitman, and Ronald Prescott Loui. Logical models of argument. *ACM Computing survey*, 32(4):337–383, 2000.

[CS07]   Carlos Iván Chesñevar and Guillermo Ricardo Simari. A lattice-based approach to computing warranted beliefs in skeptical argumentation frameworks. In *IJ-*

CAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007, pages 280–285, 2007.

[DB04]   Paul E. Dunne and Trevor J. M. Bench-Capon. Complexity in value-based argument systems. In *Logics in Artificial Intelligence, 9th European Conference, JELIA 2004, Lisbon, Portugal, September 27-30, 2004, Proceedings*, pages 360–371, 2004.

[DBC01]   Paul E. Dunne and T.J.M. Bench-Capon. Complexity and combinatorial properties of argument systems. In *Technical report, Dept. of Comp. Sci., Univ. of Liverpool (July 2001)*, 2001.

[DBC02]   Paul E. Dunne and Trevor J. M. Bench-Capon. Coherence in finite argument systems. *Artif. Intell.*, 141(1/2):187–203, 2002.

[DHM$^+$11]   Paul E. Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artif. Intell.*, 175(2):457–486, 2011.

[DKT06]   Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. Dialectic proof procedures for assumption-based, admissible argumentation. *Artif. Intell.*, 170(2):114–159, 2006.

[DNT99]   Yannis Dimopoulos, Bernhard Nebel, and Francesca Toni. Preferred arguments are harder to compute than stable extension. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, IJCAI 99, Stockholm, Sweden, July 31 - August 6, 1999. 2 Volumes, 1450 pages*, pages 36–43, 1999.

[DNT00]   Yannis Dimopoulos, Bernhard Nebel, and Francesca Toni. Finding admissible and preferred arguments can be very hard. In *KR 2000, Principles of Knowledge*

Representation and Reasoning Proceedings of the Seventh International Conference, Breckenridge, Colorado, USA, April 11-15, 2000., pages 53–61, 2000.

[DNT02]   Yannis Dimopoulos, Bernhard Nebel, and Francesca Toni. On the computational complexity of assumption-based argumentation for default reasoning. *Artif. Intell.*, 141(1/2):57–78, 2002.

[DPPS09]   Jürgen Dix, Simon Parsons, Henry Prakken, and Guillermo Ricardo Simari. Research challenges for argumentation. *Computer Science - R&D*, 23(1):27–34, 2009.

[DS95]   Phan Minh Dung and Tran Cao Son. Nonmonotonic inheritance, argumentation and logic programming. In *Logic Programming and Nonmonotonic Reasoning, Third International Conference, LPNMR'95, Lexington, KY, USA, June 26-28, 1995, Proceedings*, pages 316–329, 1995.

[DS01]   Phan Minh Dung and Tran Cao Son. An argument-based approach to reasoning with specificity. *Artif. Intell.*, 133(1-2):35–85, 2001.

[DT96]   Yannis Dimopoulos and Alberto Torres. Graph theoretical structures in logic programs and default theories. *Theor. Comput. Sci.*, 170(1-2):209–244, 1996.

[Dun95a]   Phan Minh Dung. An argumentation-theoretic foundations for logic programming. *J. Log. Program.*, 22(2):151–171, 1995.

[Dun95b]   Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, 77(2):321–358, 1995.

[DW09]   Paul E. Dunne and Micheal Wooldridge. Complexity of abstract argumentation. *Argumentation in Artificial Intelligence (Ed: I. Rahwan and G. Simari)*, pages 85–104, 2009.

[FKIS02]   Marcelo A. Falappa, Gabriele Kern-Isberner, and Guillermo Ricardo Simari. Explanations, belief revision and defeasible reasoning. *Artif. Intell.*, 141(1/2):1–28, 2002.

[FNL13]   Norman E. Fenton, Martin Neil, and David A. Lagnado. A general structure for legal arguments about evidence using bayesian networks. *Cognitive Science*, 37(1):61–102, 2013.

[GMAB01]   Guido Governatori, Michael J. Maher, Grigoris Antoniou, and David Billington. Representation results for defeasible logic. *ACM Transactions on Computational Logic*, 2:255–287, 2001.

[GMAB04]   Guido Governatori, Michael J. Maher, Grigoris Antoniou, and David Billington. Argumentation semantics for defeasible logic. *J. Log. Comput.*, 14(5):675–702, 2004.

[GRS07]   Alejandro Javier García, Nicolás D. Rotstein, and Guillermo Ricardo Simari. Dialectical explanations in defeasible argumentation. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 9th European Conference, ECSQARU 2007, Hammamet, Tunisia, October 31 - November 2, 2007, Proceedings*, pages 295–307, 2007.

[GW09]   Thomas F. Gordon and Douglas Walton. Proof burdens and standards. *Argumentation in Artificial Intelligence, , I. Rahwan and G. Simari, Eds. Springer-Verlag, Berlin, Germany*, pages 239–260, 2009.

[HPW04]   Chris Reed Henry Prakken and Douglas Walton. Argumentation schemes and burden of proof. In *Working Notes of the 4th International Workshop on Computational Models of Natural Argument (CMNA2004), ed. F. Grasso, C. Reed and G. Carenini, Valencia*, volume 4, pages 81–86, 2004.

[KMD94]   Antonis C. Kakas, Paolo Mancarella, and Phan Minh Dung. The acceptability se-
mantics for logic programs. In *Logic Programming, Proceedings of the Eleventh
International Conference on Logic Programming, Santa Marherita Ligure, Italy,
June 13-18, 1994*, pages 504–519, 1994.

[KR04]   Joel Katzav and Chris Reed. On argumentation schemes and the natural classifi-
cation of arguments, argumentation. *Argumentation*, 18(2):239–259, 2004.

[KT96]   Robert A. Kowalski and Francesca Toni. Abstract argumentation. *Artif. Intell.
Law*, 4(3-4):275–296, 1996.

[KT99]   Antonis C. Kakas and Francesca Toni. Computing argumentation in logic pro-
gramming. *J. Log. Comput.*, 9(4):515–562, 1999.

[Kyb61]   Henry E. Kyberg. *Probability and the Logic of Rational Belief*. Wesleyan Uni-
versity Press, 1961.

[Lou87]   Ronald Prescott Loui. Defeat among arguments: a system of defeasible inference.
Computational Intelligence, 3:100–106, 1987.

[Lou89]   Ronald Prescott Loui. Defeasible decisions: What the proposal is and isn't. In
*UAI '89: Proceedings of the Fifth Annual Conference on Uncertainty in Artifi-
cial Intelligence, Windsor, Ontario, Canada, August 18-20, 1989*, pages 99–116,
1989.

[McC80]   John McCarthy. Circumscription - A form of non-monotonic reasoning. *Artif.
Intell.*, 13(1-2):27–39, 1980.

[McC86]   John McCarthy. Applications of circumscription to formalizing common-sense
knowledge. *Artif. Intell.*, 28(1):89–116, 1986.

[MGS08]   Diego C. Martínez, Alejandro Javier García, and Guillermo Ricardo Simari. An
abstract argumentation framework with varied-strength attacks. In *Principles of*

Knowledge Representation and Reasoning: Proceedings of the Eleventh International Conference, KR 2008, Sydney, Australia, September 16-19, 2008, pages 135–144, 2008.

[Mod06]   Sanjay Modgil. Value based argumentation in hierarchical argumentation frameworks. In *Computational Models of Argument: Proceedings of COMMA 2006, September 11-12, 2006, Liverpool, UK*, pages 297–308, 2006.

[Mod09]   Sanjay Modgil. Reasoning about preferences in argumentation frameworks. *Artif. Intell.*, 173(9-10):901–934, 2009.

[Moo84]   Robert C. Moore. Possible-world semantics for autoepistemic logic. In *NMR*, pages 344–354, 1984.

[NP06]   Søren Holbech Nielsen and Simon Parsons. A generalization of dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *Argumentation in Multi-Agent Systems, Third International Workshop, ArgMAS 2006, Hakodate, Japan, May 8, 2006, Revised Selected and Invited Papers*, pages 54–73, 2006.

[NP07]   Søren Holbech Nielsen and Simon Parsons. An application of formal argumentation: Fusing bayesian networks in multi-agent systems. *Artif. Intell.*, 171(10-15):754–775, 2007.

[Nut01]   Donald Nute. Defeasible logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 353–395. Oxford University Press, 2001.

[Par98]   Simon Parsons. A proof theoretic approach to qualitative probabilistic reasoning. *Int. J. Approx. Reasoning*, 19(3-4):265–297, 1998.

[Pea88]   Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

[PL92]   Gadi Pinkas and Ronald Prescott Loui. Reasoning from inconsistency: A taxonomy of principles for resolving conflict. In *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR'92). Cambridge, MA, October 25-29, 1992.*, pages 709–719, 1992.

[Pol67]   John L. Pollock. Criteria and our knowledge of the material world. *Philosophical Review*, 76:28–62, 1967.

[Pol87]   John L. Pollock. Defeasible reasoning. *Cognitive Science*, 11(4):481–518, 1987.

[Pol90]   John L. Pollock. Interest driven suppositional reasoning. *J. Autom. Reasoning*, 6(4):419–461, 1990.

[Pol92]   John L. Pollock. How to reason defeasibly. *Artif. Intell.*, 57(1):1–42, 1992.

[Pol94]   John L. Pollock. Justification and defeat. *Artif. Intell.*, 67(2):377–407, 1994.

[Pol96]   John L. Pollock. Reason in a changing world. In *Practical Reasoning, International Conference on Formal and Applied Practical Reasoning, FAPR '96, Bonn, Germany, June 3-7, 1996, Proceedings*, pages 495–509, 1996.

[Pol01]   John L. Pollock. Defeasible reasoning with variable degrees of justification. *Artif. Intell.*, 133(1-2):233–282, 2001.

[Pol02]   John L. Pollock. Defeasible reasoning with variable degrees of justification. *Artif. Intell.*, 133(1):233–282, 2002.

[Pol10]   John L. Pollock. Defeasible reasoning and degrees of justification. *Argument & Computation*, 1(1):7–22, 2010.

[Poo88]   David Poole. A logical framework for default reasoning. *Artif. Intell.*, 36(1):27–47, 1988.

[Poo91]   David Poole. The effect of knowledge on belief: conditioning, specificity and the lottery paradox in default reasoning. In *Artif. Intell.*, volume 49, pages 281–307, 1991.

[Poo92]   David Poole. Logic programming, abduction and probability. In *FGCS*, pages 530–538, 1992.

[Poo93a]   David Poole. Probabilistic horn abduction and bayesian networks. *Artif. Intell.*, 64(1):81–129, 1993.

[Poo93b]   David Poole. The use of conflicts in searching bayesian networks. In *UAI '93: Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence, The Catholic University of America, Providence, Washington, DC, USA, July 9-11, 1993*, pages 359–367, 1993.

[Poo97]   David Poole. Probabilistic partial evaluation: Exploiting rule structure in probabilistic inference. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence, IJCAI 97, Nagoya, Japan, August 23-29, 1997, 2 Volumes*, pages 1284–1291, 1997.

[Poo00]   David Poole. Logic, knowledge representation, and bayesian decision theory. In *Computational Logic*, pages 70–86, 2000.

[Pra01a]   Henry Prakken. Modelling defeasibility in law: Logic or procedure? *Fundam. Inform.*, 48(2-3), 2001.

[Pra01b]   Henry Prakken. Modelling reasoning about evidence in legal procedure. In *Proceedings of the Eigths International Conference on Artificial Intelligence and*

Law, ICAIL 2001, St. Louis, Missouri, USA, May 21-25, 2001, pages 119–128, 2001.

[Pra10]   Henry Prakken. An abstract framework for argumentation with structured arguments. *Argument & Computation*, 1(2):93–124, 2010.

[PRW03]   Henry Prakken, Chris Reed, and Douglas Walton. Argumentation schemes and generalizations in reasoning about evidence. In *Proceedings of the 9th International Conference on Artificial Intelligence and Law, ICAIL 2003, Edinburgh, Scotland, UK, June 24-28, 2003*, pages 32–41, 2003.

[PRW05]   Henry Prakken, Chris Reed, and Douglas Walton. Dialogues about the burden of proof. In *The Tenth International Conference on Artificial Intelligence and Law, Proceedings of the Conference, June 6-11, 2005, Bologna, Italy*, pages 115–124, 2005.

[PS96a]   Henry Prakken and Giovanni Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artif. Intell. Law*, 4(3-4):331–368, 1996.

[PS96b]   Henry Prakken and Giovanni Sartor. A system for defeasible argumentation, with defeasible priorities. In *Practical Reasoning, International Conference on Formal and Applied Practical Reasoning, FAPR '96, Bonn, Germany, June 3-7, 1996, Proceedings*, pages 510–524, 1996.

[PS97]   Henry Prakken and Giovanni Sartor. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics*, 7(1):25–75, 1997.

[PS98]   Henry Prakken and Giovanni Sartor. Modelling reasoning with precedents in a formal dialogue game. *Artif. Intell. Law*, 6(2-4):231–287, 1998.

[PS99]   Henry Prakken and Giovanni Sartor. A system for defeasible argumentation, with defeasible priorities. In *M.J. Wooldndge and M. Veloso (Eds.): Artificial Intelligence Today, LNAI 1600*, pages 365–379. Springer-Verlag Berlin Heidelberg, 1999.

[PS04]   Henry Prakken and Giovanni Sartor. The three faces of defeasibility in the law. In *Ratio Juris, March 2004*, volume 17, pages 118–139. John Wiley & Sons Ltd, 2004.

[PV01]   Henry Prakken and Gerard A. W. Vreeswijk. Logics for defeasible argumentation. In Dov M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 4, pages 219–319. Kluwer Academic Publishers, 2001.

[QBVT05]   Foo N. Y. Quoc Bao Vo and Thurbon. Semantics for a theory of defeasible reasoning. *Annals of Mathematics and Artificial Intelligence*, 44(1-2):87–119, 2005.

[Rei80]   Raymond Reiter. A logic for default reasoning. *Artif. Intell.*, 13(1-2):81–132, 1980.

[Res77]   Nicholas Rescher. *Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge*. State University of New York Press, 1977.

[RG01]   Chris Reed and Floriana Grasso. Computational models of natural language argument. In *Computational Science - ICCS 2001, International Conference, San Francisco, CA, USA, May 28-30, 2001. Proceedings, Part I*, pages 999–1008, 2001.

[RW06]   Chris Reed and Douglas Walton. Towards a formal and implemented model of argumentation schemes in agent communication. *Autonomous Agents and Multi-Agent Systems*, 11(2):173–188, 2006.

[SCG+94]   Guillermo R. Simari, Carlos I. Chesevar, Alejandro J. Garcia, Benjamin C. Pierce, and David N. Turner. The role of dialectics in defeasible argumentation. *In Proceedings of the XIV Conferencia Internacional de la Sociedad Chilena para Ciencias de la Computacion*, pages 111–121, 1994.

[Sim11]   Guillermo Ricardo Simari. A brief overview of research in argumentation systems. In *Scalable Uncertainty Management - 5th International Conference, SUM 2011, Dayton, OH, USA, October 10-13, 2011. Proceedings*, pages 81–95, 2011.

[SL92]   Guillermo Ricardo Simari and Ronald Prescott Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artif. Intell.*, 53(2-3):125–157, 1992.

[TDH12]   Phan Minh Thang, Phan Minh Dung, and Nguyen Duy Hung. Towards argument-based foundation for sceptical and credulous dialogue games. In *Computational Models of Argument - Proceedings of COMMA 2012, Vienna, Austria, September 10-12, 2012*, pages 398–409, 2012.

[TK95]   Francesca Toni and Antonis C. Kakas. Computing the acceptability semantics. In *Logic Programming and Nonmonotonic Reasoning, Third International Conference, LPNMR'95, Lexington, KY, USA, June 26-28, 1995, Proceedings*, pages 401–415, 1995.

[Tou58]   Stephen E. Toulmin. *The Uses of Argument*. Cambridge University Press, 1958.

[vE95]   Frans Hendrik van Eemeren. The rich state of argumentation theory. *INFORMAL LOGIC*, 17(2):144–158, 1995.

[Ver01a]   Bart Verheij. Accrual of arguments in defeasible argumentation. *Dutch/German Workshop on Nonmonotonic Reasoning*, pages 217–224, 2001.

[Ver01b]   Bart Verheij. Douglas walton, the new dialectic, conversational contexts of argument, toronto: University of toronto press (review). *Artif. Intell. Law*, 9(4):305–313, 2001.

[Ver05]   Bart Verheij. Evaluating arguments based on toulmin's scheme. In *David Hitchcock, Bart Verheij (Ed.), Argumentation*, volume 19, pages 347–371. Kluwer Academic Publishers, 2005.

[Ver07]   Bart Verheij. A labeling approach to the computation of credulous acceptance in argumentation. In *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007*, pages 623–628, 2007.

[Ver09]   Bart Verheij. The toulmin argument model in artificial intelligence. In *I. Rahwan, G. R. Simari (eds.), Argumentation in Articial Intelligence*, pages 219–238. Springer Publishing Company, Incorporated, 2009.

[Vre97]   Gerard Vreeswijk. Abstract argumentation systems. *Artif. Intell.*, 90(1-2):225–279, 1997.

[Vre04]   Gerard Vreeswijk. Argumentation in bayesian belief networks. In *Argumentation in Multi-Agent Systems, First International Workshop, ArgMAS 2004, New York, NY, USA, July 19, 2004, Revised Selected and Invited Papers*, pages 111–129, 2004.

[Vre06]   Gerard Vreeswijk. An algorithm to compute minimally grounded and admissible defence sets in argument systems. In *Computational Models of Argument: Proceedings of COMMA 2006, September 11-12, 2006, Liverpool, UK*, pages 109–120, 2006.

[Wal97]   Douglas Walton. How can logic best be applied to arguments? *Logic Journal of*

the IGPL, 5(4):603–614, 1997.

[Wal02]   Douglas N. Walton. *Legal argumentation and evidence*. Pennsylvania State Univ. Pr., 2002.

[Wal03a]   Douglas Walton. Is there a burden of questioning? *Artif. Intell. Law*, 11(1):1–43, 2003.

[Wal03b]   Douglas N. Walton. *Relevance in argumentation*. Lawrence Erlbaum, 2003.

[Wal07]   Douglas Walton. Dialogical models of explanation. In *Explanation-Aware Computing, Papers from the 2007 AAAI Workshop, Vancouver, British Columbia, Canada, July 22-23, 2007*, pages 1–9, 2007.

[Wal08]   Douglas Walton. The three bases for the enthymeme: A dialogical theory. *J. Applied Logic*, 6(3):361–379, 2008.

[Wal10]   Douglas Walton. Types of dialogue and burdens of proof. In *Computational Models of Argument: Proceedings of COMMA 2010, Desenzano del Garda, Italy, September 8-10, 2010.*, pages 13–24, 2010.

[WMP05]   Michael Wooldridge, Peter McBurney, and Simon Parsons. On the meta-logic of arguments. In *Argumentation in Multi-Agent Systems, Second International Workshop, ArgMAS 2005, Utrecht, The Netherlands, July 26, 2005, Revised Selected and Invited Papers*, pages 42–56, 2005.

Index of Definitions

A

accepted by, 34

active

    active attacker, 79

    active defender, 79

active attack sequences

    active, 82

    intercepted, 82

    negatively active, 82

    partially intercepted, 82

    positively active, 82

adjacent intercepting sub-frameworks, 120

admissibility status, 35

admissible, 35

argumentation framework operations $+^N$, $-^N$, 111

attack cycle, 39

B

backings

    γ, 55

    admissibility backing, 47

C