Alma Mater Studiorum – Università di Bologna
in cotutela con University of Luxembourg - Universitè du Luxembourg

DOTTORATO DI RICERCA IN

LAW, SCIENCE AND TECHNOLOGY

Ciclo 35

Settore Concorsuale: 01/B1 – INFORMATICA

Settore Scientifico Disciplinare: INF/01 - INFORMATICA

DISTRIBUTED ARGUMENTATION TECHNOLOGY: ADVANCING RISK ANALYSIS AND REGULATORY COMPLIANCE OF DISTRIBUTED LEDGER TECHNOLOGIES FOR TRANSACTION AND MANAGEMENT OF SECURITIES

Presentata da: Liuwen YU

Coordinatore Dottorato

Monica PALMIRANI

Supervisore

Leon VAN DER TORRE

Supervisore

Monica PALMIRANI

Co-supervisore

Massimo DURANTE

Esame finale anno 2023

UNIVERSITÉ DU LUXEMBOURG

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

PhD--FSTM-2023-127
The Faculty of Science, Technology and Medicine

The Department of Legal Studies

DISSERTATION

Defence held on 26/10/2023 in Bologna

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG
EN INFORMATIQUE
AND
DOTTORE DI RICERCA
IN LAW, SCIENCE AND TECHNOLOGY

by

Liuwen YU
Born on 27 September 1994 in Nei Mongol (China)

DISTRIBUTED ARGUMENTATION TECHNOLOGY:
ADVANCING RISK ANALYSIS AND REGULATORY COMPLIANCE OF DISTRIBUTED LEDGER TECHNOLOGIES (DLT) FOR TRANSACTION AND MANAGEMENT OF SECURITIES

Dissertation defence committee

Dr. Bart VERHEIJ, Chairman
Professor, University of Groningen

Dr. Massimiliano GIACOMIN
Professor, University of Brescia.

Dr. Christian STRASSER
Professor, Ruhr University Bochum

Dr. Leon VAN DER TORRE, Supervisor
Professor, University of Luxembourg

Dr. Monica PALMIRANI, Supervisor
Professor, University of Bologna

Dr. Massimo DURANTE, Co-supervisor
Professor, University of Turin

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Leendert van der Torre, who leads the Individual and Collective Research Group (ICR). His full support, constructive feedback, and invaluable inspiration have been instrumental in my development and research progress. I would also like to extend my gratitude to Prof. Monica Palmirani, the coordinator in charge of the LAST-JD project, for her persistent endeavors.

I would like to express my gratitude to Dr. Réka Markovich, my advisor and the head of the Computational Law and Machine Ethics group (CLAiM). The discussions with her have enriched my understanding and broadened my horizons. Through her, I also witnessed the profound strength and influence of women in academia. I am also grateful to Dr. Davide Liga. Together with Dr. Markovich, they took great care of the LAST-JD students in Luxembourg.

Prof. Beishui Liao, together with my supervisor, initiated the Zhejiang University – the University of Luxembourg Joint Lab on Advanced Intelligent Systems and Reasoning (ZLAIRE). I am grateful that ZLAIRE provided me with a platform to experience international collaborations in scientific research, teaching, and outreaches.

I want to thank all my co-authors: Caren Al Anaissy, Dongheng Chen, Chenchen, Dov Gabbay, Réka Markovich, Amro Najjar, Xu Li, Pere Pardo, Lisha Qiao, Yiqi Shen, Leendert van der Torre, Srdjan Vesic, and Mirko Zichichi. Their expertise, feedback, and commitment were essential in completing our papers.

Given the nature of my PhD project, I was lucky to have the support of two colleague groups: the Luxembourg group and the Bologna group. The ICR& CLAiM group in Luxembourg consistently stood out, fostering a warm and enriching atmosphere. I'd like to thank all the members. I would also like to thank my colleagues from LAST-JD in Bologna for their companionship. Especially during the challenges of the pandemic, they offered invaluable help.

Last but not least, I want to thank my family. Being 9,558 km away, especially during the pandemic, I can only imagine their concerns for me. Studying abroad is as much an adventure for me as it is a journey of faith and trust for my family.

Abstract

Distributed argumentation technology is a computational approach incorporating argumentation reasoning mechanisms within multi-agent systems. For the formal foundations of distributed argumentation technology, in this thesis we conduct a principle-based analysis of structured argumentation as well as abstract multi-agent and abstract bipolar argumentation. The results of the principle-based approach of these theories provide an overview and guideline for further applications of the theories. Moreover, in this thesis we explore distributed argumentation technology using distributed ledgers. We envision an Intelligent Human-input-based Blockchain Oracle (IHiBO), an artificial intelligence tool for storing argumentation reasoning. We propose a decentralized and secure architecture for conducting decision-making, addressing key concerns of trust, transparency, and immutability. We model fund management with agent argumentation in IHiBO and analyze its compliance with European fund management legal frameworks. We illustrate how bipolar argumentation balances pros and cons in legal reasoning in a legal divorce case, and how the strength of arguments in natural language can be represented in structured arguments. Finally, we discuss how distributed argumentation technology can be used to advance risk management, regulatory compliance of distributed ledgers for financial securities, and dialogue techniques.

Contents

Chapter 1

Introduction

This thesis is involved in the Law, Science, and Technology Joint Doctoral Program (LAST-JD) funded by Marie Skłodowska-Curie Actions. The advent of new AI technologies brings forth technical, legal, and ethical challenges in a society that may not yet be fully equipped to address them. Recognizing this, the project aims to emphasize that a comprehensive understanding of these challenges extends beyond just engineering and technology. It necessitates a variety of different scientific and interdisciplinary approaches as well as technical, legal, and economic points of view. In this thesis, we propose distributed argumentation technology, a computational approach incorporating argumentation reasoning mechanisms within multi-agent systems. For the formal foundations of distributed argumentation technology, in this thesis we conduct a principle-based analysis of structured argumentation as well as abstract multi-agent and abstract bipolar argumentation. The results of the principle-based approach of these theories provide an overview and guideline for further applications of the theories. Moreover, in this thesis, we explore distributed argumentation technology using distributed ledgers. We envision an Intelligent Human-input-based Blockchain Oracle (IHiBO), an artificial intelligence tool for storing argumentation reasoning. IHiBO provides a decentralized and secure architecture for conducting decision-making, addressing key concerns of legal, trust, transparency, and auditability aspects. Both the LAST-JD program and this thesis share a focus on encompassing the key elements from various interdisciplinary fields.

1.1 Background

The background of this thesis is illustrated in Figure 1.1. Computer science is a dynamic branch of science that focuses on the study of algorithms, computation, and information processing. A key subfield, artificial intelligence (AI), aims to create intelligent systems with human-like capabilities [283, 171]. Its progressive strides in recent years have propelled AI to the forefront of technological advancement, significantly influencing many fields such as healthcare, transportation, finance, etc. [342]. The Association for the Advancement of Artificial Intelligence (AAAI) and the Computer Science Teachers Association (CSTA) [1] proposed five big ideas of AI, namely the *AI for K-12 Initiative*. It is dedicated to AI education to K-12 students, who engage in the concepts and practices.

[1] https://ai4k12.org/

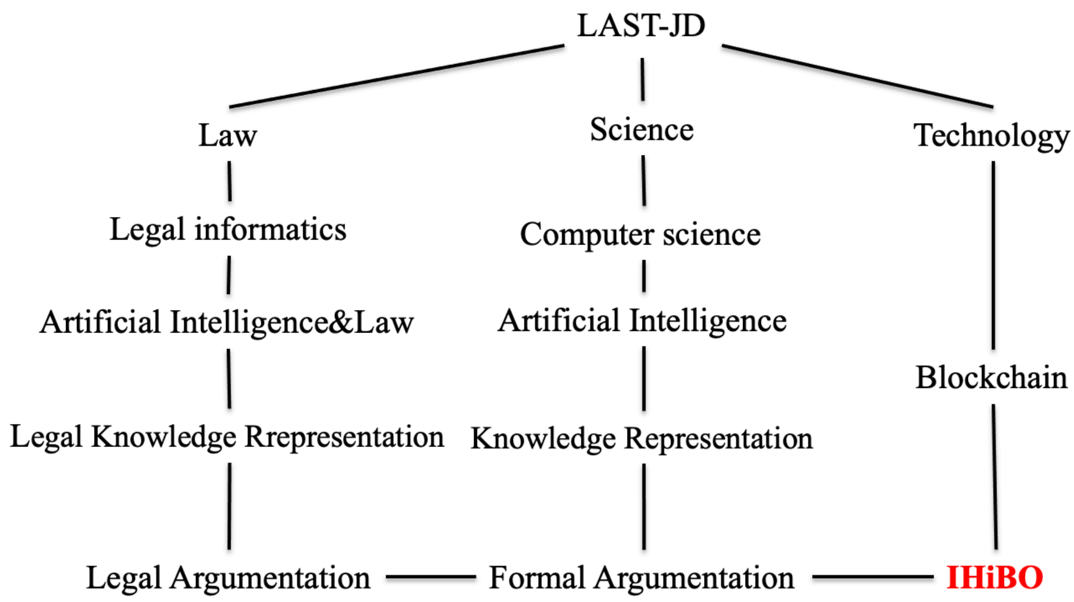


Figure 1.1: Background of IHiBO

These five ideas are Perception, Representation and Reasoning, Learning, Natural Interaction, and Societal Impact. In the early 21st century, the five ideas were primarily focused on research and were not widely implemented in practical applications. However, over time, advancements in computing power and AI have led to significant progress in various domains. Perception, which involves tasks such as computer vision and speech recognition, has seen tremendous development and is now being applied in real-world applications like autonomous vehicles [169], facial recognition systems [15], and voice assistants [162]. Learning algorithms have also become more sophisticated, enabling AI systems to improve their performance through experience and training on large datasets [203]. This has paved the way for practical applications like recommendation systems, fraud detection, and personalized user experiences [286]. Natural interaction, involving the ability of AI systems to understand and respond to human language and gestures, has seen significant progress as well. Chatbots like ChatGPT exemplify the practical implementation of natural interaction in AI. Societal impact, which encompasses the ethical and societal implications of AI, has become an increasingly important consideration in the development and deployment of AI systems [173]. Industry and researchers are actively working on addressing issues like bias, privacy, and fairness in AI applications [193]. While these areas have seen significant progress and practical implementations, representation and reasoning, still remain more focused on research and comparatively lack widespread practical applications.

Knowledge representation and reasoning are concerned with applying reasoning in the form of logic. The origins of formal logic can be traced back to Aristotle, over two millennia ago. In the present era of computer science, formal logic assumes a vital role in reasoning about the correctness and behavior of algorithms, programs, and systems. Within the domain of computer science, mathematical logic, including propositional logic and first-order logic, is predominantly employed for tasks such as specification and verification [165]. The utilization of modal logic becomes paramount when addressing more intricate concepts, necessitating dedicated languages that surpass the generality of first-order logic. Modal logic finds extensive application not only in philosophical logic but also in computer science, facilitating specialized education in this realm [141].

Parallel to the development of formal logic, the exploration of reasoning through argumentation has been a distinctive pursuit since the time of Aristotle. Argumentation, deeply rooted in human history as a fundamental practice, has captivated scholarly attention, particularly within the realm of theoretical philosophy. Toulmin and his influential book The *Uses of Argument* challenging the "classic logicians" view for neglecting many features of common-sense reasoning marked a turning point [311]. Toulmin argued that not all practical arguments need to be deductively valid. Inspired by Toulmin's critique, the field of informal logic, encompassing non-monotonic logic [199], had become fashionable around 1980. Non-monotonic logic acknowledges that in common-sense reasoning, acquiring additional knowledge may necessitate the revision or abandonment of earlier conclusions, departing from the monotonic property. For instance, the well-known "Tweety" example demonstrates that while birds typically fly, if Tweety is identified as a penguin, the earlier conclusion that Tweety can fly must be withdrawn. The extensive work on argumentation from Aristotle to today's computational argumentation in AI shows how far research in argumentation has come. Notably, the two volumes of the *Handbook of Formal Argumentation* [31, 134] have played a pivotal role in advancing our understanding and practical implementation of argumentation, both in philosophical and computational realms.

Law, an intricate and ever-evolving domain, plays a pivotal role in shaping and safeguarding society. As technology continues its inexorable march, the convergence of law and technology has given birth to new frontiers of study and practice. One such frontier is legal informatics [291], which delves into the application of information technology and computer science to legal systems and processes. Within this realm, AI and Law stand as a prominent subfield, emphasizing the utilization of AI technologies in the legal domain [151, 287, 321]. The transformative potential of AI in the legal profession spans a wide spectrum, encompassing tasks ranging from legal research and document analysis to the bedrock of legal reasoning and decision-making. By harnessing the power of machine learning, natural language processing, and other AI techniques, legal professionals can elevate their efficiency, precision, and overall effectiveness in delivering legal services.

At the heart of the legal system lie two fundamental pillars: legal representation and reasoning. These pillars converge in the realm of legal argumentation, where persuasive arguments, fortified by legal principles, precedents, and factual evidence, are crafted to bolster a specific legal position. The advent of computational argumentation has further augmented this process, with AI technologies now offering assistance in the analysis of legal arguments [151]. These technologies excel at evaluating the strengths and weaknesses of arguments, and can even generate counter-arguments. This symbiotic integration not only expedites legal research endeavors but also engenders more robust and well-informed legal discourses. Law by its very nature serves as an ideal test bed for the exploration and refinement of formal argumentation methodologies. In this context, the Handbook of Legal AI assumes paramount importance [319]. This comprehensive resource serves as a guiding beacon, providing invaluable insights and frameworks for the study and application of AI in the legal domain.

In the context of technology, this thesis leverages distributed ledger technology (DLT). DLT operates as a decentralized system, wherein multiple replicas of a shared ledger are distributed among network participants [299]. This ensures transparency, security, and consensus without necessitating a central authority [118]. A prominent and universally acknowledged manifestation of DLT is blockchain. In blockchain, a specific form of DLT, data is compartmentalized into blocks, with each block cryptographically linked to its predecessor [224]. This architecture guarantees data immutability and transparency, establishing blockchain as a favored network. The applications of blockchain are vast, such as cryptocurrency, financial services, risk management, and more [205].

Having introduced the broader context of this thesis, it's essential to introduce the core of distributed argumentation technology: formal argumentation.

1.2 Formal Argumentation

The field of argumentation research, spanning philosophy, AI, linguistics, and more, is vast and multifaceted. This section offers a succinct overview of the historical progression of formal argumentation during the late 20th century. For a comprehensive examination of the foundations of argumentation, as well as its connection with the broader literature, we refer the reader to the Handbook of Argumentation Theory [316] and two volumes of the *Handbook of Formal Argumentation* [31, 134]. In this section, we delineate the evolution of formal argumentation, present the three major branches of this field, and illustrate how argumentation dialogue transitions into an abstract argumentation framework, using a running example.

1.2.1 Brief History of Formal Argumentation

In his book *The Uses of Arguments* [311], Toulmin pointed out that deductive logic is not sufficient for modeling human reasoning, given its limitation to cover all aspects, for instance, the existence of counterarguments, i.e., inconsistent information in argumentation. Different from deductive reasoning which aims at preserving certainty of conclusion from true premises, argumentation in the first place is a form of nonmonotonic logic, that became fashionable around 1980. It helps find a conclusion in reasoning with inconsistent and incomplete information, which is an important aspect of intelligence. Pollock published his seminal paper in 1987 [253], arguably, since then, the idea arose in the field of argumentation that non-monotonic inference rules can be used to model arguments. Pollock proposed the important philosophical notion of defeasible reasoning that is closer to the character of human argumentation and commonsense reasoning. Also in 1987, Loui explicitly designed nonmonotonic logic in the argumentation way [200], which was extended and fully formalized by Simari and Loui in 1992 [295]. Their work in turn led to the development of Defeasible Logic Programming [138, 139]. Another relevant early work was the work by Nute [230], later developed into so-called Defeasible Logic [231]. Around the year 1990, there were some papers that proposed argumentation as a proof theory for model-theoretic notions of nonmonotonic consequence [29, 142]. Such works see an argument as a set of consistent assumptions with conclusions, and an attack arises when there is a negation of the attacked argument or one of the assumptions. This idea later became the basis for assumption-based argumentation (ABA) [58].

Then in 1995, as witnessed as a turning point of modern argumentation, Dung presented his influential formalism of abstract argumentation frameworks (AAF) [110]. AAF is a directed graph in which the nodes represent arguments and the arrows the defeat relations among these arguments, based on which, the argument evaluation, i.e., acceptability status can be formalized using the notion of semantics. One of the strengths of AAF is its powerful generality. Dung compares the idea of AAF with non-monotonic logic, and logic programming semantics, showing that they are special cases of AAF. Besides, it is also a general framework as an instance of game theory and social choice. In other words, various formalisms for defeasible reasoning can be represented using graphs-based argumentation in a manner that is both intuitive and adequate when applying abstract argumentation semantics to the resulting graphs. The process of argumentation transforms a knowledge base into an argumentation graph and uses formal argumentation methods to derive a set of acceptable conclusions. This process involves constructing arguments based on the knowledge base, establishing attack relations between the arguments, and using argumentation semantics to obtain sets of acceptable arguments and conclusions. The exact way of argument and attack construction as well as applied semantics can be varied based on the specific reasoning context, which was identified as a way to understand and reconcile differences (as well as point out similarities) between various formalisms for non-monotonic, default reasoning as studied in AI.

In 2004, Governatori et al. studied to which extent defeasible logic can be reformulated in terms of Dung's theory of abstract argumentation frameworks [152]. An important development is the study of rationality postulates introduced by Caminada and Amgoud [71, 72], and later extended by Caminada et al. in 2012 [75] and Wu and Podlaszewski in 2015 [333]. They proposed several properties that any argumentation system should fulfill. Structured argumentation is a family of formal approaches for the handling of defeasible and potentially inconsistent information. For this, many models of structured argumentation distinguish between strict and defeasible inference rules. In the 2010s, presented by Prakken and Modgil, ASPIC+ became known as a form of structured argumentation that unifies defeasible reasoning by specifying how to construct arguments and attacks (defeats with preference consideration) structurally [218]. Actually, the work of Caminada and Amgoud in 2007 [72] was inspired by the predecessor of ASPIC+. Since then, it has been shown that the specific families of nonmonotonic logic can indeed be embedded or expressed in argumentation, by the hard work on translations between theories.

As stated in the Handbook of Argumentation Theory: "The general objective of argumentation theory is, in the end, a practical one: to provide adequate instruments for analyzing, evaluating, and producing argumentative discourse." Formal argumentation can present the non-monotonic notion of logical consequence in the form of argument construction, argument relations, and argument evaluation with the aim of resolving conflicts among arguments. Both Pollock [253] and Dung [110] introduced the two key ideas of the formal study of argumentation as inference. Pollock introduced the notion of a defeasible reason, while Dung showed that argument evaluation can be formalized by assuming just two primitive notions of argument and attack. Neither of these ideas on their own defines the field; it is their combination that makes the argumentation way of doing nonmonotonic logic so powerful [316].

1.2.2 Three Branches of Formal Argumentation

Formal argumentation is developed to be a rich and multidimensional field that encompasses various perspectives and approaches to the study of reasoning, persuasion, and decision-making. It offers a systematic framework for analyzing and evaluating arguments, taking into account their logical structure, the dynamics of dialogue and negotiation, and the need to strike a balance between conflicting viewpoints. In formal argumentation, different branches emerge, including argumentation as inference, which includes abstract and structured argumentation [110, 218, 310]; argumentation as dialogue [17], which explores multiagent systems and strategic interactions; and argumentation as balancing [148], which finds applications in domains such as law and ethics, as shown in Figure 1.2.

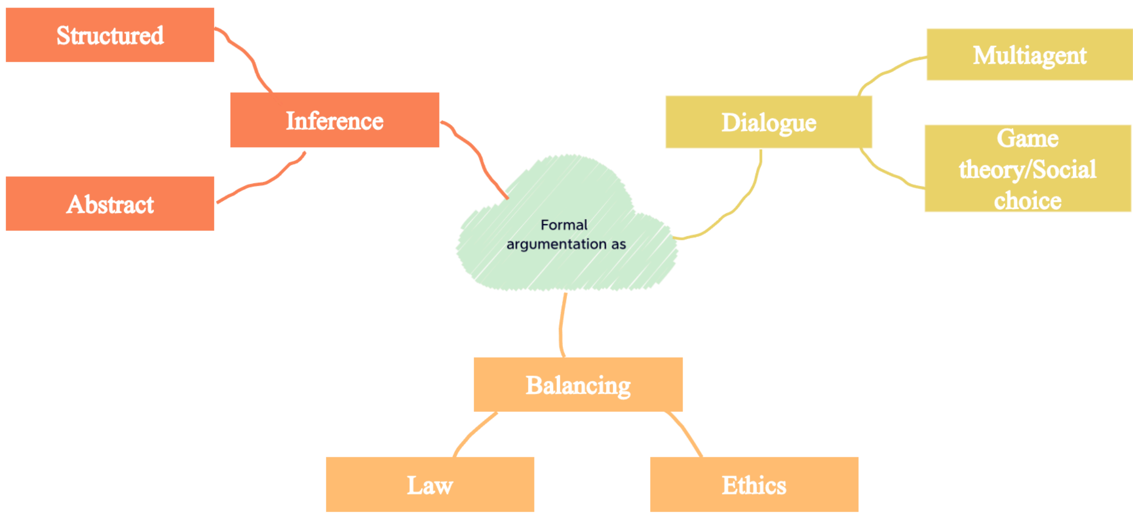


Figure 1.2: Three branches in formal argumentation

As categorized by Henry Prakken in the first *Handbook of Formal Argumentation*, formal argumentation can be distinguished by argumentation as inference and argumentation as dialogue [31, Chap.2]. In systems for argumentation as inference, the focus lies in determining the conclusions that can be derived from a given body of information, which may be incomplete, inconsistent, or uncertain. These systems establish a nonmonotonic notion of logical consequence, employing concepts such as argument construction, argument attack, and argument evaluation. Arguments are regarded as constellations of premises, conclusions, and inferences, forming the foundation for reasoning within this framework. On the other hand, systems for argumentation as dialogue conceptualize argumentation as a form of verbal interaction aimed at resolving conflicts of opinion. These systems define argumentation protocols, which serve as the rules of the argumentation game, and address strategic aspects that guide effective engagement in the game. The exploration of strategies involves understanding how to engage in productive discourse and effectively present arguments. Both aspects of formal argumentation, inference, and dialogue, are examined through formal and computational models, with a review of their historical influences to provide a comprehensive understanding of their development and application.

In Chapter 3 of the first *Handbook of Formal Argumentation*, Thomas Gordon proposes an alternative definition of argumentation with the importance of argumentation for making justified decisions, not only when resolving conflicts of opinion in persuasion dialogues, but also, e.g., when deciding courses of action in deliberation dialogues [31, Chap.3]. He then gives a new definition of argumentation: Argumentation is a rational process, typically in dialogues, for making and justifying decisions of various kinds of issues, in which arguments pro and con alternative resolutions of the issues (options or positions) are put forward, evaluated, resolved and balanced. Argumentation as balancing finds significant applications in the realms of law and ethics. In these domains, the objective is not merely to assess the validity or strength of individual arguments, but to strike a balance between conflicting viewpoints or interests. Balancing involves weighing different considerations, evaluating the relative importance of arguments, and reaching decisions that are ethically sound and legally justifiable.

1.2.3 From Argumentation Dialogue to Argumentation Inference

We start this section with a dialogue example to present the intuition behind formal argumentation, which later can be presented by an abstract argumentation framework.

Example 1.1. *Alice and Lucy are talking about a divorce case, concerning the child's best interest that she lives with her mother or that the child's best interest is that she lives with her father. The dialogue is illustrated in Figure 1.3.*

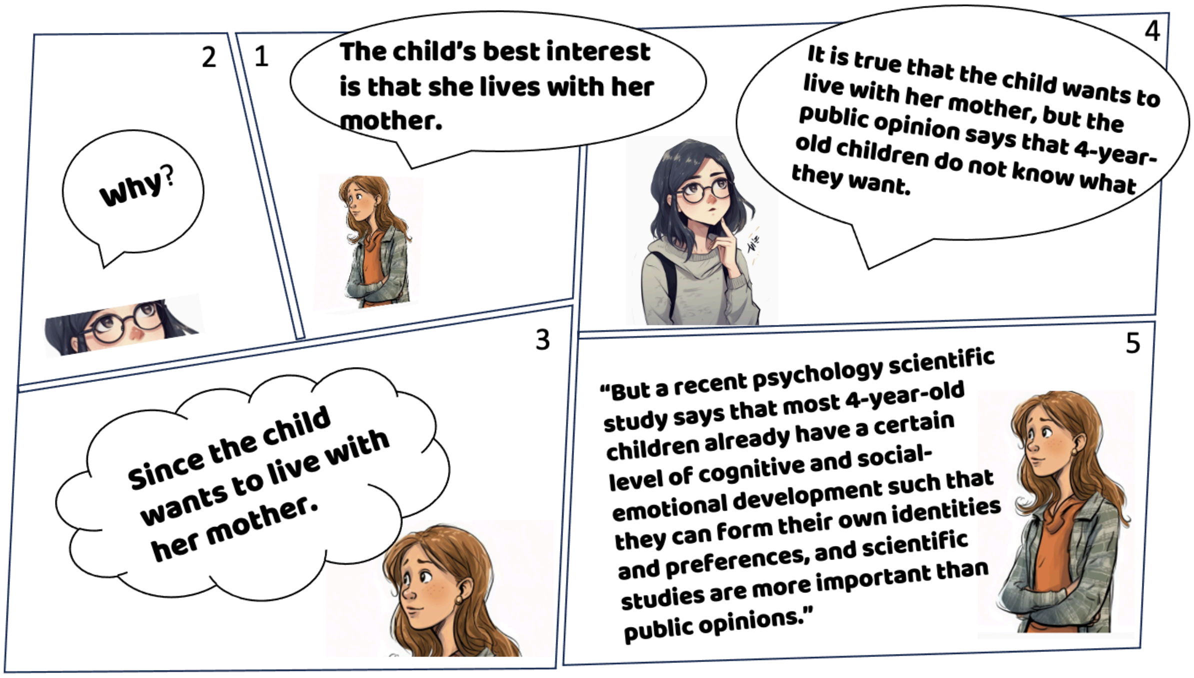


Figure 1.3: The dialogue between two agents on child custody

Having considered the example dialogue, we next show how structured argumentation can represent such discourses. ASPIC+ is a system of structured argumentation that specifies how arguments are constructed relative to a premise set and a number of inference rules [218]. Premises are formulas in a given formal language. They represent the evidence or information on the basis of which we build arguments. Besides premises, we have rules at our disposal for inferring new formulas from others. Arguments are considered the result of applying inference rules to the given premises and, possibly, of chaining such applications. In applications of automated reasoning, a given set of premises and a given set of inference rules form a knowledge base from which arguments are generated. Such a knowledge base can be used to model, for example, legal argumentation. ASPIC+ allows for a distinction between strict and defeasible rules. The acceptable arguments are determined using the semantics of abstract argumentation. To be able to apply these semantics, an argumentation graph has to be constructed on the basis of the argumentation system. This is done by building arguments using the premises and rules of the argumentation system and specifying attacks between these arguments on the basis of the contrariness function and the preference order of the argumentation system. Consider Example 1.1 presented in ASPIC+.

Example 1.2 (Example 1.1 Continued). *The dialogue between Alice and Lucy can be illustrated formally by structured argumentation, where arguments are a chain of applied rules to the premise. As shown in Figure 1.4.*

$a : want(mother) \Rightarrow livewith(mother), b : say(po, \neg know(want)) \Rightarrow \neg know(want) \Rightarrow \neg(wantmother \Rightarrow livewith(mother)), c : say(ss, know(want)) \Rightarrow know(want).$

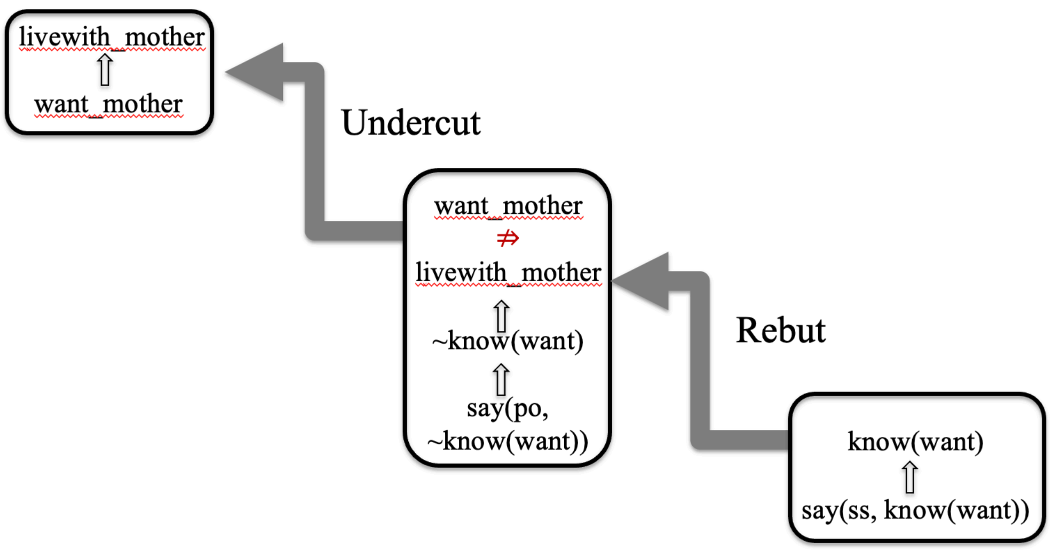


Figure 1.4: Structured argumentation formalizing the dialogue

Example 1.3. *The discussion between Alice and Lucy can be illustrated formally by a directed graph. In the associated graph each node indicates an argument and each arrow shows an attack between arguments. In Figure 1.5 a directed arrow from c to a represents that there is a conflict between these two arguments and argument c attacks argument a. The graph in Figure 1.5 is a formal way of presenting the discussion between Alice and Lucy. The reasonable extension of arguments is thus, $\{a, c\}$, given that c is not being disputed, i.e., there is no argument attacking c, although a is attacked by b, c in turn attacks b, such that it reinstate a.*

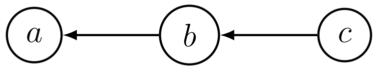


Figure 1.5: Abstract argumentation framework illustrating the dialogue

1.3 Research Questions

Having set the stage with the background of this thesis and a brief overview of formal argumentation, we now turn to the research questions addressed in this thesis. They are as follows.

Question 1 How to construct argumentation frameworks? Argumentation framework construction involves building arguments and assigning attacks. For the latter, it's crucial not only to identify conflicting viewpoints but also to evaluate and consider the relative strength of each argument. The strength of arguments largely depends on the reasoning context, which is key for resolving the conflicts between arguments. For example, in epistemic contexts, an argument is considered stronger since it is based on more reliable information and the inferences from this information have more capacity to conserve such reliability. In the legal reasoning context, it comes to the degrees of deontic or legal urgency (e.g., in view of the authority which issued them, etc.). Formal argumentation offers ways of argument construction and evaluation. The strength of arguments can be compared by the inferences applied in its structured construction, by assigning a (qualitative or quantitative) degree of strength to the defeasible rules in our knowledge base. When obtaining a measure of argument strength concerns multiple rules applied, we need so-called lifting principles. The difference in lift principles may give rise to quite different outcomes. For example, weakest link principle is based on the intuition that an argument is as strong as its weakest rule, it is defined purely in terms of the strength of the defeasible rules used in argument construction. While the last link principle concerns not only the strength of the defeasible rules but also the order in which these rules are applied in the construction of the argument. There is no consensus in the formal argumentation field on which lift principles should be used in a certain context. For instance, Pollock argued that weakest link is appropriate in an epistemic setting [256], while Prakken argues last link is more appropriate in a legal setting [218]. The subquestion we need to answer in this part is: *How to define and compare the attack relation assignments corresponding to weakest link?* Chapter 2 studies how to construct arguments and their relations when looking into their structure, i.e., by comparing the strength of rules applied in argument constructions. We introduce a new definition of the weakest link attack relation assignment and compare this new definition with two existing ones in the literature using a principle-based analysis. We follow Dung's two seminal papers in 2016 [113] and 2018 [115], which will be discussed in detail in Chapter 2.

Question 2 How to accept arguments in agent argumentation? Argumentation typically unfolds within the context of multiple agents, concerning the exchange of viewpoints, justifications, and counterarguments among diverse participants. For instance, legal reasoning can be a challenging task, particularly when it involves the need to balance opposing arguments from different agents involved in a court case. In a constitutional court, a panel of judges may be required to make a decision on a case. In other scenarios, a judge might need to conduct collective argumentation based on the individual argumentation of the accused, prosecutors, witnesses, lawyers, and experts. Besides, agents involved may have partial knowledge of the arguments. The second research question is thus how to accept arguments that are proposed, trusted, or known by multi-agents. The subquestions studied here are: *What kind of semantics can be defined for agent argumentation frameworks? Which of the principles proposed in the literature [35, 314] does not hold for such agent semantics? What new principles can we define to distinguish the varieties of agent semantics?* In Chapter 3, we study abstract agent argumentation, that extends Dung's theory with agents.

Question 3 How to accept arguments in bipolar argumentation? Real-world argumentation often involves not only conflicts but also cooperative aspects where arguments support each other. In evaluating legal argumentation and judicial decisions, one significant role is con arguments which make part of the justification of the standpoint [249]. In other words, it is not solely focused on the arguments put forward in favor of a standpoint, but also on its opposing arguments. At an abstract level, it seems that pro and con arguments and their relation can be represented more easily in so-called bipolar argumentation frameworks [79, 81, 82, 83] containing besides attack also a support relation among arguments. This suggests a notion of bipolarity, i.e. the existence of two independent kinds of information which have a diametrically opposed nature and which represent repellent forces. When an argument is aimed at establishing the truth, empirical evidence can be used to support alleged facts. For instance, a witness's testimony can provide evidence for the claim that the suspect was at the scene of a crime, a clinical test can provide evidence against a medical diagnosis, and the outcome of a laboratory experiment can be evidence confirming (or falsifying) a psychological phenomenon [186]. The research subquestions are *Which of the principles proposed in the literature [35, 314] do not hold for such agent semantics? What new principles can we define to distinguish the varieties of agent semantics?* All these questions are answered in Chapter 4.

Question 4 How to distribute argumentation framework? Formal argumentation is a crucial aspect of decision-making processes in various domains, including law, ethics, and AI. When multiple agents are involved, it's imperative to distribute the argumentation process to sidestep the constraints of centralized systems. By distributing formal argumentation reasoning, we can overcome the limitations of centralized systems. Blockchain technology provides a decentralized and transparent platform that ensures the integrity and immutability of data. By integrating formal argumentation with blockchain, we can enhance the trustworthiness of the reasoning process. The transparent nature of blockchain enables stakeholders to audit and verify the argumentation reasoning process, promoting accountability and mitigating concerns of bias or manipulation. This question will be answered in Chapter 5, which involves an experimental platform utilizing blockchain technology. The research discusses architectures that enable knowledge representation, and reasoning within a distributed system.

1.4 Methodology

A principle-based approach This thesis generally adopts a principle-based approach, which is developed to analyze formal systems. This methodology used in formal argumentation could be seen as analogous to other fields. For example, in game theory that appeals to multiple equilibrium states in their characterization of rationality [222]. Social choice deals with the question of how to aggregate individual preferences into a collective choice. When going from a single agent to multiple agents, it is challenging to find a single outcome that satisfies everyone. The properties then play a crucial role in social choice theory as they provide a framework for evaluating and comparing different voting rules and mechanisms [23]. These properties, also known as criteria or axioms, define the desirable attributes that an ideal collective decision-making process should possess. In belief revision, there are the standard AGM properties that an operator that performs revision should satisfy in order for that operator to be considered rational [5]. In 1995 when Dung proposed the graph-based argumentation frameworks, and also the various semantics functions mapping from the graphs to sets of acceptable arguments, many other semantics have been proposed [33]. This gives a reason to use a principle-based approach for giving an overview and classifying all of these approaches. Principles conceptualize the behavior of a system at a higher level of abstraction. Furthermore, in the absence of a standard approach, principles can serve as guidelines for choosing appropriate definitions and semantics depending on various requirements. They can also guide the search for new argumentation semantics. Since argumentation is growing with the inspirations by logic, philosophy, or insights from law and other fields, such methodology is needed for developing sets of criteria that argumentation satisfies in order to qualify as rational or reasonable [317]. As argued in [135]: the study of representation and (im)possibility results for abstract argumentation must be extended for a principle-based approach for extended argumentation such as bipolar frameworks, preference-based frameworks, abstract dialectical frameworks, weighted frameworks, and input/output frameworks. In formal argumentation, principles are often more technical. Agent argumentation and bipolar argumentation typically introduce various aspects such as coalitions, knowledge, uncertainty, support, and so on. In line with common practice in the principle-based approach, this thesis uses a minimal extension of Dung as a common core to these approaches. We only introduce an abstract set of agents for agent argumentation, positive support relation for bipolar argumentation, and nothing else. The aim of this thesis is not to direct the reasoning agent to a unique set of appropriate decisions but to characterize different, possibly conflicting conclusion sets as rational outcomes within a particular formalism. Concerning the application of formal argumentation, such as in the field of law, the principle-based approach can be beneficial for the legal context, offering a more abstract conceptualization and analysis of the fundamental concept of formal argumentation.

Methodology of Architecture In this thesis, we propose IHiBO, an innovative DLT-based architecture designed specifically for distributed argumentation. The objective of our research is to address the challenges faced in decentralized argumentation systems and provide a robust framework that facilitates efficient and transparent argumentation reasoning that involves multi-agents. The IHiBO architecture incorporates the properties of DLT to enable decentralized decision-making processes and promote trust and consensus among participants. By leveraging DLT, our architecture ensures the immutability, transparency, and security of argumentation data. This allows participants to engage in argumentative exchanges with confidence, knowing that the integrity of the information is preserved. One of the key features of the IHiBO architecture is its ability to facilitate the seamless integration of distributed argumentation protocols. By leveraging smart contracts and decentralized consensus mechanisms, the architecture enables participants to engage in argumentation across different platforms and protocols, promoting interoperability and enhancing the overall efficiency of the system. In addition to its contributions to distributed argumentation, the IHiBO architecture also plays a significant role. By harnessing the capabilities of DLT used in AI. IHiBO introduces new possibilities for bringing formal argumentation frameworks to practical applications and real-world scenarios.

1.5 Relevance, Scope, and Interdisciplinary Aspects

There are two crucial suggestions stemming from the Dagstuhl workshop [135], the first is to assess the existing argumentation formalisms: ”Specifically, there needs to be more exploration on how to implement the formalisms investigated in formal computational argumentation in practical reasoning contexts, such as legal and ethical reasoning. This exploration may give rise to a more structured research area known as argumentation analysis. Within the realm of formal argumentation, it is necessary to develop procedures, approaches, and utilities for identifying, precisely representing, and formally evaluating key aspects of argumentation, ranging from basic to extremely complex situations.” The second is the recommendation for people working in AI to recognize the interdisciplinary nature of their theories and to establish connections with concepts and terminology from other fields. For instance, formal argumentation and computational social choice intersect in several intriguing research inquiries. One such question is examining the interconnection between voting and argumentation semantics, as well as how various democratic structures relate to argumentation semantics. Our exploration of agent argumentation semantics and bipolar argumentation semantics has led us to ponder the possibility of incorporating the notion of voting in social choice frameworks. Trying to follow the two crucial recommendations, this thesis studies different aspects of a formalism of argumentation, namely argument strength in structured argumentation, abstract multi-agent argumentation, and bipolar argumentation. We analyze these foundational formalisms of distributed argumentation technology using principles, some of them are introduced from other fields. We also explore the application aspects of distributed argumentation technology with the proposal of IHiBO.

Interdisciplinary aspects This thesis integrates knowledge from three disciplines: Law (focusing on AI and Law), Computer Science (specifically Knowledge Representation and Reasoning), and Technology (adopting Distributed Ledger Technology). The case studies on divorce action (in Chapter 4) apply KR to the law. The principles of formal argumentation are inspired by other fields, e.g. game theory, social choice, etc. IHiBO showcases a methodological approach from the technology domain. We focus on the current status of KR compared with other AI components (perception, machine learning, computer-human interaction, and social impacts), which is not much applied. IHiBO introduces new possibilities for bringing formal argumentation frameworks to practical applications and real-world scenarios. IHiBO as a mechanism for distributed argumentation via blockchain is an innovative solution that could potentially transform how we approach trust, transparency, and explainability in AI decision-making processes, specifically in the legal and financial domains.

1.6 Contributions

The contributions of this thesis are categorized into three distinct, yet interconnected, parts.

Formal argumentation First, our contribution is the study of weakest link. We propose a new definition of weakest link attack relation assignment based on lookahead, and compare this new lookahead definition with two existing ones in the literature using a principle-based analysis. The key result of our paper is that the lookhead definition does not satisfy the principle of context independence [112]. We therefore introduce a new principle called weak context independence, and show that it does satisfy weak context independence. We also show that lookahead weakest link is the closest approximation to Brewka's prioritised default logic PDL, also known as the greedy approach. For PDL, we prove an impossibility result under Dung's axioms. Our results generalize earlier findings restricted to total orders to the more general case of modular orders. Second, we extend the concept of abstract agent argumentation by Ryuta et al. [17]. we study four types of semantics for them. First, agent defense semantics replaces Dung's notion of defense with agent defense. Second, social agent semantics prefers arguments that belong to more agents. Third, agent reduction semantics considers the perspective of individual agents. Fourth, agent filtering semantics are inspired by a lack of knowledge. We study five existing principles and we introduce twelve new ones. In total, we provide a full analysis of fifty-two agent semantics and the seventeen principles. Specifically, we focus on the multi-agent argumentation system and the extension by Pere et al. [195], which involves a set of agents with their own normative knowledge base, combined in four different ways. Our work is concerned with the second step, which combines the argumentation frameworks into a common one, where we provide several variants of argument extensions stable sets for this kind of framework and a principle-based analysis of these variants. Third, we conduct a principle-based analysis of bipolar argumentation. Bipolar argumentation semantics is limited in the literature to reduction-based approaches. We introduce three types of defence-based semantics by adapting the notions of defence, and two types of selection-based semantics. We provide a full analysis of twenty-eight bipolar argumentation semantics and ten principles in total.

To legal reasoning This thesis contributes to the discussion on the formalization of legal interpretation in the following way. The role of interpretation is crucial in law, but it is also a source of criticism for using logic-based methods in modeling legal reasoning. For example, Leith warns that the knowledge engineer's interpretation when formalizing is necessarily premature, as the authority of interpretation of the law is assigned to the judiciary [262]. Addressing this criticism, the literature on legal interpretation has discussed the possibility that legal knowledge-based systems contain alternative syntactic formalizations. Prakken observes that while on the syntactic level formalization commits us to a given interpretation, on the conceptual level, classification of factual situations as legal concepts is not an issue of the logical form [262, p.14]. Alternatively, we can restrict the investigation by saying that "the only aspects of legal reasoning which can be formalized are those aspects which concern the following problem: *given* a particular interpretation of a body of legal knowledge, and *given* a particular description of some legal problem, what are then the general rational patterns of reasoning with which a solution to the problem can be obtained?" [262, p.4]. We analyze a formal framework itself offering different interpretations, though, then using it is directly exploitable to the comparison of the different possibilities and routes of reasoning given each interpretation.

To decision-making application Moreover, the use of DLT in the IHiBO architecture opens up opportunities for integrating Knowledge Representation (formal argumentation) with other techniques and technologies. For instance, by utilizing smart contracts and decentralized consensus mechanisms, IHiBO can enable automated reasoning and decision-making based on formal argumentation. This integration paves the way for the development of intelligent systems that can process and evaluate complex reasoning, leading to more informed and robust decision-making processes, this forms new-generation AI used in practice. Additionally, the practical implementation of formal argumentation through the IHiBO architecture provides a means to capture, represent, and exchange knowledge in a distributed manner. By combining the formal argumentation with the immutability and traceability features of blockchain technology, IHiBO facilitates a distributed argumentation reasoning process that can be accessed and verified by multiple participants. This contributes to the development of collective intelligence systems, where diverse perspectives and expertise can be effectively integrated into decision-making processes.

1.7 Layout of This Thesis

The layout of this paper is as follows. In Chapter 2, we introduce a new definition of weakest link attack relation assignment based on lookahead, and compare this new lookahead definition with two existing ones in the literature using a principle-based analysis. In Chapter 3, we introduce abstract agent argumentation frameworks. We study four types of semantics for them. First, agent defense semantics replaces Dung's notion of defense by some kind of agent defense. Second, social agent semantics prefers arguments that belong to more agents. Third, agent reduction semantics considers the perspective of individual agents. Fourth, agent filtering semantics are inspired by a lack of knowledge. We study five existing principles and we introduce twelve new ones. In total, we provide a full analysis of fifty-two agent semantics and the seventeen principles. In Chapter 4, we introduce and study seven types of semantics for bipolar argumentation frameworks, each extending Dung's interpretation of attack with a distinct interpretation of support. First, we introduce three types of defence-based semantics by adapting the notions of defence. Second, we examine two types of selection-based semantics that select extensions by counting the number of supports. Third, we analyze two types of traditional reduction-based semantics under deductive and necessary interpretations of support. We provide a full analysis of twenty-eight bipolar argumentation semantics and ten principles. We discuss a divorce case with different ways of calculating the semantics based on different interpretations of legal rules. In Chapter 5, we propose IHiBO, as a means for distributed argumentation technology, and we give a comprehensive discussion of the legal and trust considerations of the use of IHiBO in the fund management scenario. Chapter 6 follows Chapter 5, offers a thorough overview of DLT-based securities markets, and provides an exhaustive legal analysis. Outside the related works embedded in individual chapters, Chapter 7 offers a broader perspective on the related work in formal argumentation. In Chapter 8, we present an outlook for distributed argumentation technology. Finally, Chapter 9 concludes this thesis.

Chapter 2

A Principle-based Analysis to Weakest link

In this chapter, we study how to construct arguments and their relations when looking into their structure, i.e., by comparing the strength of rules applied in argument constructions. In particular, we introduce a new definition of the weakest link attack relation assignment and compare this new definition with two existing ones in the literature using a principle-based analysis. We adopt the formal framework for such attack relation assignments introduced by Dung and co-authors in a series of papers over the past decade. We show that our new definition does not satisfy context independence, we introduce a new principle called weak context independence, and we show that our new definition satisfies weak context independence. We also compare the three attack relation assignments with Brewka's prioritized default logic, also known as the greedy approach. Our results generalize earlier findings restricted to total orders, to the much more general case of modular orders. The most general case of partial orders studied in ASPIC+ is left to further research.

2.1 Introduction

The saga of weakest link is one of the great stories of defeasible argumentation. The idea that a chain of reasoning is as strong as its weakest link was used by John Pollock in 1995 as a way to compare the strength of arguments [256]. In Pollock's words: the strength of each conclusion is the minimum of the strengths of the inference with which it was derived and of the premises or intermediate conclusions from which it was derived [256, p. 99]. Pollock wrote a series of influential articles on defeasible reasoning that laid the foundations of formal argumentation [254, 255, 256, 257, 258]. Also in 1995, Dung published a seminal paper on abstract argumentation that became as well part of the foundations of formal argumentation [110]. It has been used as a general framework for instantiating (prioritised) default logic [110, 334] and defeasible logic [153], among other non-monotonic systems. These logics can be formalised in structured argumentation (e.g. ASPIC+) to generate abstract argumentation frameworks.[1] In ASPIC+, the attack relation is defined by a notion of argument strength based on weakest link or last link [216, 219].

[1] Structured argumentation builds arguments from the rules and facts of a knowledge base. Abstract argumentation just assumes an attack relation to define sets of arguments that are collectively acceptable, while ignoring the underlying logic that defines attacks as logical conflicts.

Whether one agrees or not with Modgil and Prakken that *weakest link* is appropriate for epistemic scenarios while *last link* suits better normative scenarios [219], this choice has an impact on queries to knowledge bases and normative systems: *Do fitness-loving Scots like whisky? Should snoring professors get access to the library?* [216, 219]:

The fitness-lover Scot	Snoring professor at library
$\left\{\begin{array}{c} bornInScotland \Rightarrow scottish \\ scottish \Rightarrow likesWhisky \\ fitnessLover \Rightarrow \neg likesWhisky \end{array}\right\}$	$\left\{\begin{array}{c} snores \Rightarrow misbehaves \\ misbehaves \Rightarrow accessDenied \\ professor \Rightarrow \neg accessDenied \end{array}\right\}$

Pollock's work and the distinction between weakest and last link in particular played a central role in formal models of structured argumentation. This important distinction between weakest and last link necessitates in fact the possibility of representing default rules —compare e.g. with Assumption-Based Argumentation (ABA) [58] or classical logic-based argumentation [50]. Principle-based analyses [69, 111, 112, 115, 150] have recently studied general properties of attack relations under various approaches to structured argumentation. However, given the long history of weakest link, it may come as a surprise that there have been few developments characterising how it can be used to instantiate abstract argumentation frameworks that capture a given logic. Starting with traditional weakest link [216, 219, 256], and the variant called disjoint weakest link [334], we explain this saga and its relation to prioritised default logic (PDL) [62] using three benchmark examples and study the following research question: how to axiomatize the attack relations that correspond to each variant of weakest link?

We use the formal framework for attack relation assignments introduced by Dung and Thang [111, 112, 113, 115, 116]. Their principle-based analyses of last link pointed out how weakest link must differ from last link at the level of axioms. In this chapter, we propose a new lookahead weakest link attack and compare it with existing definitions also using a principle-based analysis. An important result of this chapter is that the lookahead definition does not satisfy the principle of context independence [112]. We therefore introduce a new principle called weak context independence, and show that it does satisfy weak context independence. Another key result is an impossibility theorem for Dung's axioms [112] in the context of prioritised default logic [62].

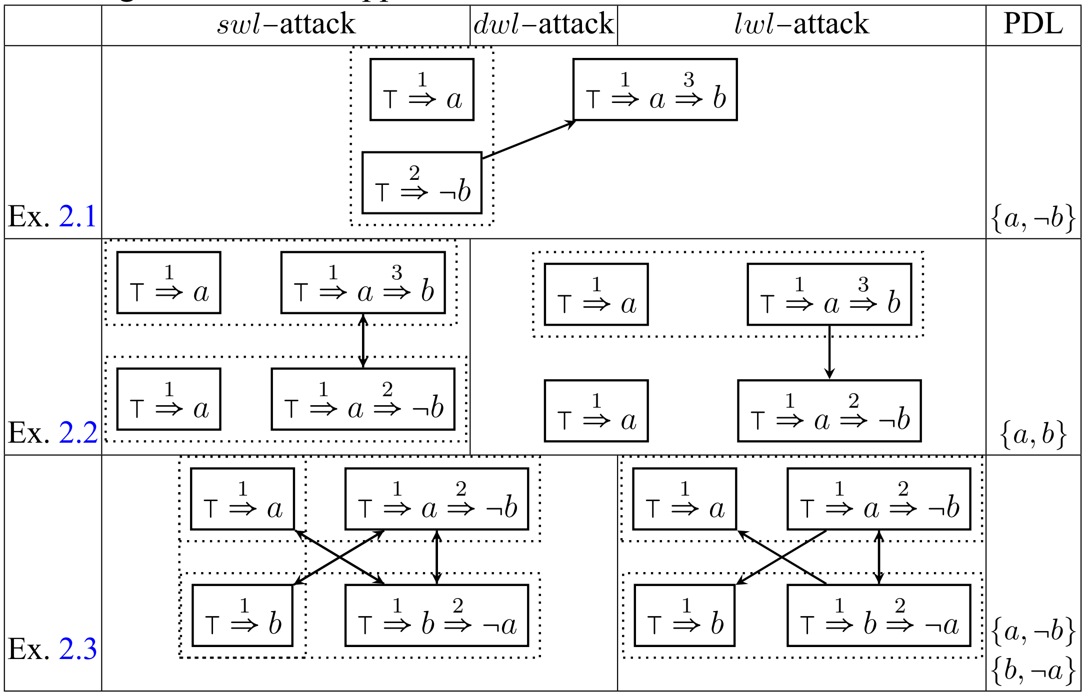
Structure of the chapter. Section 2.2 informally presents three key historical examples illustrating how to reason on weakest link. Section 2.3 gives the preliminary formal settings and our new attack relation. Section 2.4 offers the principle-based analyses. Section 2.5 shows that no attack relation assignment that captures PDL [62] can satisfy context independence. Section 2.6 discusses related work and we conclude with Section 2.7.

2.2 Three Benchmark Examples on Weakest Link

The history of weakest link evolves around three key examples which are visualised in Figure 2.1 and described as Examples 2.1–2.3. Note that the examples illustrate the role of formal argumentation in the context of PDL. All formal definitions are introduced later in Section 2.3. Here, we discuss Examples 2.1–2.3. informally.

Given a knowledge base with prioritised defaults $a \stackrel{n}{\Rightarrow} b$ and facts (including $\top$). A prioritised default $a \stackrel{n}{\Rightarrow} b$ reads as: *if* a *then normally* b. A higher number n means a higher

Figure 2.1: Approximating PDL in structured argumentation: a comparison of three attacks (columns) for three examples (rows). Columns are not marked when adjacent notions of attack agree on the induced attack relation at a given row. Dotted rectangles are argument extensions. Rightmost attacks approximate PDL better.



priority for the default rule $a \Rightarrow b$. These numerical priorities correspond to a preference relation among defaults defined by a modular order. A prioritised logic selects sets of defaults and extracts their conclusions into the so-called extensions of the logic —see Figure 2(1). A PDL extension, for example, obtains from selecting a consistent set of strongest applicable defaults. But what does a stronger *priority* mean for a default? Under the prescriptive reading, it means priority in the order of application: PDL iteratively adds the strongest applicable consistent default (Definition 2.18). Under the descriptive reading, the priority of a default is its contribution to the overall status of any extension containing this default [105]. The two readings clash in the most discussed example in defeasible reasoning with prioritised rules.

Example 2.1 (Weakest *vs* last link)**.** *Consider the three defaults:* $\top \overset{1}{\Rightarrow} a$*,* $a \overset{3}{\Rightarrow} b$*,* $\top \overset{2}{\Rightarrow} \neg b$*.*

(Prescriptive.**)** *One must select* $\{\top \overset{2}{\Rightarrow} \neg b, \top \overset{1}{\Rightarrow} a\}$ *based on application order, as shown in Figure 1. (The first choice for* $\top \Rightarrow \neg b$ *precludes* $a \Rightarrow b$ *from being selected.) This results in the extension* $\{a, \neg b\}$*, which is also a PDL extension.*

(Descriptive.**)** *This reading favours the set* $\{\top \overset{1}{\Rightarrow} a, a \overset{3}{\Rightarrow} b\}$ *as its priorities are globally better, i.e.* $\{1,3\}$ vs. $\{1,2\}$*. This gives the extension* $\{a, b\}$*, not shown in Fig. 1.*[2]

Argumentation serves as a tool for representing these two interpretations of prioritised default logic using an indirect path to conclusions, as shown in Figure 2.2(2–4). Argumentation systems add the structure that turns collections of rules into arguments [114, 219].

[2] Example 1, without priorities, represents the well-known Tweety scenario: *penguin* ⇒ *bird*, *bird* ⇒ *flies*, *penguin* ⇒ ¬*flies*. One can adduce reasons of specificity (of *penguin* over *bird*) for the standard solution: *birds fly* is overruled by the more specific rule *penguins do not fly*.

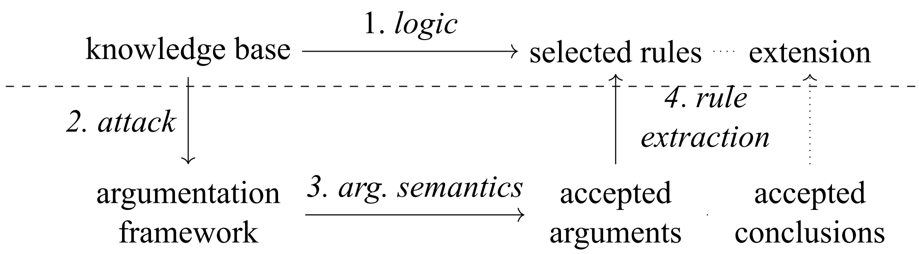


Figure 2.2: Two approaches to non-monotonic inference: (1) logic systems; (2)–(4) argumentation systems. With appropriate choices on the elements (2)–(3) one can obtain exactly the same conclusions as a given logic (1).

An attack relation (2) among arguments, together with a semantics (3), determines the acceptance status of arguments (and their conclusions). To capture a logic, the sets of accepted arguments must correspond to the sets of defaults selected by this logic (4). **Attack relations** have thus become a major subject of study in logic-based argumentation. The direction of an attack between two conflicting arguments is often determined by their relative strengths.

Example 2.1 (cont'd)**.** *Suppose we want the arrow in Figure 1 (top) giving the extension* $\{a, \neg b\}$ *corresponding to the prescriptive reading of Example 1. This attack relation is induced by* simple weakest link *(*swl*): the strength of an argument is the lesser priority of its defaults. Under the attack relation induced by* swl*, called* att_{swl}*,* $\top \Rightarrow \neg b$ *attacks* $\top \Rightarrow a \Rightarrow b$ *since* $2 > 1$*. In fact, the three notions of weakest link considered in Figure 1 agree upon this attack relation for Example 2.1.*

For the descriptive reading, the extension $\{a, b\}$ *obtains if the attack relation is induced by* last link*, i.e. if the strength of an argument is the priority of its last default. Under last link,* $\top \Rightarrow a \Rightarrow b$ *attacks* $\top \Rightarrow \neg b$ *since* $3 > 2$*.*[3]

For Examples 2.2–2.3, the three variants of weakest link swl, dwl and lwl no longer agree on the attacks or argument extensions. For each variant, Figure 2.1 depicts its attacks and extensions in the argumentation framework that falls under its column.

Example 2.2 (Simple vs. Disjoint weakest link)**.** *Let* $\top \overset{1}{\Rightarrow} a$*,* $a \overset{3}{\Rightarrow} b$*,* $a \overset{2}{\Rightarrow} \neg b$ *define our knowledge base. Note that the two arguments* $\top \Rightarrow a \Rightarrow b$ *and* $\top \Rightarrow a \Rightarrow \neg b$ *share a default with minimum priority* $\top \Rightarrow a$*. See the mid row in Figure 2.1.*

(Simple weakest link.) *Pollock's definition assigns the same strength* 1 *to these two arguments. This strength gives the mutual* swl*-attack in Figure 2.1 (mid, left).*

(Disjoint weakest link.) *A more intuitive attack relation ignores all defaults shared by two arguments in order to exploit a potential asymmetry in the remaining defaults' strengths. A relational measure of strength for such an attack is disjoint weakest link* dwl *[334].* dwl *assigns strengths* $3 > 2$ *to these arguments, and generates the* dwl*-attack in Figure 2.1 (mid, right) that breaks the symmetry of* swl*-attacks.*

[3] These priorities give the same outputs for the *fitness-loving Scot* and *snoring professor* [216, 219], which are just variants of Example 1 with facts. Other variants of Example 1 with facts and strict rules [63, 65] give the (non-)teaching dean professor scenario [112], see Example 2.5 below. For further variants of Example 1 defined by partial orders we refer to Dung's paper in 2018 [115]. A brief discussion for the case of partial orders can be found in Section 2.7.

Pollock's definition of weakest link swl [257] was adopted and studied for ASPIC+ by Modgil and Prakken [216, 219]. Young et al. [334, 336] introduced dwl and proved that argument extensions under the dwl-attack relation correspond to PDL extensions under total orders; see also the results by Liao et al. [194] or Pardo and Straßer [242]. For knowledge bases with modular orders, a new attack relation is needed for more intuitive outputs and also for a better approximation of PDL —that is, better than dwl.

Example 2.3 (Beyond dwl)**.** *Let $\top \overset{1}{\Rightarrow} a$, $\top \overset{1}{\Rightarrow} b$, $a \overset{2}{\Rightarrow} \neg b$, and $b \overset{2}{\Rightarrow} \neg a$ be the defaults.*

(swl, dwl**)** *The induced attacks admit $\{\top \Rightarrow a, \top \Rightarrow b\}$ as one of the argument extensions in Figure 1 (bottom, left). This fits neither the prescriptive interpretation nor PDL: as these two defaults are the weakest, selecting either of them ought to be followed by a stronger default, namely $a \Rightarrow \neg b$ and resp. $b \Rightarrow \neg a$. In other words, swl and dwl can select applicable defaults concurrently, leading to sub-optimal outputs.*

(lwl**)** *A sequential selection of defaults, more in line with PDL, is enforced by the attack relation in Figure 1 (bottom, right), induced by* lookahead weakest link *(lwl).*

The new attack we propose (lwl) decides an attack from an argument by looking ahead to any superargument and its attacks: if both coincide at attacking a third argument, the former attack is disabled and only that of the superargument remains. For Example 3, this is how in Figure 1(bottom) lwl prevents the undesired swl- and dwl-based extension $\{a, b\}$.

2.3 Attack Assignments Based on Weakest Link

Preliminaries. This chapter uses basic setting similar to that of Dung [112]. We assume a non-empty set $\mathscr{L}$ of ground atoms and their classical negations. An atom is also called a positive literal while a negative literal is the negation of a positive literal. A set of literals is said to be **contradictory** if it contains a pair $a, \neg a$, i.e. an atom a and its negation $\neg a$.

Definition 2.1 (Rule)**.** *A defeasible rule is of the form $b_1, \ldots, b_n \Rightarrow h$ where $b_1, \ldots, b_n, h$ are domain literals. A strict rule is of the form $b_1, \ldots, b_n \to h$ where h is now either a domain literal or a non-domain atom ab_d for some defeasible rule d.*

We also define the body and head of rule r as $bd(r) = \{b_1, \ldots, b_n\}$ and $hd(r) = h$.

Instead of just assuming transitivity for the preference order among defeasible rules, as in Dung's work, in this chapter, we use modular orders $\leq$ and their equivalent ranking functions $rank$. In fact, we will use the two notions indistinctly throughout the chapter.

Definition 2.2 (Rule-based system)**.** *A rule-based system is defined as a triple $RBS = (RS, RD, rank)$, where RS is a set of strict rules, RD is a finite set of defeasible rules, and $rank$ is a function $RD \to \mathbb{N}$ that assigns a priority $n = rank(d)$ to each rule $d \in RD$.*

A ranking $rank : RD \to \mathbb{N}$ corresponds to a modular preorder $\leq \subseteq RD \times RD$, i.e. a reflexive, transitive relation satisfying: $rank(d) \leq rank(d')$ iff $d \leq d'$. A **base of evidence** BE is a (consistent) set of ground domain literals containing $\top$ and representing unchallenged facts.

Remark 2.1. *Given the scope of our discussion and examples, our framework is less expressive than that of Dung [112]. We assume an empty set $RS = \varnothing$ of strict rules, and keep the set RS in Definition 2.2 only for notational coherence with the literature.*[4]

Definition 2.3 (Knowledge base)**.** *A knowledge base is a pair $K = (RBS, BE)$ containing a rule-based system $RBS = (RS, RD, rank)$ and a base of evidence $BE \subseteq \mathscr{L} \cup \{\neg a : a \in \mathscr{L}\}$. For convenience, we often write $K = (RS, RD, rank, BE)$ instead of $K = (RBS, BE)$.*

Example 2.4. *The knowledge base $K = (RS, RD, rank, BE)$ for Example 3 is defined by: $RS = \varnothing$; $RD = \{d_1 : \top \Rightarrow a,\ d_2 : \top \Rightarrow b,\ d_3 : a \Rightarrow \neg b,\ d_4 : b \Rightarrow \neg a\}$, the function $rank$ mapping $\{d_1, d_2\} \mapsto 1$ and $\{d_3, d_4\} \mapsto 2$, and finally $BE = \{\top\}$. Equivalently, we can write $K = (RS, RD, \preceq, BE)$ with $\preceq = \{d_1, d_2\}^2 \cup \{d_3, d_4\}^2 \cup (\{d_1, d_2\} \times \{d_3, d_4\})$.*

Example 2.5 (Dean scenario)**.** *For an example with strict rules, the dean scenario asks whether the dean teaches. The knowledge base $K = (RS, RD, rank, BE)$ is given by:*

$$\begin{aligned} RS &= \{dean \rightarrow administrator\} \\ RD &= \{dean \overset{1}{\Rightarrow} professor,\ professor \overset{3}{\Rightarrow} teach,\ administrator \overset{2}{\Rightarrow} \neg teach\} \\ BE &= \{dean\}. \end{aligned}$$

Definition 2.4 (Argument)**.** *Given a knowledge base $K = (RS, RD, rank, BE)$, an **argument** wrt K is defined inductively as follows:*

1. *For each $\alpha \in BE$, $[\alpha]$ is an argument with conclusion α.*
2. *Let r be a rule of the form $\alpha_1, \ldots, \alpha_n \rightarrow / \Rightarrow \alpha$ (with $n \geq 0$) from K. Further suppose that $A_1, \ldots, A_n$ are arguments with conclusions $\alpha_1, \ldots, \alpha_n$ respectively. Then $A = [A_1, \ldots, A_n \rightarrow / \Rightarrow \alpha]$, also denoted $A = [A_1, \ldots, A_n, r]$, is an argument with conclusion $cnl(A) = \alpha$ and last rule $last(A) = r$.*
3. *Each argument wrt K is obtained by finitely many applications of the steps 1–2.*

Example 2.6. *The arguments wrt the knowledge base K from Example 2.4 are $A_0 = [\top]$ plus:*

$$A_1 = [[\top] \Rightarrow a] \quad A_2 = [[\top] \Rightarrow b] \quad A_3 = [[[\top] \Rightarrow a] \Rightarrow \neg b] \quad A_4 = [[[\top] \Rightarrow b] \Rightarrow \neg a].$$

Definition 2.5 (Argumentation framework)**.** *The set of all arguments induced by a knowledge base K is denoted by AR_K. An **argumentation framework** (AF) induced by K is a pair $AF = (AR_K, att(K))$ where $att(K) \subseteq AR_K \times AR_K$ is called an attack relation.*

Definition 2.6. *A knowledge base K is **consistent** if the closure of BE under RS is not a contradictory set. The set of **conclusions** of arguments in $\mathscr{E} \subseteq AR_K$ is denoted by $cnl(\mathscr{E})$.*

*A **strict** argument is an argument containing no defeasible rule. An argument is **defeasible** iff it is not strict. **The set of defeasible rules** appearing in an argument A is denoted by $dr(A)$.*

[4] As a consequence, the atoms in $\mathscr{L}$ here only consist of *domain atoms* representing propositions about the concerned domains. Dung also considers *non-domain atoms* ab_d for the non-applicability of a defeasible rule d, and undercuts as strict rules $b_1, \ldots, b_n \rightarrow ab_d$ that act against the applicability of a defeasible rule d in RD [112]. We leave for future work the extension of our current results to knowledge bases with strict rules and undercutting arguments.

*An argument B is a **subargument** of an argument A, denoted as $B \in sub(A)$ or $B \sqsubseteq A$, iff $B = A$ or $A = [A_1, \ldots, A_n, r]$ and B is a subargument of some A_i. B is a **superargument** of A, denoted as $B \in super(A)$ or $B \sqsupseteq A$, iff $A \in sub(B)$.*

Definition 2.7 (Sensible class)**.** *A class $\mathscr{K}$ of knowledge bases is **sensible** iff $\mathscr{K}$ is a non-empty class of consistent knowledge bases K, and for any knowledge base $K = (RBS, BE)$ in $\mathscr{K}$, all consistent knowledge bases of the form (RBS, BE') also belong to $\mathscr{K}$.*

Definition 2.8 (Attack relation assignment)**.** *Given a sensible class of knowledge bases $\mathscr{K}$, an **attack relation assignment** is a function att mapping each $K \in \mathscr{K}$ to an attack relation $\mathrm{att}(K) \subseteq AR_K \times AR_K$.*

Definition 2.9 (Stable semantics)**.** *Given an argumentation framework $(AR_K, \mathrm{att}(K))$, we say that $\mathscr{E} \subseteq AR_K$ is a **stable extension** if: (1) $\mathscr{E}$ is conflict-free $\mathrm{att}(K) \cap (\mathscr{E} \times \mathscr{E}) = \varnothing$, and (2) $\mathscr{E}$ attacks all the arguments in $AR_K \smallsetminus \mathscr{E}$. This is also denoted $\mathscr{E} \in stb(AR_K, \mathrm{att}(K))$.*

While many other semantics exist, we follow Dung [112] and study attack relations mostly under stable semantics. Only Principle 2.5 mentions the complete semantics. Recall that a set $\mathscr{E} \subseteq AR_K$ **defends** an argument A iff $\mathscr{E}$ attacks all attackers of A. A **complete** extension $\mathscr{E}$ is defined by: $\mathscr{E}$ is conflict free (no attack occurs within $\mathscr{E}$) and $A \in \mathscr{E}$ iff $\mathscr{E}$ defends A. Our main result does not depend on the choice for the stable semantics: for Examples 2.1–2.3 and the proof of Theorem 2.1, one can indistinctly use the complete semantics or the preferred semantics (i.e. $\subseteq$-maximally complete extensions).

Definition 2.10 (Belief set)**.** *A set $S \subseteq \mathscr{L}$ is said to be a stable belief set of knowledge base K wrt an attack relation assignment att iff $\mathrm{att}(K)$ is defined and there is a stable extension $\mathscr{E}$ of $(AR_K, \mathrm{att}(K))$ such that $S = cnl(\mathscr{E})$.*

Attacks based on weakest link. We now present three attack relation assignments based on weakest link. All our attacks are rebuts, i.e. they contradict (sub-)conclusions. (Recall that we have neither non-domain literals nor defeasible premises that would define undercutting and resp. undermining attacks.)

Definition 2.11 (Contradicting attack)**.** *Let $A, B \in AR_K$ for a knowledge base K. A **contradicts** B (at B') iff $B' \in sub(B)$ and the conclusions of A and B' are contradictory.*

Definition 2.12 (Weakest link)**.** *The **weakest link** of a set of rules R, denoted as $wl(R)$, is the rank of the lowest rank rule in R. Formally, $wl(R) = \min_{r \in R} rank(r)$. Abusively, we also use $wl(A)$ for arguments A, simply defined by $wl(dr(A))$.*

Weakest link thus provides an absolute measure wl of strength for arguments—for strict arguments A, we just define $wl(A) = \infty$. This measure defines the first attack, based on Pollock's traditional idea [257].

Definition 2.13 (Simple weakest link attack)**.** *Let $A, B \in AR_K$ for a knowledge base K. We say that A **swl-attacks** B (at B'), denoted as $(A, B) \in \mathrm{att}_{swl}(K)$ iff A contradicts B at B' and $wl(A) \nless wl(B')$ (that is, $wl(A) \geq wl(B')$ for modular orders).*

Note that a defeasible argument A can contradict a strict argument B —a fact, in the present context. In those cases, $wl(A) < wl(B)$ and so the ordering $<$ is well-defined.

The second attack was introduced by Young et al. [334] for total orders. dwl was motivated by the unintuitive outputs of swl in scenarios with shared rules, like Example 2.2.

Definition 2.14 (Disjoint weakest link attack)**.** *Let $A, B \in AR_K$ for some K. A* ***dwl-attacks*** *B (at B'), denoted $(A, B) \in att_{dwl}(K)$ iff A contradicts B at B' and $wl(dr(A) \smallsetminus dr(B')) \nless wl(dr(B') \smallsetminus dr(A))$.*

The third attack, newly introduced in this chapter, is a refinement of disjoint weakest link. It aims to better approximate the extensions of PDL, a paradigmatic implementation of the idea of weakest link. The motivation for a new attack was given in Example 3. We call it *lookahead attack* since an attack from an argument may be cancelled if a superargument of it also attacks the same target, so this new attack looks ahead to superarguments before deciding whether an attack from the subargument ultimately exists or not.

Definition 2.15 (Lookahead weakest link attack)**.** *Let $A, B \in AR_K$ for a knowledge base K. We say that $(A, B) \in att_{dwl}(K)$ is* ***maximal*** *if A is $\sqsubseteq$-maximal in AR_K with the property $(\cdot, B) \in att_{dwl}$. We also define: A* ***lwl-attacks*** *B at B', denoted as $(A, B) \in att_{lwl}(K)$, iff A dwl-attacks B at B' and*

1. *either $(B', A) \notin att_{dwl}(K)$*

2. *or, in case $(B', A) \in att_{dwl}(K)$, if (A, B) is not maximal then neither is (B', A).*

Informally, att_{lwl} obtains from att_{dwl} by removing, in each bidirectional attack, the attacker that is not $\sqsubseteq$-maximal, in case the other attacker is. With more detail, one must (1) compute $att_{dwl}(K)$; (2) for each $(A, B'), (B', A) \in att_{dwl}(K)$, if (A, B') is not maximal while (B', A) is, then remove as attacks all pairs (A, B) with $B \sqsupseteq B'$.[5]

Let us stress that our definition of lookahead attack lwl overrides the notion of contradicting attack (Definition 2.11). As a result, the principle of subargument structure will fail for att_{lwl}, while in general it holds for all ASPIC+ attacks in the literature.

Each of the above definitions (Defs. 2.13–2.15) of an attack relation $att(K)$ over a knowledge base K extends into an attack relation assignment att over a sensible class $\mathscr{K}$ of knowledge bases. This is simply the function $att : K \mapsto att(K)$ for each $K \in \mathscr{K}$.

2.4 Principle-based Analysis

In this section, we offer a principle-based analysis of the three attack relation assignments, using the eight principles proposed by Dung [112] plus a new principle. In the following, $\mathscr{K}$ denotes a sensible class of knowledge bases, and att an attack relation assignment defined for $\mathscr{K}$. Some of the following results for Principles P 2.1–P 2.9 were partly proved by Dung [113]. With detail, our results on swl are also proved in Theorem 7.10 (for P 2.1), Lemma 7.6 (for P 2.2, P 2.6–P 2.8) and Theorem 7.8 (for P 2.4).

Credulous cumulativity states that turning accepted conclusions Ω of a knowledge base K into facts preserves stable extensions and consistency. This operation is denoted as an expansion of K into $K + \Omega = (RBS, BE \cup \Omega)$.

Principle 2.1 (Credulous cumulativity)**.** *We say that att satisfies* ***credulous cumulativity*** *for $\mathscr{K}$ iff for each $K \in \mathscr{K}$ and each stable belief set S of K, any finite subset $\Omega \subseteq S$ satisfies:*

[5] A reader might wonder why Definition 2.15 does not simply state: $(A, B) \in att_{lwl}(K)$ iff $(A, B) \in att_{dwl}(K)$ and A is $\sqsubseteq$-maximal with $(\cdot, B) \in att_{dwl}(K)$. The reason is that, under these attacks, one can define some K whose stable belief sets include logically contradictory sets.

1. *$K + \Omega$ is a consistent knowledge base (i.e. $K + \Omega$ belongs to $\mathscr{K}$), and*

2. *S is a stable belief set of $K + \Omega$ wrt att.*

Proposition 2.1. *Credulous cumulativity (P1) is not satisfied by any of $att_{swl}, att_{dwl}, att_{lwl}$.*

Proof. For a counterexample, let a sensible class $\mathscr{K}$ contain the knowledge base K corresponding to Example 1. As depicted in Fig. 1(top), $S = \{a, \neg b\}$ is a stable belief set of K wrt att_{swl}, att_{dwl} and att_{lwl}. However, S is not a stable belief set of $K + \{a\}$ wrt any of these three attacks. □

Context independence states that the attack relation between two arguments depends only on the rules that appear in them and their preferences [112].

Principle 2.2 (Context independence)**.** *We say that att satisfies **context independence** for $\mathscr{K}$ iff for any two $K, K' \in \mathscr{K}$ with preference relations $\leq$ and resp. $\leq'$ and any two arguments A, B belonging to $AR_K \cap AR_{K'}$, if the restrictions of $\leq$ and $\leq'$ on $dr(A) \cup dr(B)$ coincide, then it holds that $(A, B) \in att(K)$ iff $(A, B) \in att(K')$.*

Proposition 2.2. *Context independence (P 2.2) is satisfied by att_{swl} and att_{dwl}, while it is not satisfied by att_{lwl}.*

Proof. **For att_{swl}.** Let $K, K' \in \mathscr{K}$ have preference relations $\leq$ and resp. $\leq'$. Suppose that for $A, B \in AR_K \cap AR_{K'}$, the restrictions of $\leq$ and $\leq'$ on $dr(A) \cup dr(B)$ coincide. If $(A, B) \in att_{swl}(K)$, by Def. 2.13 the conclusions of A and a subargument $B' \in AR_K$ of B are contradictory and $wl(A) \nless wl(B')$ for K. Since B' is a subargument of $B \in AR_{K'}$, $B' \in AR_{K'}$ and $dr(B) \supseteq dr(B')$. Hence the restrictions of $\leq$ and $\leq'$ on $dr(A) \cup dr(B')$ also coincide. So for K' it also holds that $wl(A) \nless wl(B')$. Hence, $(A, B) \in att_{swl}(K')$. The same reasoning applies in the other direction, and so we conclude that $(A, B) \in att_{swl}(K)$ iff $(A, B) \in att_{swl}(K')$.

For att_{dwl}. The proof is analogous to the proof for att_{swl}: Let $K, K' \in \mathscr{K}$ have preference relations $\leq$ and resp. $\leq'$. Suppose that for $A, B \in AR_K \cap AR_{K'}$, the restrictions of $\leq$ and $\leq'$ on $dr(A) \cup dr(B)$ coincide. If $(A, B) \in att_{dwl}(K)$, by Definition 2.14 the conclusions of A and a subargument $B' \in AR_K$ of B are contradictory and $wl(dr(A) \setminus dr(B')) \nless wl(dr(B') \setminus dr(A))$ for K. Since B' is a subargument of $B \in AR_{K'}$, $B' \in AR_{K'}$ and $dr(B) \supseteq dr(B')$. Hence, the restrictions of $\leq$ and $\leq'$ on $dr(A) \cup dr(B')$ also coincide. So for K' it also holds that $wl(dr(A) \setminus dr(B')) \nless wl(dr(B') \setminus dr(A))$. Hence, $(A, B) \in att_{dwl}(K')$. The same reasoning applies in the other direction, and so it holds that $(A, B) \in att_{dwl}(K)$ iff $(A, B) \in att_{dwl}(K')$.

For att_{lwl}. Let $K' = \{\top \overset{1}{\Rightarrow} a, \top \overset{1}{\Rightarrow} b, b \overset{2}{\Rightarrow} \neg a\}$ obtain from removing $a \overset{2}{\Rightarrow} \neg b$ from the knowledge base K in Example 2.3. This is a counterexample, since the arguments $[\top \Rightarrow a]$ and $[\top \Rightarrow b \Rightarrow \neg a]$ belong to $AR_K \cap AR_{K'}$, and the restrictions of $\leq$ and $\leq'$ to the set $dr([\top \Rightarrow a]) \cup dr([\top \Rightarrow b \Rightarrow \neg a])$ coincide. However, $([\top \Rightarrow a], [\top \Rightarrow b \Rightarrow \neg a]) \notin att_{lwl}(K)$ while $([\top \Rightarrow a], [\top \Rightarrow b \Rightarrow \neg a]) \in att_{lwl}(K')$. □

For a weaker version of context independence, one can state that an attack also depends on the superarguments. Let us define: $super_K(A) = \{A^+ \in AR_K : A^+ \sqsupseteq A\}$.

Principle 2.3 (Weak context independence)**.** *We say that att satisfies **weak context independence** for $\mathscr{K}$ iff for any two $K, K' \in \mathscr{K}$ with preferences $\leq$ and resp. $\leq'$ and any two arguments $A, B \in AR_K \cap AR_{K'}$:*

$$\textit{if} \begin{cases} \preceq, \preceq' \text{ agree upon } dr(A) \cup dr(B) \\ \text{and } super_K(A) = super_{K'}(A) \\ \text{and } super_K(B) = super_{K'}(B) \end{cases} \textit{then } (A,B) \in att(K) \textit{ iff } (A,B) \in att(K').$$

Proposition 2.3. *Weak context independence (P 2.3) is satisfied by the three attacks* att_{swl}, att_{dwl}, att_{lwl}.

Proof. **For** att_{swl}, att_{dwl}**.** Clearly, the set of pairs $\{K, K'\}$ in $\mathscr{K}$ that need to be tested for (P 2.3) are a subset of those pairs that to be tested for (P 2.2): the former are all pairs validating Def. 2.3(i)–(ii) while the latter also include the pairs that only validate (i). Hence, if att satisfies (P 2.2), then it also satisfies (P 2.3). From this and the above proofs for (P 2.2), we conclude that att_{swl}, att_{dwl} satisfy (P 2.3).

For att_{lwl}**.** Let $K, K' \in \mathscr{K}$ have preference relations $\preceq$ and resp. $\preceq'$. Suppose that for $A, B \in AR_K \cap AR_{K'}$, $\preceq$ and $\preceq'$ agree upon $dr(A) \cup dr(B)$ and $super_K(A) = super_{K'}(A)$ and $super_K(B) = super_{K'}(B)$. Towards a contradiction, assume that $(A,B) \in att_{lwl}(K)$ at B', but $(A,B) \notin att_{lwl}(K')$. Because $(A,B) \in att_{lwl}(K)$ at B', according to Def. 2.15, $(A,B) \in att_{dwl}(K)$ at B'. Since att_{dwl} satisfies context independence, $(A,B) \in att_{dwl}(K')$ at B'. As a result, $(\star)$ $(B',A) \in att_{dwl}(K')$, and so (A,B) is not maximal in $att_{dwl}(K')$ and (B',A) is maximal in $att_{dwl}(K')$. Because $super_K(A) = super_{K'}(A)$ and $super_K(B) = super_{K'}(B)$, by $(\star)$ and (P 2.3) we obtain $(B',A) \in att_{dwl}(K)$, and so (A,B) is not maximal in $att_{dwl}(K)$ and (B',A) is maximal in $att_{dwl}(K)$. Hence, $(A,B) \notin att_{lwl}(K)$. This is in contradiction with $(A,B) \in att_{lwl}(K)$. □

The principle of attack monotonicity (defined below)reflects the intuition that the more reliable the foundation of an argument is, the stronger the argument becomes. Suppose the defeasible information on which an argument is based is confirmed by unchallenged observations. Replacing the defeasible bits by the observed facts should result in a strengthened argument: whatever is attacked by the original argument should also be attacked by the strengthened one, and whatever attacks the strengthened one, attacks the original one.

Definition 2.16 (Strengthening operation)**.** *Let* $A \in AR_K$ *and* $\Omega \subseteq BE$ *be a finite set of domain literals. The strengthening of* A *wrt* Ω *denoted by* $A \uparrow \Omega$ *is defined inductively as follows:*

$$A \uparrow \Omega = \begin{cases} \{[\alpha]\} & \textit{if } A = [\alpha] \textit{ and } \alpha \in BE \\ AS \cup \{[hd(r)]\} & \textit{if } A = [A_1, \ldots, A_n, r] \textit{ and } hd(r) \in \Omega \\ AS & \textit{if } A = [A_1, \ldots, A_n, r] \textit{ and } hd(r) \notin \Omega \end{cases}$$

where $AS = \{[X_1, \ldots, X_n, r] \mid \forall i : X_i \in A_i \uparrow \Omega\}$

Principle 2.4 (Attack monotonicity)**.** *Let* att *be an attack relation assignment defined for a sensible class* $\mathscr{K}$ *of knowledge bases. We say* att *satisfies the property of attack monotonicity for* $\mathscr{K}$ *iff for each knowledge base* $K \in \mathscr{K}$ *and each finite subset* $\Omega \subseteq BE$*, the following assertions hold for arbitrary* $A, B \in AR_K$ *and* $X \in A \uparrow \Omega$.

1. *If* $(A,B) \in att(K)$ *then* $(X,B) \in att(K)$.

2. *If* $(B,X) \in att(K)$ *then* $(B,A) \in att(K)$.

Proposition 2.4. *Attack monotonicity is satisfied by* att_{swl} *and* att_{dwl}*. It is not satisfied by* att_{lwl}.

Proof. **For att_{swl}.** (1) Let $K \in \mathscr{K}$, $\Omega \subseteq BE$, $A, B \in AR_K$ and $X \in A \uparrow \Omega$. From $(A, B) \in att_{swl}(K)$, A contradicts B at some B' with $wl(A) \nless wl(B')$. Because $X \in A \uparrow \Omega$, X also contradicts B at B' with $dr(X) \subseteq dr(A)$, so $wl(X) \geq wl(A)$. As a result, $wl(X) \nless wl(B')$ and so $(X, B) \in att_{swl}(K)$. (2) From $(B, X) \in att_{swl}(K)$, B contradicts X at some X' with $wl(B) \nless wl(X')$. Because $X \in A \uparrow \Omega$, there is $A' \in sub(A)$ with $cnl(X') = cnl(A')$ and $dr(X') \subseteq dr(A')$, so $wl(X') \geq wl(A')$. As a result, B contradicts A at A' with $cnl(B)$ and $cnl(A')$ being contradictory and $wl(B) \nless wl(A')$. Thus, $(B, A) \in att_{swl}(K)$.

For att_{dwl}. The proofs are analogous to the att_{swl} case. (1) From $(A, B) \in att_{dwl}$ to $(X, B) \in att_{dwl}$: since $dr(X) \subseteq dr(A)$ we get $wl(dr(X) \setminus dr(B')) \geq wl(dr(A) \setminus dr(B')) \geq wl(dr(B') \setminus dr(A)) \geq wl(dr(B') \setminus dr(X))$. As a result, $wl(dr(X) \setminus dr(B')) \nless wl(dr(B') \setminus dr(X))$, and so $(X, B) \in att_{dwl}(K)$. (2) From $(B, X) \in att_{swl}(K)$, B contradicts X at some X' with $wl(dr(B) \setminus dr(X')) \nless wl(dr(X') \setminus dr(B))$. Because $X \in A \uparrow \Omega$, there is $A' \in sub(A)$ with $cnl(X') = cnl(A')$ and $dr(X') \subseteq dr(A')$, so $wl(dr(B) \setminus dr(A')) \geq wl(dr(B) \setminus dr(X')) \geq wl(dr(X') \setminus dr(B)) \geq wl(dr(A') \setminus dr(B))$. As a result, $(B, A) \in att_{dwl}(K)$.

For att_{lwl}. Let K contain $BE = \{a\}$ and a set RD rules of strength 1 that give: $A = [[\top \Rightarrow a] \Rightarrow b]$, $B = [[\top \Rightarrow \neg c] \Rightarrow \neg b]$, $X = [[a] \Rightarrow b]$ and also $A^+ = [A \Rightarrow c]$, $B^+ = [B \Rightarrow \neg a]$, $X^+ = [X \Rightarrow c]$. Since B^+ cannot attack X, $(A, B) \in att_{lwl}(K)$ is not preserved into $(X, B) \in att_{lwl}(K)$ although X is a strengthening of A with $\{a\}$.

□

The next principle, irrelevance of redundant defaults, states that adding redundant defaults into the knowledge base does not result in a change of beliefs (outputs).

Notation 2.1. *For any defeasible rule d, denote $K + d = (RS, RD \cup \{d\}, \preceq, BE)$ where $K = (RS, RD, \preceq, BE)$. For convenience, for any evidence $\omega \in BE$ we also denote the default $\Rightarrow \omega$ by d_ω.*

Principle 2.5 (Irrelevance of redundant defaults)**.** *Let $\mathscr{K}$ be a sensible class of knowledge bases such that for each $K = (RSB, BE) \in \mathscr{K}$, for each evidence $\omega \in BE$, $K + d_\omega$ belongs to $\mathscr{K}$. Further let att be an attack relation assignment defined for $\mathscr{K}$.*

We say the attack relation assignment att satisfies irrelevance of redundant defaults for $\mathscr{K}$ iff for each knowledge base $K = (RSB, BE) \in \mathscr{K}$, for each evidence $\omega \in BE$:

1. *the stable belief sets of K and $K + d_\omega$ coincide, and*
2. *the complete belief sets of K, $K + d_\omega$ coincide.*

Proposition 2.5. *Irrelevance of redundant defaults (P 2.4) is satisfied by the three attacks $att_{swl}, att_{dwl}, att_{lwl}$.*

Proof. **For att_{swl}.** First, $AR_K \subset AR_{K+d_\omega}$. Let $AR^+ = AR_{K+d_\omega} \setminus AR_K$, representing arguments that are newly added into AR_{K+d_ω} due to the addition of d_ω. For each argument $A' \in AR^+$, there exists an argument $A \in AR_K$, such that $A = A' \uparrow \{\omega\}$. Hence, $cnl(A) = cnl(A')$ and $wl(A) \nless wl(A')$. Then, for each $B \in AR_K$ such that $(B, A) \in att_{swl}(K)$, we have $(B, A), (B, A') \in att_{swl}(K + d_\omega)$. Hence, A' can not be in any stable or complete extension $\mathscr{E}$ unless $A \in \mathscr{E}$. As a result, each stable or complete extension $\mathscr{E}'$ of $K + d_\omega$ is of the form $\mathscr{E} \cup \{A' \in AR^+ : A \in \mathscr{E}\}$ where $\mathscr{E}$ is an extension of K.

For att_{dwl}. The proof is analogous and only changes in statements of the form $wl(A \setminus B) \nless wl(A' \setminus B)$. Again, $cnl(A) = cnl(A')$ for any argument $A \in \mathscr{E}$ in an extension and its weakening $A' \in AR_{K+d_\omega}$. By the definition of stable and complete extensions, in the new AF these must be of the form $\mathscr{E}' = \mathscr{E} \cup \{A' : A \in \mathscr{E}\}$.

For att_{lwl}**.** The proof is also analogous. For each argument $A' \in AR^+$, there exists an argument $A \in AR_K$, such that $A = A' \uparrow \{\omega\}$. By definition of att_{lwl}, $(A,B) \in att_{lwl}(K)$ iff $(A,B),(A,B') \in att_{lwl}(K+d_\omega)$. As a result, each stable or complete extension $\mathscr{E}'$ of $K+d_\omega$ is of the form $\mathscr{E} \cup \{A' : A \in \mathscr{E}\}$ where $\mathscr{E}$ is an extension of K.

□

The next two principles state basic properties of argumentation. Subargument structure and attack closure are two basic principles. Subargument structure states that if an argument attacks a subargument, it attacks the entire argument. Attack closure says that attacks are either based on undercuts[6] or contradicting arguments.

Principle 2.6 (Subargument structure)**.** *Let $\mathscr{K}$ be a sensible class of knowledge bases and att be an attack relation assignment defined for $\mathscr{K}$. Then att is said to satisfy the property of subargument structure for $\mathscr{K}$ iff for each $K \in \mathscr{K}$, for all $A, B \in AR_K$,*

$$(A,B) \in att(K) \textit{ iff there is a defeasible subargument } B' \textit{ of } B \textit{ such that} (A,B') \in att(K).$$

Proposition 2.6. *Subargument structure (P 2.6) is satisfied by att_{swl} and att_{dwl}, while it is not satified by att_{lwl}.*

Proof. **For** att_{swl}**.** (⇒) From $(A,B) \in att_{swl}$(K), A contradicts some $B' \sqsubseteq B$ with $wl(A) \nless wl(B')$. If B' was strict, so would be A, contradicting that K is consistent, i.e. that $K \in \mathscr{K}$. (⇐) If A contradicts a defeasible B' at B'' with $wl(A) \nless wl(B'')$, then for any $B \sqsupseteq B'$ we have $wl(B) \leq wl(B') \leq wl(B'')$ and so $(A,B) \in att_{swl}(K)$.

For att_{dwl}**.** The two directions of the proof are analogous, now using $wl(A \smallsetminus B') \nless wl(B' \smallsetminus A)$ for (⇒); and $wl(B \smallsetminus A) \leq wl(B' \smallsetminus A) \leq wl(B'' \smallsetminus A)$ for (⇐).

For att_{lwl}**.** For a counterexample to (⇐), let $\top \overset{2}{\Rightarrow} a$, $\top \overset{2}{\Rightarrow} \neg a$, $a \overset{1}{\Rightarrow} b$, $\neg a \overset{1}{\Rightarrow} \neg b$ be the rules of AR_K. Then, $[\top \Rightarrow a]$ does not lwl-attack $[\top \Rightarrow \neg a \Rightarrow \neg b]$ but lwl-attacks $[\top \Rightarrow \neg a]$; finally, note that $[\top \Rightarrow \neg a]$ is a subargument of $[\top \Rightarrow \neg a \Rightarrow \neg b]$. □

Principle 2.7 (Attack closure)**.** *Let $\mathscr{K}$ be a sensible class of knowledge bases and att be an attack relation assignment defined for $\mathscr{K}$. Then att is said to satisfy the property of attack closure for $\mathscr{K}$ iff for each $K \in \mathscr{K}$, for all $A, B \in AR_K$, it holds that:*

1. *If A attacks B wrt $att(K)$ then A undercuts B or A contradicts B.*

2. *If A undercuts B then A attacks B wrt $att(K)$.*

Proposition 2.7. *Attack closure (P 2.7) is satisfied by att_{swl}, att_{dwl} and att_{lwl}.*

Proof. Since we do not consider strict rules (undercuts), this principle reduces to: $(A,B) \in att$ implies A contradicts B which is immediate from Definitions 2.13–2.15. □

The principle of effective rebuts enforces a natural interpretation of priorities under conflict: when two defeasible rules lead to a contradiction and so cannot be applied together, then the preferred one should be applied.

[6]The notion of undercut from Principle 7 is the same as in Pollock [253] and ASPIC+ [219]: an argument A undercuts B at $B' \in sub(B)$ iff the last rule $d = last(B') \in RD$ and A states that this defeasible rule d is not applicable $cnl(A) = ab_d$.

Principle 2.8 (Effective rebut). *Let $\mathscr{K}$ be a sensible class of knowledge bases and att be an attack relation assignment defined for $\mathscr{K}$. Then att is said to satisfy the property of effective rebut for $\mathscr{K}$ iff for each $K \in \mathscr{K}$, for all $A_0, A_1 \in AR_K$ containing each exactly one defeasible rule $dr(A_0) = \{d_0\}$ and $dr(A_1) = \{d_1\}$, if A_0 contradicts A_1 then*

$$(A_0, A_1) \in att(K) \;\; \textit{iff} \;\; d_0 \nless d_1.$$

Proposition 2.8. *Effective rebut is satisfied by att_{swl} and att_{dwl}, but not by att_{lwl}.*

Proof. **For att_{swl}.** Let $dr(A) = \{d_1\}$ and $dr(B) = \{d_2\}$ contain each one defeasible rule with A, B contradicting each other. Note that $wl(A) = rank(d_1)$ and $wl(B) = rank(d_2)$. For ($\Rightarrow$), suppose that $(A, B) \in att_{swl}(K)$. As a result, A contradicts B at B' and $wl(A) \nless wl(B')$. Since $RS = \varnothing$, $B = B'$. So, $wl(A) \nless wl(B)$. That is to say, $d_1 \nless d_2$. For ($\Leftarrow$), suppose $d_1 \nless d_2$. So, $wl(A) \nless wl(B)$. Because A contradicts B at B, $(A, B) \in att_{swl}(K)$.

For att_{dwl}. The proof is analogous. Since $RS = \varnothing$, A contradicts B at B' implies $B = B'$ and $d_1 \neq d_2$. Hence, $wl(dr(A) \setminus dr(B)) = wl(A)$ and $wl(dr(B) \setminus dr(A))) = wl(B)$.

For att_{lwl}. Let AR_K contain $A = [\top \Rightarrow a]$, $B = [\top \Rightarrow \neg a]$ and $A^+ = [A \Rightarrow a]$, where all RD rules have strength 1. Then, $dr(A) = \{d_1\}$ and $dr(B) = \{d_2\}$ satisfy $d_1 \nless d_2$ but $(A, B) \notin att_{lwl}(K)$, since (A, B) is not maximal while (B, A) is maximal. □

The last principle, called link orientation (see below for its definition), directs attacks against those links in an argument that are identified as responsible for the argument's weakness.

Definition 2.17 (Weakening operation). *Let $A \in AR_K$ and $AS \subseteq AR_K$. The weakening of A by AS, denoted $A \downarrow AS$ is the set inductively defined by:*

$$A \downarrow AS = \begin{cases} \{[\alpha]\} \cup \{X \in AS : cnl(X) = \alpha\} & \textit{if } A = [\alpha] \textit{ and } \alpha \in BE \\ \{[X_1, \ldots, X_n, r] \mid X_i \in A_i \downarrow AS\} & \textit{if } A = [A_1, \ldots, A_n, r]. \end{cases}$$

Principle 2.9 (Link orientation). *Let $\mathscr{K}$ be a sensible class of knowledge bases and att be an attack relation assignment defined for $\mathscr{K}$. att satisfies link-orientation iff for each $K \in \mathscr{K}$, if $A, B, C \in AR_K$ are such that $C \in B \downarrow AS$, then*

$$\left\{ \begin{array}{c} (A, C) \in att(K) \textit{ and} \\ \forall X \in AS, (A, X) \notin att(K) \end{array} \right\} \quad \textit{implies} \quad (A, B) \in att(K).$$

That is, wrt $att(K)$, if A attacks C (the weakening of B by AS) but none of AS, then A attacks the original argument B.

Proposition 2.9. *Link orientation is not satisfied by any of the attacks att_{swl}, att_{dwl}, att_{lwl}.*

Proof. A counterexample for $att_{swl}, att_{dwl}, att_{lwl}$ can be found by expanding Example 1 with a new fact: $BE = \{a\}$. **For att_{swl}.** Let K consist of:

$$RD = \{\top \overset{1}{\Rightarrow} a, \top \overset{2}{\Rightarrow} \neg b, a \overset{3}{\Rightarrow} b\} \quad \text{and} \quad BE = \{a\}.$$

Let $AS = \{D = [\top \Rightarrow a]\}$, $A = [\top \Rightarrow \neg b]$, $B = [[a] \Rightarrow b]$ and $C = [D \Rightarrow b]$. Note that $C \in B \downarrow AS$, and that $wl(A) = 2$, $wl(B) = 3$ and $wl(C) = 1$. Finally, observe that $(A, C) \in att_{swl}(K)$ and $(A, D) \notin att_{swl}(K)$ for $AS = \{D\}$ while $(A, B) \notin att_{swl}(K)$. **For att_{dwl}.** The same example holds, since for all the previous pairs (X, Y), $wl(X \setminus Y) = wl(X)$. **For att_{lwl}.** The same example works as in att_{dwl}, since all of A, B, C are $\sqsubseteq$-maximal attackers in K and so $att_{lwl}(K) = att_{dwl}(K)$. □

Attack Assignment	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
swl-attack (Def. 2.13)	□	■	■	■	■	■	■	■	□
dwl-attack (Def. 2.14)	□	■	■	■	■	■	■	■	□
lwl-attack (Def. 2.15)	□	□	■	□	■	□	■	□	□

Table 2.1: Principles satisfied (■) by each attack relation assignment. Each number n refers to the Principle Pn listed next: (P 2.1) credulous cumulativity, (P 2.2) context independence, (P 2.3) weak context independence, (P 2.4) attack monotonicity, (P 2.5) irrelevance of redundant defaults, (P 2.6) subargument structure, (P 2.7) attack closure, (P 2.8) effective rebut, and (P 2.9) link orientation.

Theorem 2.1. *The principles satisfied by each attack relation are listed in Table 2.1.*

Proof. This result follows from Propositions 2.1–2.9. □

Discussion of The Principle-based Analysis. Weakest link presumes that the evaluation of an argument depends on that of its subarguments, namely their weakest components. Towards a characterization of PDL, the att_{lwl} attack relation assignment captures this idea by making the attacks from subarguments to depend on its superarguments. This results in a less compositional and more holistic view of attacks, which affects some of the intuitive principles proposed by Dung [113]. This should not be surprising at all, and instead it should be seen as part of the ongoing debate on how intuitive some of these principles are. For the popular notion of weakest link, we have a clash of intuitions. On the one hand, our intuitions on the legitimacy of weakest link and on some of our examples and, on the other, the *prima facie* intuitive principles from Dung. Following Nelson Goodman's notion of *reflective equilibrium* [145], this principle-based analysis should prompt us to search for a balance between intuitions on principles and intuitions on cases. Let us take a detailed look at look-ahead weakest link in Table 2.1.

(P 2.1) Credulous cumulativity has also been challenged by Prakken and Vreeswijk [266, Sec. 4.4], by Prakken and Modgil [218, Sec. 5.2]. Intuitively, the strengthened defeasible conclusion may gain the ability to defeat other arguments that they did not defeat before, which causes the stable extensions to change, thus leading to the violation of credulous cumulativity.

(P 2.2)–(P 2.3) Given our aim to characterize PDL and vindicate its role in non-monotonic reasoning, Context independence (P 2.2) has to be relativized to take part of the context into account, namely the superarguments of an argument. Attack relations based on lookahead weakest link are still independent from external arguments.

(P 2.4) The violation of one of the two directions of Attack monotonicity might be seen as the least palatable consequence of lookahead weakest link. Still, our conjecture is that the other direction (P 2.4, item 2) holds for att_{lwl}.

(P 2.5), (P 2.7) The principle of Irrelevance of redundant defaults (P 2.5) results in an intuitive property of ASPIC+, i.e. a semantic invariance under the weakening of facts into (irrelevant) defaults. Attack closure (P 2.7) captures our understanding of how attacks in ASPIC+ should be defined. Both principles are preserved by att_{lwl}.

(P 2.6), (P 2.8) Despite their intuitive character, Subargument structure and Effective rebuts seem to exclude a relational notion of attacks based on the global structure of an argument, that is, the superarguments it is part of. The violation of these two principles might be a necessary step in any characterization of PDL in terms of attack relation assignments.

(P 2.9) Link orientation is, in view of the counterexample in Proposition 2.9, one of the most disputable principle in the list. It clashes, as (P 2.1) does, with all attack relation assignments inspired by the idea of weakest link. For anyone considering the possibility of argumentation based on weakest link, this counterexample shows that (P 2.9) makes little sense as a general principle.

2.5 PDL and Dung's Principles: an Impossibility Result

As discussed in the previous section, most of the principles proposed by Dung [113] seem indisputable, yet some others hide a partisan view on what argumentation can or cannot be. Context independence, for example, could be used to rule out Brewka's PDL from argumentation altogether. For another example, credulous cumulativity is used by Dung [112, Ex. 7.1] against elitist orderings. In turn, this principle has been further discussed and disputed by Modgil and Prakken [218].

In this section, we offer more evidence against Context independence, in the form of an impossibility result (Theorem 2.2). Any attempt to realize PDL in ASPIC+ should preserve the definitional principle of Attack closure (P 2.7). Theorem 2.2 explains how this is incompatible with the principle of Context independence (P 2.2).

Recall that PDL inductively applies a default of maximal priority amongst those rules that: (i) have not been applied yet, (ii) can be applied and (iii) their application does not raise an inconsistency [194, 334]. We adapt the definitions to structured argumentation.

Definition 2.18 (PDL). *Let $K = (RS, RD, \preceq, BE)$ be a knowledge base. For a set of defeasible rules $R \subseteq RD$, let $K \upharpoonright R = (RS, R, \preceq, BE)$ and define the following sets:*

$$\begin{aligned} cl(K,R) = {}& cnl(AR_{K\upharpoonright R}) \\ appl(K,R) = {}& \{d \in RD \setminus R : bd(d) \subseteq cl(K,R) \text{ and } cl(K, R\cup\{d\}) \text{ is consistent}\}. \end{aligned}$$

A PDL construction for K is any set $\bigcup_{i=0}^{\omega} R(i)$ built inductively as follows:

$$R(0) = \varnothing \qquad \text{and} \qquad R(i+1) = R(i) \cup \{d\} \quad \text{for some } d \in \max\nolimits_{\preceq} appl(K, R(i))$$

*where $\max_{\preceq} \Gamma = \{d \in \Gamma \mid \forall d' \in \Gamma (d \npreceq d')\}$. Then, S is a **PDL extension** of K, denoted as $S \in pdl(K)$, if $S = cnl(K,R)$ for some PDL construction R for K.*

Example 2.7. *Recall the set $RD = \{d_1 : \top \overset{1}{\Rightarrow} a, d_2 : \top \overset{1}{\Rightarrow} b, d_3 : a \overset{2}{\Rightarrow} \neg b, d_4 : b \overset{2}{\Rightarrow} \neg a\}$ in knowledge base $K = (RS, RD, \preceq, BE)$ from Examples 3–2.4. The PDL constructions for K are: $R_1 = \{d_1, d_3\}$ and $R_2 = \{d_2, d_4\}$. These constructions give the PDL extensions $S_1 = \{a, \neg b\}$ and $S_2 = \{b, \neg a\}$ respectively. Figure 2.3 shows an argumentation framework for K. (Note that we omit $\top$ from the PDL extensions and the argument A_0 from AR_K.)*

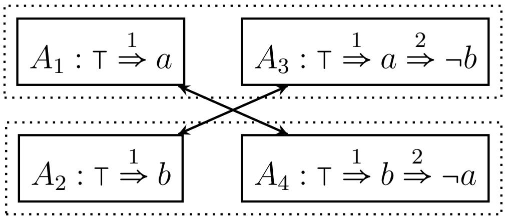


Figure 2.3: AF constructed from Example 2.7. Arrows describe all the possible individual attacks at the subarguments. An attack relation $att(K)$ cannot contain both (A_1, A_4) and (A_2, A_3) if it is to capture the PDL extensions.

Example 2.8. *Let $K_1 = (RS, RD_1, \leq_1, BE)$ be the fragment of K consisting of $RD_1 = RD \setminus \{d_4\}$ with the preference $\leq_1$ given by restricting $\leq$ to the set RD_1. PDL (Def. 2.18) gives: $R_1 = \{d_1, d_3\} \longmapsto S_1 = \{a, \neg b\}$, and $R_3 = \{d_2, d_1\} \longmapsto S_3 = \{a, b\}$.*

The PDL extensions S_1, S_3 are also obtained as (sets of the conclusions of) the stable extensions under the attack relation $att_1 = \{(A_2, A_3), (A_3, A_2)\}$. See Figure 2.4.

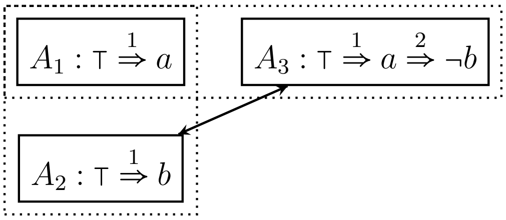


Figure 2.4: Under the attack relation att_1, the stable extensions of $AF_1 = (AR_{K_1}, att_1)$ match the PDL extensions of K_1.

Example 2.9. *Let $K_2 = \{RS, RD_2, \leq_2, BE\}$ now be the fragment of K defined by $RD_2 = RD \setminus \{d_3\}$ and the preference $\leq_2$ obtained by restricting $\leq$ to RD_2. Now PDL gives: $R_2 = \{d_2, d_4\} \longmapsto S_2 = \{b, \neg a\}$, and $R_3 = \{d_1, d_2\} \longmapsto S_3 = \{a, b\}$.*

The PDL extensions S_2, S_3 are also obtained as (sets of the conclusions of) the stable extensions under the attack relation $att_2 = \{(A_1, A_4), (A_4, A_1)\}$. See Figure 2.5.

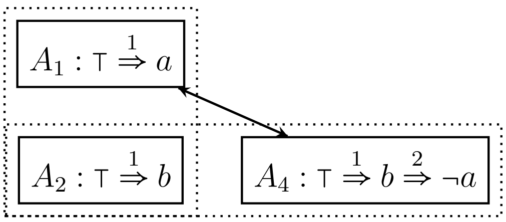


Figure 2.5: Similarly, for $AF_2 = (AR_{K_2}, att_2)$, we have $stb(AF_2) = pdl(K_2)$.

Now we are in a position to prove an impossibility result for Dung's axioms and PDL, under the assumption that the axioms hold for any sensible class of knowledge bases —akin to the universal domain axiom in Arrow's impossibility theorem [21].

Theorem 2.2. *Let att be an attack relation assignment capturing the PDL extensions (say, under stable semantics) and satisfying attack closure (P 2.7). Then att does not satisfy context independence (P 2.2).*

Proof. Let $\mathscr{K}$ be a sensible class of knowledge bases containing K, K_1 and K_2 from Examples 2.7–2.9. Let also att be the attack relation assignment capturing the PDL extensions under stable semantics. Given this attack relation assignment att, the stable extensions must be the following. For $AF_0 = (AR_K, att(K))$: $\mathscr{E}_1 = \{A_1, A_3\}$ and $\mathscr{E}_2 = \{A_2, A_4\}$; for $AF_1 = (AR_{K_1}, att(K_1))$, $\mathscr{E}_1$ and $\mathscr{E}_3 = \{A_1, A_2\}$; for $AF_2 = (AR_{K_2}, att(K_2))$, $\mathscr{E}_2$ and $\mathscr{E}_3$.

The proof is by contradiction. Assume context independence (2.2). Using attack closure (P 2.7), it is only the case that $\mathscr{E}_3 \in stb(AF_1)$ if $(A_2, A_3) \in att(K_1)$. Similarly, $\mathscr{E}_3 \in stb(AF_2)$ can only hold if $(A_1, A_4) \in att(K_2)$. Observe that AR_K contains all these arguments: $\{A_1, A_2, A_3, A_4\}$, and that the preference $\leq$ from K coincides with $\leq_1$ from K_1 on the set $\{A_1, A_2, A_3\}$ and also with $\leq_2$ from K_2 on the set $\{A_1, A_2, A_4\}$. Hence, by context independence, we conclude that $(A_2, A_3), (A_1, A_4) \in att(K)$. But this is impossible: then $\mathscr{E}_3 = \{A_1, A_2\}$ would then become a stable extension of $AF_0 = (AR_K, att(K))$ without being a PDL extension of K. Hence, context independence is not satisfied. □

2.6 Related Work

There is a lot of work in the nonmonotonic logic and logic programming literature on prioritised rules, see e.g. Delgrande et al. [103] for an overview. Pardo and Straßer give an overview of argumentative representations of prioritized default logic, concerning weakest link, they mainly consider dwl [242]. Various authors discussed the dilemma between weakest link and last link [69, 194, 216, 219]. The analysis of weakest link related to swl indicates that it is more complicated and ambiguous than it seems at first sight. With partial orders, ASPIC+ tries to accommodate both in combination with democratic and elitist orders [216, 219], but neither of them is clearly better than the other. Young et al. [334, 336] show that even for total and modular orders, swl cannot always give intuitive conclusions. They also show the correspondence between the inferences made in prioritised default logic (PDL) and dwl with strict total orders. Then they raise the question of the similarity between weakest link and PDL for modular and partial orders. Moreover, Liao et al. [194] give similar results but use other examples to demonstrate that the approach of Young et al. [334, 336] cannot be extended to preorders [194]. Liao et al. [194] use an order puzzle in the form of Example 3 to show that even with modular orders, selecting the correct reasoning procedure is challenging. This leads them to introduce auxiliary arguments and defeats on weakest arguments. Beirlaen et al. [41] point out that weakest link is defined purely in terms of the strength of the defeasible rules used in argument construction. More recently, Lehtonen et al. present novel complexity results for ASPIC+ with preferences that are based on weakest link (swl in this chapter) [189], they rephrase stable semantics in terms of subsets of defeasible elements.

2.7 Summary

In this chapter, we introduced a new weakest link attack relation assignment (lwl) and compared it with the traditional (swl) and disjoint (dwl) versions. We showed that lwl gets the right result for an important example (Ex. 2.3), at the price of loosing context independence —but this seems necessary for weakest link anyway, as shown in Table 2.1. As an alternative, we proposed a weaker context independence principle that is satisfied by lwl. A fine-grained characterization of a class of weakest link attack relation assignments, in the style of the characterizations proposed by Dung [112, 115] for last link, would also help us deepen our understanding of weakest link and vindicate its use in argumentation and non-monotonic reasoning. The core idea behind weakest link is, in our opinion, at least as important as last link for general applications in AI. On this last question, these principle-based analyses might shed some light on long-time debates between weakest link and last link, namely which one suits better each area of application of non-monotonic reasoning. Our principle-based analysis has several original insights, it presents the difference of several kinds of attack relation assignment, explains the nature of weakest link principle and reveals there is still some potential for weakest link attack to improve. By the way, it also has a tight relation with some conceptual and philosophical questions and discussions: We also proved the impossibility of satisfying context independence by any attack relation assignment that captures Brewka's prioritised default logic.

As for future work, following the results presented so far, an immediate goal would be to strengthen the principle-based analysis to knowledge bases containing strict rules (and undercutting attacks). Our conjecture is that the principles satisfied by each attack relation shown in Table 2.1 will be preserved after the addition of strict rules. One main open question for the future of ASPIC+-style structured argumentation is which way to go: introduce auxiliary arguments like Liao et al. [194], or weaken context independence as in this chapter? From a representation point of view, total orders give only one extension, while under partial orders we may have multiple extensions. Thus, another major challenge is how to generalize all the recent insights in this chapter and related work to partial orders as studied in ASPIC+. While the impossibility result immediately extends from modular to partial orders, the affirmative results in our principle-based analysis need not be preserved in the latter. We thus leave for future work deciding whether this is the case for the attack relation assignments we introduced: lwl.

Finally, Table 2.1 also shows that the current principles fail to distinguish swl from dwl, while in practice they behave quite differently. Hence, another goal would be to identify a principle that separates these two attack relation assignments.

Chapter 3

A Principle-based Analysis to Agent Argumentation Semantics

In this chapter, we delve into the acceptance of arguments, which examines the role of agents and their contributions to the argumentation process. Traditionally, argument evaluation has focused on the inherent strength of the arguments themselves, which might overlook the significance of the individuals or entities presenting those arguments. However, recognizing the influence of agents can provide valuable insights into the dynamics of argumentation and help decision-makers make more informed and impartial choices. Ultimately, understanding how the count of agents impacts the balance of pro and con arguments offers a unique perspective on decision-making processes. By considering the sources and entities behind the arguments, decision-makers can effectively navigate the complexities and ambiguities inherent in evaluating opposing viewpoints. Through this exploration, we aim to contribute to the advancement of argumentation theory and provide practical insights for decision-makers facing multifaceted and contentious issues.

Abstract agent argumentation frameworks extend Dung's theory with agents, and in this chapter, we study four types of semantics for them. First, agent defense semantics replaces Dung's notion of defense by some kind of agent defense. Second, social agent semantics prefers arguments that belong to more agents. Third, agent reduction semantics considers the perspective of individual agents. Fourth, agent filtering semantics are inspired by a lack of knowledge. We study five existing principles and we introduce twelve new ones. In total, we provide a full analysis of fifty-two agent semantics and the seventeen principles.

3.1 Introduction

The two volumes of the Handbook of Formal Argumentation [34, 134] explain the central role of Dung's theory of abstract argumentation [110] and many of its variants proposed over the past few decades. However, whereas several papers have proposed *agent-based* variants [27, 176, 137], so far an overview of these variants is lacking. Moreover, the semantics of agent argumentation is related to merging argumentation frameworks [95, 106, 73]. We address the following research questions:

1. What kind of semantics can be defined for agent argumentation frameworks?

2. Which of the principles proposed in the literature [35, 314] do not hold for such agent semantics?

3. What new principles can we define to distinguish the varieties of agent semantics?

For comparison, we distinguish four kinds of semantics for agent argumentation frameworks:

Agent defense approaches adapt Dung's notion of defense for argumentation semantics.

Social approaches [190] are based on counting the number of agents and a reduction to preference-based argumentation [8].

Agent reductions take the perspective of individual agents and create extensions accordingly [143].

Filtering methods are inspired by the knowledge or trust of the agents [19] and leave out some arguments or attacks because they do not belong to any agents.

We make two important observations about the way the principle-based approach is used in formal argumentation in general, and in this paper in particular.

Minimality First, agent-based extensions typically introduce various aspects such as coalitions, knowledge, uncertainty, support, and so on. In line with common practice in the principle-based approach, this paper uses a minimal extension of Dung as a common core to these approaches. We only introduce an abstract set of agents, and we associate arguments with agents and nothing else.

Distinguishability Principles and axioms can be used in many ways. Often, they conceptualize the behavior of a system at a higher level of abstraction. Moreover, in the absence of a standard approach, principles can be used as a guideline for choosing the appropriate definitions and semantics depending on various needs. Therefore, in formal argumentation, principles are often more technical. The most discussed principles are admissibility, directionality and SCC decomposibility, which also play a central role in this paper. In this paper, we focus on principles distinguishing kinds of agent semantics.

The paper is organized as follows. In the following section, we introduce agent argumentation frameworks. We discuss four kinds of semantics for agent argumentation frameworks in the next four sections. As for principles, traditional principles are introduced in the following section and thereafter variants of traditional principles are introduced. Then we introduce eight new agent principles. Finally, we discuss related work, future work, and the conclusions of the paper. Due to space limitations, we only sketch a few proofs.

3.2 Agent Argumentation Framework

This section introduces agent argumentation frameworks. They generalize argumentation frameworks studied by Dung (1995), which are directed graphs, where the nodes are arguments, and the arrows correspond to the attack relation.

Definition 3.1 (Argumentation framework [110])**.** *An* argumentation framework *(AF) is a pair $\langle \mathscr{A}, \rightarrow \rangle$ where $\mathscr{A}$ is a set called arguments, and $\rightarrow \subseteq \mathscr{A} \times \mathscr{A}$ is a binary relation over $\mathscr{A}$ called attack. For a set $S \subseteq \mathscr{A}$ and an argument $a \in \mathscr{A}$, we say that S attacks a if there exists $b \in S$ such that b attacks a, a attacks S if there exists $b \in S$ such that a attacks b, $a^- = \{b \in \mathscr{A} | b \text{ attacks } a\}$, and $S^-_{out} = \{a \in \mathscr{A} \backslash S | \text{ a attacks } S\}$.*

Dung's admissibility-based semantics is based on the concept of defense. A set of arguments defends another argument if they attack all its attackers.

Definition 3.2 (Admissible [110])**.** *Let $\langle \mathscr{A}, \rightarrow \rangle$ be an AF. $E \subseteq \mathscr{A}$ is* conflict-free *iff there are no arguments a and b in E such that a attacks b. $E \subseteq \mathscr{A}$* defends c *iff for all arguments b attacking c, there is an argument a in E such that a attacks b. $E \subseteq \mathscr{A}$ is* admissible *iff it is conflict-free and defends all its elements.*

For their principle-based analysis, Baroni and Giacomin [35] define semantics as a function from argumentation frameworks to sets of subsets of arguments.

Definition 3.3. *Dung semantics [35] Dung semantics is a function σ that associates with an argumentation framework $AF = \langle \mathscr{A}, \rightarrow \rangle$ a set of subsets of $\mathscr{A}$, and the elements of $\sigma(AF)$ are called extensions.*

Dung distinguishes between several definitions of extension.

Definition 3.4 (Extensions [110])**.** *Let $\langle \mathscr{A}, \rightarrow \rangle$ be an AF. $E \subseteq \mathscr{A}$ is a* complete extension *iff it is admissible and it contains all the arguments it defends. $E \subseteq \mathscr{A}$ is a* grounded extension *iff it is the smallest complete extension (for set inclusion). $E \subseteq \mathscr{A}$ is a* preferred extension *iff it is the maximal complete extension (for set inclusion). $E \subseteq \mathscr{A}$ is a* stable extension *iff it is conflict-free, and it attacks each argument that does not belong to E.*

Each kind of extension may be seen as an acceptability semantics that formally rules the argument evaluation process. In this article, we use $\sigma \in \{c, g, p, s\}$ to represent Dung semantics {complete, grounded, preferred, stable}.

Example 3.1 (Two conflicts)**.** *Consider the argumentation framework visualized on the left in Figure 1, where $\mathscr{A} = \{a, b, c, d\}$, $\rightarrow = \{a \rightarrow b, b \rightarrow a, c \rightarrow d, d \rightarrow c\}$. Each argument defends itself. There are nine admissible sets – $\{a\}, \{b\}, \{c\}, \{d\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \varnothing$ – which are all complete extensions. The grounded extension is $\varnothing$. The preferred extensions $\{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}$ are also stable extensions. For example, in an oft-used dinner scenario, we may choose between fish (a) or meat (b), and we may choose between eating at home (c) or going out (d), and these two choices are independent. In structured argumentation, these arguments may have a complex structure, providing the reasons for these conclusions, but in abstract argumentation we do not detail these reasons.*

An agent argumentation framework extends an argumentation framework with a set of agents and a relation associating arguments with agents. Note that an argument can belong to no agent, one agent, or multiple agents.

We write $a \sqsubset \alpha$ for argument a belongs to agent α, or that agent α has argument a.

Definition 3.5 (Agent argumentation framework)**.** *An* agent argumentation framework *(AAF) is a 4-tuple $\langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ where $\mathscr{A}$ is a set of arguments, $\rightarrow \subseteq \mathscr{A} \times \mathscr{A}$ is a binary relation*

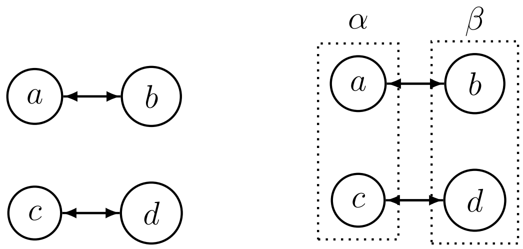


Figure 3.1: An AF and an AAF

over $\mathscr{A}$ called attack, $\mathscr{S}$ is a set of agents or sources, and $\sqsubset \subseteq \mathscr{A} \times \mathscr{S}$ is a binary relation associating arguments with agents. $\mathscr{A}_\alpha = \{a \in \mathscr{A} | a \sqsubset \alpha\}$ for all arguments that belong to agent α, $\mathscr{S}_a = \{\alpha | a \sqsubset \alpha\}$ for all agents that have argument a, $\rightarrow_a = \{x \rightarrow y | x = a \text{ or } y = a\}$ for the attack relations related to argument a, and $\sqsubset_\alpha = \{(a, \alpha) | a \sqsubset \alpha\}$ for the relation between agent α and its arguments.

Example 3.2 (Two conflicts, continued from Example 3.1)**.** *Consider the agent argumentation framework visualized on the right in Figure 3.1. This figure should be read as follows. Each dashed box contains the arguments belonging to the same agent, $\mathscr{S} = \{\alpha, \beta\}$, and $\sqsubset = \{(a, \alpha), (b, \beta), (c, \alpha), (d, \beta)\}$. For example, Alice ($\alpha$) may hold the arguments for eating fish and staying at home, and Bob (β) may hold the arguments for eating meat and going outside.*

3.3 Agent Defense Semantics

We now introduce a new kind of defense for agent argumentation frameworks, which we call agent defense. Roughly, if an agent puts forward an argument, it can only be defended by arguments from the same agent. In extensions with coalitions, we may also consider agents of the same coalition defending each others' arguments [271].

In individual agent defense, only a single agent can defend an argument, whereas in collective agent defense, a set of agents can do that.

Definition 3.6 (Agent Admissible)**.** *Let $\langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be an AAF:*

- *$E \subseteq \mathscr{A}$ is* conflict-free *iff there are no arguments a and b in E such that a attacks b.*
- *$E \subseteq \mathscr{A}$ individually agent defends (agent defends$_1$) c iff there exists an agent α in $\mathscr{S}_c$ such that for all arguments b in $\mathscr{A}$ attacking c, there exists an argument a in $E \cap \mathscr{A}_\alpha$ such that a attacks b.*
- *$E \subseteq \mathscr{A}$ collectively agent defends (agent defends$_2$) c iff for all arguments b in $\mathscr{A}$ attacking c, there exists an agent α in $\mathscr{S}_c$ and an argument a in $E \cap \mathscr{A}_\alpha$ such that a attacks b.*
- *$E \subseteq \mathscr{A}$ is* agent admissible$_i$ *iff it is conflict-free and agent defends$_i$ all its elements, for i in $\{1, 2\}$.*

The following example illustrates agent defense, and its role in so-called reinstatement. Though reinstatement is considered by many to be a desirable property, there is also a minority opinion that argues that reinstatement should not hold in general, c.f. the arguments and examples of Horty [161]. Example 3.3 shows that there is a middle way in this debate. Agent defense semantics allows for reinstatement if all the arguments belong to the same agent, but not if the arguments belong to distinct agents.

Example 3.3 (Reinstatement)**.** *Consider the agent argumentation framework visualized in Figure 3.2, where* $\mathscr{A} = \{a, b, c\}$, $\rightarrow = \{c \rightarrow b, b \rightarrow a\}$, $\mathscr{S} = \{\alpha, \beta, \gamma\}$ *and* $\sqsubset = \{(a, \alpha), (b, \beta), (c, \gamma)\}$*. Argument* c *defends argument* a*, but it does not agent defend it. For example, in the dinner scenario, Alice (*α*) may hold an argument in favor of eating meat, Bob (*β*) holds a better argument in favor of not eating meat but fish, and Cayrol (*γ*) holds an argument asking why fish is not an option (c). Assuming that Alice and Cayrol are not in a coalition, Cayrol does not agent defend the argument of Alice against the attacker of Bob.*

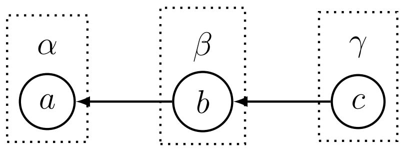


Figure 3.2: Agent Reinstatement

Definition 3.7 (Agent semantics)**.** *An agent semantics is a function* δ *that associates a set of subsets of* $\mathscr{A}$ *with an agent argumentation framework* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*, and the elements of* $\delta(AAF)$ *are called agent extensions.*

We use Sem_1 and Sem_2 to represent agent semantics based on individual defense and collective defense respectively.

Definition 3.8 (Agent extensions)**.** *Let* $\langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*:*

- $E \subseteq \mathscr{A}$ *is an* agent complete$_i$ extension *iff* E *is agent admissible*$_i$ *and it contains all the arguments it agent defends*$_i$*, for* $i \in \{1, 2\}$.
- $E \subseteq \mathscr{A}$ *is an* agent grounded$_i$ extension *iff it is the minimal agent complete*$_i$ *extension (for set inclusion), for* $i \in \{1, 2\}$.
- $E \subseteq \mathscr{A}$ *is an* agent preferred$_i$ extension *iff it is the maximal agent complete*$_i$ *extension (for set inclusion), for* $i \in \{1, 2\}$.
- $E \subseteq \mathscr{A}$ *is an* agent stable$_i$ extension *iff it is conflict-free and it attacks all the arguments in* $\mathscr{A} \backslash E$*, for* $i \in \{1, 2\}$.

The following two examples illustrate agent extensions.

Example 3.4 (Two conflicts, continued from Example 3.2)**.** *Reconsider Figure 3.1. Each argument agent defends itself, therefore the agent complete extensions are the same as the complete extensions of the corresponding extensions of the argumentation framework without considering agents. The agent grounded, preferred and stable extensions are also the same as those of the argumentation framework.*

Example 3.5 (Reinstatement, continued from Example 3.3)**.** *Reconsider Figure 3.2. The individual and collective agent complete extension is* $\{c\}$*. It is also the unique individual and collective agent grounded and preferred extension. There is no agent stable extension. When the only accepted argument is* c*, it suggests a vegetarian dinner. Using stable semantics, no agreement is reached on dinner.*

The following example illustrates the difference between individual agent defense and collective agent defense. In particular, if a set of arguments individually agent defends another argument, then it also collectively agent defends it, but the example illustrates that the opposite does not always hold.

Example 3.6 (Collective defense)**.** *Consider the agent argumentation framework visualized in Figure 3.3, where* $\mathscr{A} = \{a, b_1, b_2, c_1, c_2\}$*,* $\rightarrow = \{c_1 \rightarrow b_1, b_1 \rightarrow a, c_2 \rightarrow b_2, b_2 \rightarrow a\}$*,* $\mathscr{S} = \{\alpha, \beta, \gamma\}$*,* $\sqsubset = \{(a, \alpha), (a, \beta), (b_1, \gamma), (b_2, \gamma), (c_1, \alpha), (c_2, \beta)\}$*. For example, in the dinner scenario, Alice and Bob argue in favor of eating meat, Cayrol has two better arguments for eating fish, but Alice argues why the first argument of Cayrol cannot be accepted, and Bob argues why the second argument of Cayrol cannot be accepted.*

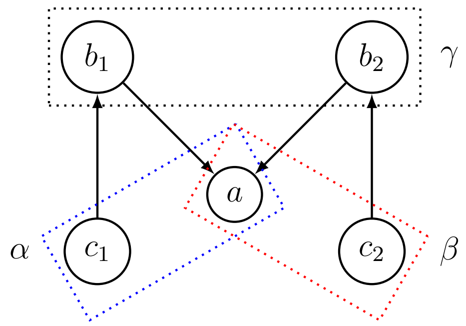


Figure 3.3: There is no single agent defending argument a

$\{c_1, c_2\}$ *collectively agent defend argument a, but they do not individually agent defend it. The agent admissible*$_1$ *extensions are* $\varnothing$*,* $\{c_1\}$*,* $\{c_2\}$ *and* $\{c_1, c_2\}$*. The only agent complete*$_1$ *extension is* $\{c_1, c_2\}$*, which is also the agent grounded*$_1$ *extension and the unique agent preferred*$_1$ *extension. There is no agent stable*$_1$ *extension. The agent admissible*$_2$ *extensions are* $\varnothing$*,* $\{c_1\}$*,* $\{c_2\}$*,* $\{c_1, c_2\}$ *and* $\{a, c_1, c_2\}$*. The only agent complete*$_2$ *extension is* $\{a, c_1, c_2\}$*, which is also the grounded*$_2$ *extension, the unique preferred*$_2$ *extension and stable*$_2$ *extension. Though Alice and Bob do not form a coalition in the sense that they defend each others' arguments, by using collective defense they can form a coalition in the sense that together they reinstate the argument in favor of eating meat.*

The following example illustrates another aspect of agent defense.

Example 3.7 (Agent defense)**.** *Consider Figure 4, where* $\mathscr{A} = \{a_1, a_2, b, c\}$*,* $\rightarrow = \{c \rightarrow b, b \rightarrow a_2, b \rightarrow a_1\}$*,* $\mathscr{S} = \{\alpha, \beta, \gamma\}$ *and* $\sqsubset = \{(a_1, \alpha), (a_2, \gamma), (b, \beta), (c, \alpha)\}$*. The unique individual (collective) agent complete extension, grounded extension and preferred extension is* $\{a_1, c\}$*. There is no stable extension. When we compute extensions using SCC-recursion, we first consider argument c, then argument b, and finally argument* a_1 *and* a_2*. When accepting c, we cannot simply remove b.*

In the following three sections, we introduce several other kinds of semantics based on various kinds of reductions.

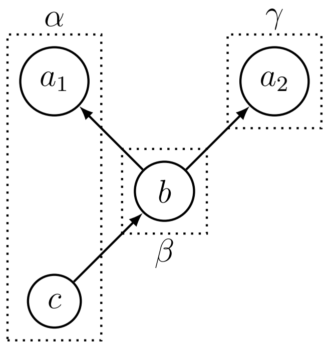


Figure 3.4: Agent defense

3.4 Social Agent Semantics

In this section, we introduce so-called social semantics, which is based on a reduction to preference-based argumentation for each argument, counting the number of agents that have the argument. It thus interprets agent argumentation as a kind of voting, as studied in social choice theory or judgment aggregation. It is not the only way to define social agent semantics, but given the formal setting we have adopted, it seems the simplest and most natural possibility.

We first give the definition of a preference-based argumentation framework.

Definition 3.9. (Preference-based argumentation framework) *A preference-based argumentation framework (PAF) is a 3-tuple* $\langle \mathscr{A}, \rightarrow, \succ \rangle$ *where* $\mathscr{A}$ *is a set of arguments,* $\rightarrow \subseteq \mathscr{A} \times \mathscr{A}$ *is a binary attack relation, and* $\succ$ *is a partial order (irreflexive and transitive) over* $\mathscr{A}$ *called preference relation.*

Amgoud and Vesic [10] introduce two different reductions of preference, while van der Torre and Vesic [314] introduce two more. We refer to these papers for an explanation and motivation, and illustrate the difference between the reductions in Example 8 below.

Definition 3.10 (Reductions of PAF to AF (PR))**.** *Given a* $PAF = \langle \mathscr{A}, \rightarrow, \succ \rangle$*:*

- $PR_1(PAF) = \langle \mathscr{A}, \rightarrow' \rangle$, *where* $\rightarrow' = \{a \rightarrow' b | a \rightarrow b, b \nsucc a\}$.
- $PR_2(PAF) = \langle \mathscr{A}, \rightarrow' \rangle$, *where* $\rightarrow' = \{(a \rightarrow' b | a \rightarrow b, b \nsucc a$ *or* $b \rightarrow a,$ *not* $a \rightarrow b, a \succ b\}$.
- $PR_3(PAF) = \langle \mathscr{A}, \rightarrow' \rangle$, *where* $\rightarrow' = \{a \rightarrow' b | (a \rightarrow b, b \nsucc a$ *or* $a \rightarrow b,$ *not* $b \rightarrow a\}$.
- $PR_4(PAF) = \langle \mathscr{A}, \rightarrow' \rangle$, *where* $\rightarrow' = \{a \rightarrow' b | a \rightarrow b, b \nsucc a,$ *or* $b \rightarrow a,$ *not* $a \rightarrow b, a \succ b,$ *or* $a \rightarrow b,$ *not* $b \rightarrow a\}$.

In social agent semantics, an argument is preferred to another argument if it belongs to more agents. The reduction from AAF to PAF is used as an intermediary step for social agent semantics.

Definition 3.11 (Social Reductions of AAF to PAF (SAP))**.** *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, $SAP(AAF) = \langle \mathscr{A}, \rightarrow, \succ \rangle$ *with* $\succ = \{a \succ b || \mathscr{S}_a| > |\mathscr{S}_b|\}$.

There are four definitions of social reduction, and σ is in $\{c, g, p, s\}$, thus, we have sixteen social agent semantics.

Definition 3.12 (Social Reductions of AAF to AF (SR)). *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, $SR_i(AAF) = PR_i(SAP(AAF))$*, and* PR_i *is one of the four reductions of PAF to AF, where the semantics* $\delta(AAF) = \sigma(SR_i(AAF)) = \sigma(PR_i(SAP(AAF)))$ *for* $i \in \{1, 2, 3, 4\}$.

Example 3.8 (Social reasoning). *Consider the agent argumentation framework (AAF) on the left in Figure 3.5, where* $\mathscr{A} = \{a, b\}$, $\rightarrow = \{a \rightarrow b\}$, $\mathscr{S} = \{\alpha, \beta\}$ *and* $\sqsubset = \{(a, \alpha), (b, \alpha), (b, \beta)\}$. *Argument* b *is preferred to argument* a *because it belongs to more agents. The preference-based argumentation framework (PAF) is visualized to the right of the AAF in Figure 3.5:* $\mathscr{A} = \{a, b\}$, $\rightarrow = \{a \rightarrow b\}$*, and* $> = \{b > a\}$. *To the right of PAF, there are four corresponding argumentation frameworks (AFs) after* SR_1 *to* SR_4*, the extensions of each are listed in Table 3.1.*

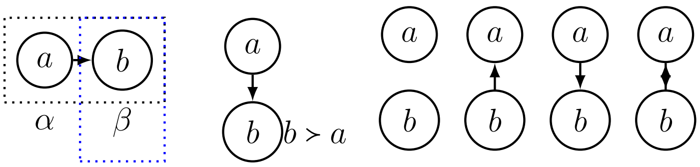


Figure 3.5: Social reduction

Sem.	ℂ	𝔾	ℙ	𝕊
SR$_1$	$\{\{a, b\}\}$	$\{\{a, b\}\}$	$\{\{a, b\}\}$	$\{\{a, b\}\}$
SR$_2$	$\{\{b\}\}$	$\{\{b\}\}$	$\{\{b\}\}$	$\{\{b\}\}$
SR$_3$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
SR$_4$	$\{\emptyset, \{a\}, \{b\}\}$	$\emptyset$	$\{\{a\}, \{b\}\}$	$\{\{a\}, \{b\}\}$

Table 3.1: The semantics of four corresponding argumentation frameworks (AFs) after SR_1 to SR_4. We refer to Dung's semantics as follows: Complete (ℂ), Grounded (𝔾), Preferred (ℙ), Stable (𝕊), and the same convention holds for all the others.

3.5 Agent Reduction Semantics

In this section, we introduce the third class of semantics. Agent reductions take the perspective of each agent and create extensions accordingly. In an abstract sense, an agent prefers its own arguments over the arguments of the other agents. It is again based on a reduction of agent argumentation frameworks to preference-based argumentation frameworks, just like social agent semantics, but now in a completely different way. One difference between social reductions in the previous section and the agent reductions in this section is that in the previous section, there is only reduction AF for every AAF, whereas in this section there is a set of such reductions, one for each agent, and then we take the union of all the reductions. Again, as in the previous section, the four kinds of reduction of preference-based argumentation frameworks lead to four kinds of agent reductions.

Definition 3.13 (Agent Reductions of AAF to PAF(AAP)). *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, $AAP(AAF, \alpha) = \langle \mathscr{A}, \rightarrow, > \rangle$ *with* $> = \{a > b | a \sqsubset \alpha$ *and not* $b \sqsubset \alpha\}$.

As in social agent semantics, there are four definitions of agent reductions, and σ is in $\{c, g, p, s\}$. Thus, we have sixteen agent reduction semantics.

Definition 3.14 (Agent Reductions of AAF to AF (AR)). *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, *for* $\alpha \in \mathscr{S}$, PR_i *is one of the four reductions of PAF to AF, where the semantics* $\delta(AAF) = \sigma(AR_i(AAF)) = \sigma(\bigcup_{\alpha \in \mathscr{S}} PR_i(AAP(AAF, \alpha)))$ *for* $i \in \{1, 2, 3, 4\}$. *For* $AF_1 = \langle \mathscr{A}_1, \rightarrow_1 \rangle$ *and* $AF_2 = \langle \mathscr{A}_2, \rightarrow_2 \rangle$, *let* $AF_1 \cup AF_2 = \langle \mathscr{A}_1 \cup \mathscr{A}_2, \rightarrow_1 \cup \rightarrow_2 \rangle$.

Example 3.9 (Agent reduction). *Reconsider the AAF on the left in Figure 3.5. Firstly, consider the reduction for agent* β. *We have that argument b is preferred over argument a, thus, we get the same PAF as in Figure 3.5, though for a very different reason compared to that from social reduction. For agent* α, *the PAF makes all arguments equivalent, and the AF is simply the same as for the trivial reduction. To compute the agent extensions of the AAF, we take the union of the reductions for each agent. The AFs of* AR_i *are the union of the AFs of* SRi *in Table 3.1 with the AF in which* a *attacks* b *(the reduction for agent* α*). Thus,* $AR_1 = AR_3 = \langle \{a, b\}, \{a \rightarrow b\} \rangle$, *and* $AR_2 = AR_4 = \langle \{a, b\}, \{a \rightarrow b, b \rightarrow a\} \rangle$. *For instance, after* AR_1, *the AF of agent* α *is* $AR_1(AAF, \alpha) = \langle \{a, b\}, \{a \rightarrow b\} \rangle$, *while* $AR_1(AAF, \beta) = \langle \{a, b\}, \{\varnothing\} \rangle$, *so the union is* $\langle \{a, b\}, \{a \rightarrow b\} \rangle$, *and then we compute the extensions of this union. The result is Table 3.2 below for the sixteen agent reduction semantics we consider.*

Sem.	$\mathbb{C}$	$\mathbb{G}$	$\mathbb{P}$	$\mathbb{S}$
AR_1	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
AR_2	$\{\varnothing, \{a\}, \{b\}\}$	$\varnothing$	$\{\{a\}, \{b\}\}$	$\{\{a\}, \{b\}\}$
AR_3	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
AR_4	$\{\varnothing, \{a\}, \{b\}\}$	$\varnothing$	$\{\{a\}, \{b\}\}$	$\{\{a\}, \{b\}\}$

Table 3.2: The semantics of four corresponding argumentation frameworks (AF) after AR_1 to AR_4.

3.6 Agent Filtering Semantics

In this section, we introduce the fourth kind of semantics for agent argumentation frameworks. Agent filtering semantics remove arguments that do not belong to an agent (OrphanReduction), or they remove attacks that do not belong to an agent (NotBothReduction), where an attack belongs to an agent if both the attacker and the attacked argument belong to the agent.

Definition 3.15 (Agent Reductions of AAF to AF). *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$:

- *OrphanRemoval (OR):* $OR(AAF) = \langle \mathscr{A}', \rightarrow' \rangle$ *where* $\mathscr{A}' = \{a | \exists \alpha \in \mathscr{S}$ *such that* $a \sqsubset \alpha, \}, \rightarrow \cap \mathscr{A}' \times \mathscr{A}'$.
- *NotBothReduction (NBR):* $NBR(AAF) = \langle \mathscr{A}, \rightarrow' \rangle$ *where* $\rightarrow' = \{(a \rightarrow b | \exists \alpha \in \mathscr{S}$ *such that* $a \sqsubset \alpha$, *and* $b \sqsubset \alpha\}$.

Example 3.10 (Epistemic reasoning)**.** *Consider the two AAFs in Figure 3.6. For the figure on the left, we may say that argument a is not known, as there is no agent that has it, and for the figure on the right, we may say that the attack is unknown, because there is no agent that has both arguments a and b. The filtering methods remove such unknown arguments (OrphanReduction) and unknown attacks (NotBothReduction).*

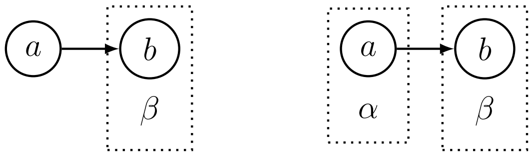


Figure 3.6: Unknown

We refer to Example 3, 7 in the paper of Cayrol and Lagasquie-Schiex [83] which illustrates semantics_8 and semantics_9.

3.7 Traditional Principles

In this section, we repeat six important principles from the literature. As the baseline for the principles, we also include Dung's semantics. It is based on the so-called trivial reduction, which simply ignores the agents and the relation between agents and arguments.

Definition 3.16 (Trivial Reduction (TR))**.** *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle, TR(AAF) = \langle \mathscr{A}, \rightarrow \rangle$.

Principle 3.1 (Conflict-free [35])**.** *An agent semantics* δ *satisfies the conflict-free principle iff for every* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*, for all* $E \in \delta(AAF)$*, there are no arguments a and b in* E *such that a attacks b.*

The conflict-free principle reflects the intuitive idea that an extension contains the arguments that can be accepted together, and that the conflicting arguments cannot be included in the same extension, while the admissibility principle reflects that all arguments are defended.

Principle 3.2 (Admissibility [35])**.** *An agent semantics* δ *satisfies the admissibility principle iff for every* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*, every* $E \in \delta(AAF)$ *is admissible in* $\langle \mathscr{A}, \rightarrow \rangle$.

Directionality and SCC-recursiveness are introduced by Baroni, Giacomin, and Guida [36]. These principles reflect the idea that we can decompose an argumentation framework into sub-frameworks so that the semantics can be defined locally. For the directionality principle, they first introduce the definition of an unattacked set.

Definition 3.17 (Unattacked Set)**.** *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*, a set* $\mathscr{U}$ *is unattacked iff there exists no* $a \in \mathscr{A} \backslash \mathscr{U}$ *such that a attacks an argument in* $\mathscr{U}$*. The set of unattacked sets in AAF is denoted as* $\mathscr{U}\mathscr{S}(AAF)$.

Definition 3.18 (Restriction)**.** *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*, and let* $\mathscr{U} \subseteq \mathscr{A}$ *be a set of arguments, the restriction of AAF to* $\mathscr{U}$ *is the agent abstract framework* $AAF{\downarrow_{\mathscr{U}}} = \langle \mathscr{U}, \rightarrow \cap \mathscr{U} \times \mathscr{U}, \mathscr{S}, \sqsubset \cap \mathscr{U} \times \mathscr{S} \rangle$.

Principle 3.3 (Directionality [35]). *An agent semantics δ satisfies the directionality principle iff for every $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, for every $\mathscr{U} \in \mathscr{U}\mathscr{S}(AAF)$, it holds that $\delta(AAF{\downarrow_{\mathscr{U}}}) = \{E \cap \mathscr{U} | E \in \delta(AAF)\}$.*

Proposition 3.1. *Agent $stable_1$ semantics and agent $stable_2$ semantics (Def. 3.8) do not satisfy Principle 3.3.*

Proof. We use a counter-example to prove Proposition 1. Assume an $AAF = \langle \{a_1, a_2, a_3, b\}, \{b \rightarrow a_3, a_3 \rightarrow a_1, a_1 \rightarrow a_2, a_2 \rightarrow a_3\}, \{\alpha\}, \{b \sqsubset \alpha, a_1 \sqsubset \alpha, a_2 \sqsubset \alpha, a_3 \sqsubset \alpha\}\rangle$. The unattacked set of arguments is $U = \{b\}$. The stable extension of $(AAF \downarrow U)$ is $\{b\}$. However, there is no stable extension of this AAF. $\delta(\text{AAF}{\downarrow_{\mathscr{U}}}) \neq \{E \cap \mathscr{U} | E \in \delta(AAF)\}$, thus, Agent $stable_1$ semantics and agent $stable_2$ semantics (Def. 8) do not satisfy Principle 3. □

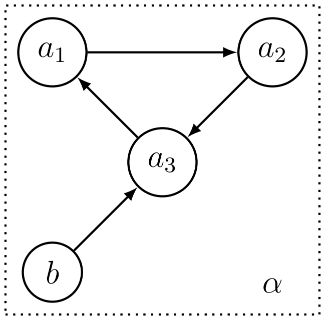


Figure 3.7: A counterexample to prove Proposition 1

The SCC-recursiveness is based on the notion of strongly connected components from graph theory.

Definition 3.19 (Strongly Connected Component). *Let an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$. The binary relation of path-equivalence between nodes, denoted as $PE_{AAF} \subseteq (\mathscr{A} \times \mathscr{A})$, is defined as follows:*

- *for every $a \in \mathscr{A}$, $(a, a) \in PE_{AAF}$*
- *given two distinct arguments $a, b \in \mathscr{A}$, we say that $((a, b) \in PE_{AAF}$ iff there is a path from a to b and a path from b to a.*

The strongly connected components of AAF are the equivalence classes of arguments under the relation of path-equivalence. The set of strongly connected components is denoted by $SCCS_{AAF}$.

Given an argument $a \in \mathscr{A}$, notation $SCCS_{AAF}(a)$ stands for the strongly connected component that contains a. In the particular case where the argumentation framework is empty, i.e., $AAF = \langle \varnothing, \varnothing, \varnothing, \varnothing \rangle$, we assume that $SCCS_{AAF} = \{\varnothing\}$. The choice of extensions of the antecedent strongly connected components determines a partition of the arguments of a strongly connected component S into three subsets: defeated (D), provisionally defeated (P) and undefeated (U) [36].

In words, the set $D_{AAF}(S, E)$ consists of the arguments of S being attacked by E from outside S, the set $U_{AAF}(S, E)$ consists of the arguments in S that are not attacked by E from outside S and are defended by E and $P_{AFF}(S, E)$ consists of the arguments in S that are not attacked by E from outside S and are not defended by E.

Definition 3.20 (D, P, U, UP)**.** *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*, a set* $E \subseteq \mathscr{A}$ *and a strongly connected component* $S \in SCCS_{AAF}$

- $D_{AAF}(S, E) = \{a \in S | (E \cap S^{-}_{out})$ *attacks* $a\}$
- $P_{AAF}(S, E) = \{a \in S | (E \cap S^{-}_{out})$ *does not attack a and* $\exists b \in (S^{-}_{out} \cap a^{-})$ *such that E does not attack* $b\}$.
- $U_{AAF}(S, E) = S \backslash (D_{AAF}(S, E) \cup P_{AAF}(S, E))$
- $UP_{AAF}(S, E) = U_{AAF}(S, E) \cup P_{AAF}(S, E)$.

We now present the notion of SCC-recursiveness, which was introduced by Baroni, Giacomin, and Guida [36].

Principle 3.4. (SCC-recursiveness [36]) *Agent semantics* δ *satisfies the SCC-recursiveness principle iff for every* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*, we have* $\delta(AAF) = \mathscr{G}(AAF, \mathscr{A})$*, where for every AAF and for every set* $\mathscr{C} \subseteq \mathscr{A}$*, the function* $\mathscr{G}(AAF, \mathscr{C}) \subseteq 2^{\mathscr{A}}$ *is defined as follows: for every* $E \subseteq \mathscr{A}$*,* $E \in \mathscr{G}(AAF, \mathscr{C})$ *iff*

- *when* $|SCCS_{AAF}| = 1$*,* $E \in \mathscr{B}(AAF, \mathscr{C})$*,*
- *otherwise,* $\forall S \in SCCS_{AAF}$*,* $(E \cap S) \in \mathscr{G}(AAF{\downarrow}_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap \mathscr{C})$*,*

where $\mathscr{B}(AAF, \mathscr{C})$ *is a function called a base function that given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$*, such that* $|SCCS_{AAF}| = 1$ *and a set* $\mathscr{C} \subseteq \mathscr{A}$ *gives a subset of* $2^{\mathscr{A}}$.

Baumann, Brewka, and Ulbricht [39] introduce the modularization principle. By definition, AAF^{E} is the sub-framework of AAF obtained by removing the so-called range of E, the corresponding attacks, and the relation with agents.

Definition 3.21 (E-reduct)**.** *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ *and* $E \subseteq \mathscr{A}$*, let* $E^{+} = \{a \in \mathscr{A} | E$ *attacks* $a\}$*,* $E^{\oplus} = E \cup E^{+}$ *and* $E^{*} = \mathscr{A} \backslash E^{\oplus}$*. The E-reduct of AAF is the* $AAF^{E} = \langle E^{*}, R \cap (E^{*} \times E^{*}), \mathscr{S}, \sqsubset \cap (\mathscr{S} \times E^{*}) \rangle$.

Principle 3.5 (Modularity)**.** *An agent semantics* δ *satisfies modularization if for any AAF, we have* $E \in \delta(AAF)$ *and* $E' \in \delta(AAF^{E})$ *implies* $E \cup E' \in \delta(AAF)$.

The modularity principle is related to the robustness principles of Rienstra et al. [278], which consider the addition and removal of arguments and attacks. We consider here only argument removal, which we call argument modularity.

Table 3.3 provides full analysis of the traditional five principles. The first line of the trivial reduction lists a well-known analysis of which of these principles hold for Dung's semantics. Unsurprisingly, several easy examples we have already discussed here show that few of the traditional principles hold for agent semantics. This is particularly a problem for SCC-recursiveness and modularity, because we cannot apply the corresponding recursive algorithm to compute the semantics. In the next section, we therefore introduce some variants of admissibility, SCC-recursion and modularity that are based on agent defense.

Sem.	P 3.1	P 3.2	P3.3	P 3.4	P 3.5
TR	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{CGPS}$	$\mathbb{CPS}$
Sem_1	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{S}$	$\mathbb{S}$
Sem_2	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{S}$	$\mathbb{S}$
SR_1	×	×	$\mathbb{CGP}$	×	×
SR_2	$\mathbb{CGPS}$	×	×	×	×
SR_3	×	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$
SR_4	$\mathbb{CGPS}$	$\mathbb{G}$	×	×	$\mathbb{G}$
AR_1	×	×	$\mathbb{CGP}$	×	$\mathbb{S}$
AR_2	$\mathbb{CGPS}$	×	×	×	×
AR_3	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$
AR_4	$\mathbb{CGPS}$	$\mathbb{G}$	×	×	$\mathbb{G}$
OR	$\mathbb{CGPS}$	×	$\mathbb{CGP}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$
NBR	×	×	$\mathbb{CGP}$	×	$\mathbb{S}$

Table 3.3: Comparison of reductions and traditional principles. When a principle is never satisfied by a certain reduction for all semantics, we use the × symbol, and we use a question mark to represent an open problem. P 3.1 refers to Principle 3.1, and the same convention holds for all the others.

3.8 Variants of Traditional Principles

The agent admissibility principle is a straightforward adaptation of the admissibility principle, in which defense is replaced by agent defense. Since there are two kinds of admissibility, one for individual defense and one for collective defense, we end up with two agent admissibility principles.

Principle 3.6 (Agent Admissibility$_1$). *An agent semantics δ satisfies the agent admissibility$_1$ principle iff for every $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, every $E \in \delta(AAF)$ is agent admissible$_1$.*

Principle 3.7 (Agent Admissibility$_2$). *An agent semantics δ satisfies the agent admissibility$_2$ principle iff for every $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, every $E \in \delta(AAF)$ is agent admissible$_2$.*

The agent SCC-recursiveness principles are also adapted by replacing defense with agent defense, and again we end up with two principles for individual and collective defense. What needs to be adapted is the definition of P, the provisionally defeated arguments. Roughly, P stands for the case that an argument is not defended against b in E outside of S. Likewise, AP stands for the case that an argument a is not agent defended$_i$ against b in E from outside S.

To define agent SCC-recursiveness, we define AD$_i$, AP$_i$, AU$_i$, and AUP$_i$ under individual agent defense and collective agent defense.

Definition 3.22 (AD$_i$, AP$_i$, AU$_i$, AUP$_i$). *Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, a set $E \subseteq \mathscr{A}$ and a strongly connected component $S \in SCCS_{AAF}$, we define:*

- $AD_{iAAF}(S, E) = D_{iAAF}(S, E)$
- $AP_{1AAF}(S, E) = \{a \in S | (E \cap S^-_{out})$ *does not attack a, and* $\forall \alpha \in \mathscr{S}_a, \exists b \in (S^-_{out} \cap a^-)$ *such that* $E \cap \mathscr{A}_{\mathscr{S}_a}$ *does not attack* $b.\}$

- $AP_{2AAF}(S,E) = \{a \in S | (E \cap S^-_{out})$ *does not attack* a *and* $\exists b \in (S^-_{out} \cap a^-)$ *such that* $\forall \alpha$ *in* $\mathscr{S}_a, E \cap \mathscr{A}_\alpha$ *does not attack* $b.\}$
- $AU_{iAAF}(S,E) = S \backslash (AD_{iAAF}(S,E) \cup AP_{iAAF}(S,E))$
- $AUP_{iAAF}(S,E) = AU_{iAAF}(S,E) \cup AP_{iAAF}(S,E)$.

Principle 3.8 (Agent SCC-recursiveness$_1$)**.** *An agent semantics δ satisfies the agent SCC-recursiveness$_1$ principle iff for every $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, we have $\delta(AAF) = \mathscr{G}(AAF, \mathscr{A})$, where for every AAF and for every set $\mathscr{C} \subseteq \mathscr{A}$, the function $\mathscr{G}(AAF, \mathscr{C}) \subseteq 2^{\mathscr{A}}$ is defined as follows: for every $E \subseteq \mathscr{A}$, $E \in \mathscr{G}(AAF, \mathscr{C})$ iff*

- *when* $|SCCS_{AAF}| = 1$, $E \in \mathscr{B}(AAF, \mathscr{C})$,
- *otherwise,* $\forall S \in SCCS_{AAF}$, $(E \cap S) \in \mathscr{G}(AAF{\downarrow}_{AUP_{1AAF}(S,E)}, AU_{1AAF}(S,E) \cap \mathscr{C})$,

Principle 3.9 (Agent SCC-recursiveness$_2$)**.** *An agent semantics δ satisfies the agent SCC-recursiveness$_2$ principle iff for every $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, we have $\delta(AAF) = \mathscr{G}(AAF, \mathscr{A})$, where for every AAF and for every set $\mathscr{C} \subseteq \mathscr{A}$, the function $\mathscr{G}(AAF, \mathscr{C}) \subseteq 2^{\mathscr{A}}$ is defined as follows: for every $E \subseteq \mathscr{A}$, $E \in \mathscr{G}(AAF, \mathscr{C})$ iff*

- *when* $|SCCS_{AAF}| = 1$, $E \in \mathscr{B}(AAF, \mathscr{C})$,
- *otherwise,* $\forall S \in SCCS_{AAF}$, $(E \cap S) \in \mathscr{G}(AAF{\downarrow}_{AUP_{2AAF}(S,E)}, AU_{2AAF}(S,E) \cap \mathscr{C})$,

Table 3.4 shows the comparison between agent semantics and agent admissibility principles and agent SCC-recursion. This is important, since it proves that we can have an efficient SCC-recursiveness algorithm for the new agent semantics. The table also shows that for P 3.7 and P 3.9, collective defense implies individual defense. Finally, the table shows that the adapted principles, like the traditional ones, are not very useful for distinguishing between the reduction-based semantics, i.e. the social agent semantics, the agent reduction semantics, and the agent filtering semantics. Therefore, we introduce some new principles in the remainder.

3.9 New Agent Principles

In this section, we introduce eight new principles to distinguish agent semantics. Principle 3.10 says that if more agents adopt an argument that is accepted, this does not affect the extension.

Principle 3.10 (AgentAdditionPersistence)**.** *An agent semantics δ satisfies AgentAddition-Persistence iff for every $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, $E \in \delta(AAF)$, $\alpha \in \mathscr{S}$ and $a \in E$, we have $E \in \delta(\langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \cup (a, \alpha) \rangle)$.*

Proposition 3.2. *AR_1 to AR_4 and SR_1 to SR_4 do not satisfy Principle 3.10 for complete semantics.*

Proof. We use a counter-example to prove Proposition 3.2. Assume $AAF_1 = \langle \{a,b\}, \{a \rightarrow b, b \rightarrow a\}, \{\alpha, \beta\}, \{a \sqsubseteq \alpha, b \sqsubseteq \beta\} \rangle$, $AR_i(AAF_1) = SR_i(AAF_1) = \langle \{a,b\}, \{a \rightarrow b, b \rightarrow a\} \rangle$. Let $AAF_2 = \langle \{a,b\}, \{a \rightarrow b, b \rightarrow a\}, \{\alpha, \beta\}, \{a \sqsubseteq \alpha, a \sqsubseteq \beta, b \sqsubseteq \beta\} \rangle$. $AR_i(AAF_2) = SR_i(AAF_2) =$

Sem.	P3.6	P3.7	P3.8	P3.9
TR	×	×	×	×
Sem_1	$\mathbb{CGP}$	$\mathbb{CGP}$	$\mathbb{S}$	$\mathbb{S}$
Sem_2	×	$\mathbb{CGP}$	$\mathbb{S}$	$\mathbb{S}$
SR_1	×	×	×	×
SR_2	×	×	×	×
SR_3	×	×	×	×
SR_4	×	×	×	×
AR_1	×	×	×	×
AR_2	×	×	×	×
AR_3	×	×	×	×
AR_4	×	×	×	×
OR	×	×	×	×
NBR	×	×	×	×

Table 3.4: Comparison of the reductions and agent admissibility principles, and agent SCC-recursion.

$\langle\{a, b\}, \{a \to b\}\rangle$ The complete extensions of AAF_1 are $\{a\}$ and $\{b\}$, while the complete extension of AAF_2 is $\{a\}$. Thus, AR_1 to AR_4 and SR_1 to SR_4 do not satisfy Principle 3.10 for complete semantics. □

Principle 3.11 reflects the same idea as Principle 3.10, but is based on the assumption that a is accepted in all extensions.

Principle 3.11 (AgentAdditionUniversalPersistence). *An agent semantics δ satisfies AgentAdditionUniveralPerisitence iff for every $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$, for $\forall E \in \delta(AAF)$, $\alpha \in \mathscr{S}$ and $a \in E$, we have $\forall E \in \delta(\langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \cup (a, \alpha)\rangle)$, $a \in E$.*

Principle 3.12 reflects a principle we expect to hold for all agent semantics. It reflects anonymity: if we permute the agents, it does not affect the extensions. It is analogous to language independence for arguments defined by Baroni and Giacomin [35].

Principle 3.12 (PermutationPersistence). *An agent semantics δ satisfies PermutationPersistence iff for every $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and $AAF' = \langle \mathscr{A}, \to, \mathscr{S}', \sqsubset' \rangle$, and where $\mathscr{S}$ and $\mathscr{S}'$ are two different ordered sets with common elements, we have $\delta(AAF) = \delta(AAF')$.*

Principle 3.13 reflects that if the arguments of two agents do not attack each other, we can merge these agents into one single agent. It does not hold for agent defense semantics, because new agent defenses may be created.

Principle 3.13 (MergeAgent). *An agent semantics δ satisfies MergeAgent iff for every $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$, $\exists \alpha, \beta \in \mathscr{S}$, for $\forall a \in \mathscr{A}_\alpha$ and $\forall b \in \mathscr{A}_\beta$, a does not attack b, b does not attack a, we have AAF' by changing $\forall\ a \sqsubset \alpha$ to $a \sqsubset \beta$, and $\delta(AAF) = \delta(AAF')$.*

Principle 3.14 reflects that if two agents have the same arguments, we can remove one of these agents without changing the extensions. This represents the opposite of social semantics, where the number of the agents makes a difference.

Sem.	P3.10	P3.11	P3.12	P3.13	P3.14	P3.15	P3.16	P3.17
TR	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	×
Sem$_1$	$\mathbb{S}$	$\mathbb{S}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×	×
Sem$_2$	$\mathbb{S}$	$\mathbb{S}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×	×
SR$_1$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×
SR$_2$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×
SR$_3$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×
SR$_4$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×
AR$_1$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×	×
AR$_2$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×	×
AR$_3$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×	×
AR$_4$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×	×
OR	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$
NBR	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	×	×

Table 3.5: Comparison between the reductions and new agent principles.

Principle 3.14 (RemovalAgentPersistence)**.** *An agent semantics δ satisfies RemovalAgentPersistence iff for every AAF $= \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, for $\mathscr{S}_\alpha = \mathscr{S}_\beta$, we have* $\delta(\langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle) = \delta(\langle \mathscr{A}, \rightarrow, \mathscr{S} \backslash \alpha, \sqsubset \setminus \sqsubset_\alpha \rangle) = \delta(\langle \mathscr{A}, \rightarrow, \mathscr{S} \backslash \beta, \sqsubset \setminus \sqsubset_\beta \rangle)$.

Principle 3.15 is inspired by social agent semantics. It states that for two argumentation frameworks with the same arguments and attacks, if for every argument the number of agents holding that argument is the same, then the extensions are the same.

Principle 3.15 (AgentNumberEquivalence)**.** *An agent semantics δ satisfies AgentNumberEquivalence iff for every AAF $= \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ and an AAF$' = \langle \mathscr{A}, \rightarrow, \mathscr{S}', \sqsubset' \rangle$, for $\forall a \in \mathscr{A}$, $|\mathscr{S}_a| = |\mathscr{S}'_a|$, we have $\delta(AAF) = \delta(AAF')$.*

Principle 3.16 is inspired by agent reduction semantics. It states that if the set of the arguments of an agent is conflict-free, then there is an extension containing those arguments.

Principle 3.16 (Conflict-freeInvolvement)**.** *An agent semantics δ satisfies Conflict-freeInvolvement iff for every AAF $= \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, for $\forall \alpha \in \mathscr{S}, \mathscr{A}_\alpha$ is conflict-free, there is an E, we have $\mathscr{A}_\alpha \subseteq E$.*

Principle 3.17 is inspired by OrphanReduction semantics. It states that if we have arguments that do not belong to any agents, then they can be removed from the framework without affecting the extensions.

Principle 3.17 (RemovalArgumentPersistence)**.** *An agent semantics δ satisfies RemovalArgumentPersistence iff for every AAF $= \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, and for $\nexists \alpha \in \mathscr{S}$ and $a \sqsubset \alpha$, we have* $\delta(\langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle) = \delta(\langle \mathscr{A} \backslash a, \rightarrow \setminus \rightarrow_a), \mathscr{S}, \sqsubset \rangle)$.

In the resulting Table 3.5, all agent semantics satisfy P 3.12. Perhaps surprisingly, both social agent semantics and agent reduction semantics does not satisfy P 3.10, while trivial reduction semantics, social agent semantics and agent filtering semantics satisfy

P 3.13. Moreover, all agent semantics except the social agent semantics satisfy P 3.14. No semantics satisfies P 3.16. As expected, only OrphanRemoval satisfies P 3.17. The only semantics that are not distinguished yet concern the use of different preference reductions, or different Dung semantics. To distinguish them, the principles proposed in preference-based argumentation and in Dung's semantics can be used. In that sense, the principle-based analysis in this chapter is complementary to the principle-based analysis in the other areas.

3.10 Related Work

Our work builds on a rich literature on formal argumentation and dialogue, and we can mention here only a few of the most directly related papers.

From the four kinds of agent argumentation semantics introduced in this chapter, we are not aware of other approaches that adapt Dung's basic concepts directly, as we have done with individual and collective agent defense. There are other variants of semantics that adapt these notions, such as weak defense for weak admissibility semantics [39], but that is not based on the agent metaphor.

The most related work is in social agent semantics. Leite and Martins [190] introduce an abstract model of argumentation where agents can vote in favor of and against an issue. They define an abstract argumentation framework as a triple $\langle \mathscr{A}, \mathscr{R}, \mathscr{V} \rangle$, where $\mathscr{V} \rightarrow N \times N$ is a total function mapping each argument to its number of positive (Pro) and negative (Con) votes. Our work, on the other hand, only considers positive votes. Caminada and Pigozzi [73] capture the notion that individual members need to defend the collective decision in order to reach a compatible outcome, and propose to address judgment aggregation by combining different individual evaluations of the situation represented by an argumentation framework. Hunter, Polberg, and Thimm [163] take an epistemic approach to probabilistic argumentation, where the arguments are believed or not believed in terms of different degrees, providing an alternative to the subtle standard Dung framework.

Concerning agent reduction semantics, several authors build on the local functions introduced by Baroni, Gia-comin, and Guida [36], and further developed by Baroni et al. [32]. Giacomin [143] shows how to use this theory in multi-agent systems. The results in these papers indicate that such generalizations often become equivalent to Dung; and Arisaka, Satoh, and van der Torre [19] extend the agent argumentation frameworks with coalitions among the agents. Rienstra et al. [276] consider the case where the agents may have different semantics, for example one agent uses grounded semantics and another agent uses preferred semantics. Furthermore, Kontarinis and Toni [181] analyse the identification of the malicious behavior of agents in the form of bipolar argumentation frameworks, which together with the work of Panisson et al. [241] may inspire work on agent reduction semantics based on trustfulness.

In this chapter, we build on the principle-based approach to preference-based argumentation developed by Amgoud and Cayrol [8] together with several co-authors over the past fifteen years. In particular, the work of Amgoud and Vesic [10] and the work of Kaci, van der Torre, and Villata [174] have inspired us, although the principles discussed here are mostly different from those studied in preference-based argumentation. In earlier work, two of the authors have related their axiomatic approach to the analysis of bipolar argumentation [338], and there are also close relations with the study of robustness principles [278].

3.11 Summary

As common in the principle-based approach to argumentation semantics, we have selected principles that distinguish agent semantics. In addition, we have added some principles that reflect important properties of agent semantics and that can be used to guide the development of future agent semantics. To be more specific, combining ideas from earlier work in abstract argumentation, our principle-based analysis has revealed several original insights. For example, a new twist was given to the fundamental role of defense and reinstatement in Dung's theory in the context of agent defense semantics.

Moreover, a new variant of SCC-recursiveness has been introduced, leading to an SCC-recursive algorithm for agent defense semantics. Since the other approaches are based on reductions, the traditional SCC-recursive algorithms can be used. Finally, the future work section illustrates that our formal framework not only serves as a tool for organizing existing work in the area, but also provides a solid foundation for further work in this direction.

There is quite some variety in agent semantics. In this chapter, the priority and hierarchy of semantics and which kind of semantics is used depends on its application. Also, they can be combined. For example, we can use both filtering and agent defense to remove unknown arguments or unknown attacks and defend an argument put forward by an agent. We can use the principles to elect the most suitable semantics for an application, and in fact, that is one of the most important uses of principles. For example, Principle 3.17 (Removal Argument Persistence) can help us to elect agent filtering semantics for an application.

Prakken [264] distinguishes between argumentation as inference and argumentation as dialogue. Abstract agent argumentation can bring elements of argumentation as dialogue into the foundations of argumentation as inference, and may help to bridge the gap between the two branches.

Within the formal setting we have adopted in this chapter, many topics of further work present themselves. As always with the principle-based approach, we can introduce more semantics, for example by combining the ideas of the four classes, guided by the existing principles. We can also study more principles. We can find relations with other branches of logic and reasoning such as axiomatic approaches in social choice. Moreover, we can try to use principles to address the standard challenges of abstract argumentation, namely relating the abstract model to more structured forms of argumentation, and applying abstract argumentation, for instance, to legal reasoning.

For example, there is no agent semantics satisfying principle P 3.16. We can define new semantics as a variant of Definition 3.14, which combines agent reductions in a new way. Instead of combining the frameworks, we can take the union of all the extensions to the individual frameworks: $\delta(AAF) = \sigma(\bigcup_{\alpha \in \mathscr{S}} PRi(AAP(AAF, \alpha)))$ for $i \in \{1, 2, 3, 4\}$. $\delta'(AAF) = \bigcup_{\alpha \in \mathscr{S}} \sigma(PRi(AAP(AAF, \alpha)))$ for $i \in \{1, 2, 3, 4\}$. If we use the definition of δ', Table 3.2 changes as follows.

A regular topic in abstract argumentation is to search for fragments with good computational properties, such as symmetric attack relations. Also with agent argumentation, we can study frameworks where: every argument is associated with at least one agent, every

Sem.	$\mathbb{C}$	$\mathbb{G}$	$\mathbb{P}$	$\mathbb{S}$
AR$_1$	$\{\{a\},\{a,b\}\}$	$\{\{a\},\{a,b\}\}$	$\{\{a\},\{a,b\}\}$	$\{\{a\},\{a,b\}\}$
AR$_2$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$
AR$_3$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
AR$_4$	$\{\varnothing,\{a\},\{b\}\}$	$\{\{a\},\varnothing\}$	$\{\{a\},\{b\}\}$	$\{\{a\},\{b\}\}$

argument is associated with at most one agent, there are at most two agents, a symmetric attack is possible, the arguments of each agent are conflict free, and so on.

Moreover, concerning the use of reductions in abstract argumentation, our work raises the question of whether we can find a reduction to Dung's argumentation frameworks for agent defense semantics. While we have presented such reductions for all the other kinds of agent semantics, we have not yet found such a reduction for agent defense semantics. For such a reduction, we might also add auxiliary arguments, or we may introduce arguments for each pair of arguments and agent.

Finally, one of the main challenges in the area of formal argumentation is the gap between abstract argumentation and dialogue. Caminada [68] presents semantics of abstract argumentation that can be interpreted with regard to structured discussion in order to fill this gap. However, how to implement abstract argumentation with dynamic agent dialogue is still an open question.

Chapter 4

A Principle-based Analysis to Bipolar argumentation

In this chapter, we study the acceptance of arguments in bipolar argumentation. Gordon's requirements analysis for formal argumentation to balance pro and con arguments seems can be represented more easily in so-called bipolar argumentation frameworks containing besides attack also a support relation among arguments at an abstract level. In this chapter, we introduce and study nine types of semantics for bipolar argumentation frameworks, each extending Dung's interpretation of attack with a distinct interpretation of support. First, we introduce five types of defence-based semantics by adapting the notions of conflict-freeness or defence. Second, we introduce two types of selection-based semantics that selects extensions by counting the number of supports. Third, we analyse two types of traditional reduction-based semantics under deductive and necessary interpretations of support. In total, we provide a full analysis of thirty-six bipolar argumentation semantics and sixteen principles.

4.1 Introduction

In this chapter, we consider so-called bipolar argumentation frameworks [79, 80, 81, 82, 83] containing not only attacks but also supports among arguments. While there is general agreement in the formal argumentation literature on how to interpret attack, even when different kinds of semantics have been defined, there is much less consensus on how to interpret support [92]. There exist very few results and studies about the role of support in abstract argumentation. Consequently, the principle-based approach is used to bring structure to the field [84, 338]. In this paper, we address the following research questions: In which ways can support affect attack, defence and argumentation semantics? Which principles can be introduced to distinguish between, and characterise, these semantics?

There exist different approaches to extending Dung's abstract theory by taking into consideration the support relation. The relation between support and attack has been studied extensively in reduction-based approaches, in the sense that deductive and necessary interpretations of support give rise to various notions of indirect attack [84], thus, they typically give opposite results. Deductive support [57] captures the intuition that if a supports b, then the acceptance of a implies the acceptance of b. This intuition is characterised by the so-called closure principle [84]. Necessary support [228] captures the intuition that if a

supports b, then the acceptance of a is necessary to obtain the acceptance of b, or equivalently, the acceptance of b implies the acceptance of a. xsIt has been characterised by the inverse closure principle [250]. Another approach to handling support is the evidence-based approach [235] where the notion of evidential support is introduced. An argument cannot stand unless it is supported by evidential support. Support can also be seen as an inference relation between the premises and the conclusion of the argument itself [261]. Moreover, in selection-based approaches [140], support is used only to select some of the extensions provided in Dung's semantics, and thus does not change the definition of attack, or defence.

Despite the relevance and significance of all the mentioned approaches, we think that there is still the need to explore other approaches that have not been yet considered for bipolar argumentation frameworks. The aim of our research is not to replace other approaches but rather to point out to the existence of other interesting ones that can be applied depending on the chosen application. Note that our approach is novel in its methodology. On one hand, reduction-based approaches can be seen as a kind of pre-processing step for Dung's theory of abstract argumentation (i.e. adding the complex attacks and then applying Dung's semantics). On the other hand, selection-based approaches can be seen as a kind of post-processing step (i.e. applying Dung's semantics and then applying the approach to select some of the extensions). Differently from those two groups of approaches, our approach (i.e. the defence-based approach) does not affect the concept of attack and conflict-freeness, but rather changes the definition of defence.

Most of the principles we introduce and use for analysing bipolar argumentation are in the same spirit as the principles used in the principle-based analysis of Dung's semantics [315]. For example, the robustness of argumentation semantics when adding or removing attacks plays a central role [278]. In this paper, we consider robustness when adding or removing support relations. We also introduce some principles specifically defined for support, such as to which extent an argument is accepted while receiving support from others.

The layout of this chapter is as follows. We first introduce three defence-based semantics, then two selection-based ones, and we study two traditional reduction-based ones. Then, we introduce ten principles, and we analyse which properties are satisfied by which semantics, before concluding and introducing the ideas for future work.

4.2 Bipolar Argumentation Framework

Bipolar argumentation frameworks extend the argumentation frameworks introduced by Dung [110] with a binary support relation among the arguments.

Definition 4.1 (Bipolar Argumentation Framework [83]). *Given a set U called the universe of arguments, a bipolar argumentation framework (BAF) is a triple $\langle Ar, att, sup\rangle$ where $Ar \subseteq U$ is a finite set called arguments, and $att, sup \subseteq Ar \times Ar$ are binary relations over Ar called attack and support respectively .*

Figure 4.1 illustrates four BAFs, where attack relations are depicted by solid arrows, support relations are depicted by dashed arrows. Given a, b in Ar, $(a, b) \in att$ standing for a attacks b, and $(a, b) \in sup$ standing for a supports b, the definitions of conflict-freeness and defence provided by Dung are called conflict-free$_0$ and defended$_0$ in this chapter.

Definition 4.2 (Conflict-free$_0$ and Defended$_0$ [110]). *Let $\mathscr{F} = \langle Ar, att, sup \rangle$ be a BAF. A set of arguments $E \subseteq Ar$ is conflict-free$_0$, written as cf$_0(\mathscr{F}, E)$, iff there are no arguments a and b in E such that a attacks b. The set of arguments defended$_0$ by E, written as $d_0(\mathscr{F}, E)$, is the set of a arguments such that for all arguments b attacking a, there is an argument c in E attacking b.*

Example 4.1 (Conflict-free$_0$ and Defended$_0$). *Consider the bipolar argumentation framework in Figure 4.2. The conflict-free$_0$ sets are $\varnothing$, $\{a\}$, $\{b\}$, $\{c\}$, $\{d\}$, $\{e\}$, and $\{b, d\}$, etc. We have $d_0(\mathscr{F}, \{a\}) = \{a\}$, $d_0(\mathscr{F}, \{b\}) = \{b, d\}$.*

4.3 Defence-based Semantics

We first define five new types of defence-based semantics for bipolar argumentation frameworks. The previous four are based on conflict-free$_0$ and the new definitions of defended$_1$, defended$_2$, etc. The fifth is based on a new notion of conflict-free and a new notion of defended. To have a generic definition of defence-based semantics (Definition 4.5), we also define conflict-free$_1$, conflict-free$_2$, etc., for each of the new type of semantics. The first three notions of defended have stronger requirements than defended$_0$. Defended$_1$ requires that the argument defending$_0$ another argument also supports it. Defended$_2$ requires that a defender is supported. Moreover, defended$_3$ requires not only that the attackers are attacked, but also that all supporters of the attackers are attacked as well. Finally, defended$_4$ has a weaker requirement than defended$_0$, as it interprets support as a sufficient condition for defence.

Definition 4.3 (Conflict-free$_{1-4}$ and Defended$_{1-4}$). *Let $\mathscr{F} = \langle Ar, att, sup \rangle$ be a BAF. We use the same definition as Dung for conflict-free, i.e. $cf_1 \equiv \ldots \equiv cf_4 \equiv cf_0$. Moreover:*

- *the set of arguments defended$_1$ by E, written as $d_1(\mathscr{F}, E)$, is the set of arguments a in Ar such that for each argument b in Ar attacking a, there exists an argument c in E attacking b and supporting a (supporting-defence, see Figure 4.1.1);*

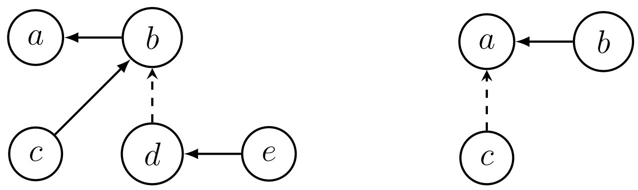
- *the set of arguments defended$_2$ by E, written as $d_2(\mathscr{F}, E)$, is the set of arguments a in Ar such that for all arguments b in Ar attacking a, there exists an argument c in E attacking b, and there is an argument d in E supporting c (supported-defence, see Figure 4.1.2);*

- *the set of arguments defended$_3$ by E, written as $d_3(\mathscr{F}, E)$, is the set of arguments a in Ar such that for all arguments b in Ar attacking a, there exists an argument c in E attacking b, and for all arguments d in Ar supporting b, there is an argument e in E attacking d (attacking-defence, see Figure 4.1.3);*

- *the set of arguments defended$_4$ by E, written as $d_4(\mathscr{F}, E)$ is the set of arguments a in Ar such that for all arguments b in Ar attacking a, there exists an argument c in E attacking b or supporting a (support-as-defence, see Figure 4.1.4).*

The following example illustrates the difference between different defended.

1.1 $d_1(\mathscr{F}, \{c\}) = \{a, c\}$ 1.2 $d_2(\mathscr{F}, \{c, d\}) = \{a, c, d\}$



1.3 $d_3(\mathscr{F}, \{c, e\}) = \{a, c, e\}$ 1.4 $d_4(\mathscr{F}, \{c\}) = \{a, b, c\}$

Figure 4.1: Four notions of defence

Example 4.2 (Defended$_{0-3}$). *Consider the bipolar argumentation framework visualized in Figure 4.2.*
(Defended$_0$) $\{a\}$ defends$_0$ a as it attacks the only attacker of a, similarly, $\{b, d\}$ defends$_0$ b, d.
(Defended$_1$) $\{a\}$ does not defend$_1$ a, although it attacks the only attacker of a, a is not supported by itself, $\{b, d\}$ does not defend$_1$ b, d, as b does not support d.
(Defended$_2$) $\{a\}$ does not defend$_2$ a, since a is not supported by itself, $\{b, d\}$ defends$_2$ b, d, as b is supported by d.
(Defended$_3$) $\{a\}$ does not defend$_3$ a, since a does not attack e and d, $\{b, d\}$ defends$_3$ b, d, as b attacks a, which supports c.

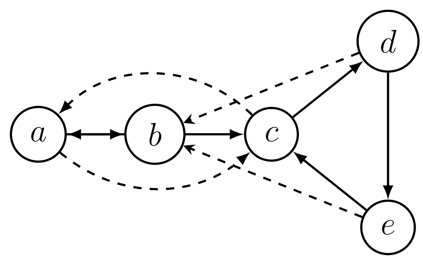


Figure 4.2: $\{a\}$ defends$_0$ a, but not defends$_{1,2,3}$ a

Following Dung's approach to characterise argumentation by a fix point theory, we say the characteristic function $d_i(\mathscr{F}, E)$ of a bipolar argumentation framework BAF is as follows:

- $d_i(\mathscr{F}, E) : 2^{Ar} \to 2^{Ar}$,
- $d_i(\mathscr{F}, E) = \{A \mid A$ is defended$_i$ by $E\}$, for $i \in \{0, 1, 2, 3\}$.

Definition 4.4 (Admissibility$_{0-3}$). *A set of arguments E in BAF $\mathscr{F} = \langle Ar, att, sup \rangle$, is said to be admissible$_i$ iff E is conflict-free$_i$ and $E \subseteq d_i(\mathscr{F}, E)$, for $i \in \{0, 1, 2, 3\}$.*

To define the complete (abbreviated as c), preferred (p) and stable (s) semantics of bipolar argumentation frameworks, the following definition is generic and can be used with any kind of conflict-freeness and defence.

Definition 4.5 (Semantics$_{0-3}$). *An extension-based semantics σ is a function that maps a BAF $\mathscr{F} = \langle Ar, att, sup \rangle$ onto a set of subsets of Ar, written as $\sigma_i^x(\mathscr{F})$, where $i \in \{0,1,2,3\}$, $x \in \{c,p,s\}$ as follows:*

- $\sigma_i^c(\mathscr{F}) = \{E \subseteq Ar \mid cf_i(\mathscr{F}, E),\ d_i(\mathscr{F}, E) = E\}$;
- $\sigma_i^p(\mathscr{F}) = \{E \subseteq Ar \mid$ *for all admissible$_i$ set* $E', E \nsubseteq E'\}$;
- $\sigma_i^s(\mathscr{F}) = \{E \subseteq Ar \mid E$ *is admissible$_i$, and for all arguments a not in E, there is an argument b in E attacking a*$\}$.
- $\sigma_i^g(\mathscr{F}) = \{E \subseteq Ar \mid E$ *is the least fix point of* $d_i(\mathscr{F}, E)\}$.

Most of the following propositions were introduced and proved for semantics$_0$ by Dung (1995). We prove that the above three new defence semantics are able to conserve the relations among complete$_i$, preferred$_i$ and grounded$_i$ for $i \in \{1,2,3\}$ and stable$_i$ for $i = 3$.

Lemma 4.1 (Fundamental Lemma). *Let E be an admissible$_i$ set of arguments, and A_1 and A_2 be two arguments which are defended$_i$ by E. Then for $i \in \{0,1,2,3\}$, we have the following:*

- *$E' = E \cup \{A_1\}$ is admissible$_i$.*
- *A_2 is defended$_i$ by E'.*

Proof.
- We will prove that $E' = E \cup \{A_1\}$ is admissible$_i$, for $i \in \{0,1,2,3\}$. We only need to prove that E' is conflict-free$_i$. We will do it by contradiction. Suppose that there exists an argument $B_1 \in E$ such that either B_1 attacks A_1 or A_1 attacks B_1.
  – If B_1 attacks A_1: Since E defends$_i$ A_1, there exists an argument $B_2 \in E$ such that B_2 attacks B_1, but E is admissible$_i$ so E is conflict-free$_i$. Contradiction
  – If A_1 attacks B_1: There exists an argument $B_2 \in E$ such that B_2 attacks A_1, but E defends$_i$ A_1, so there exists an argument $B_3 \in E$ such that B_3 attacks B_2. But E is conflict-free$_i$. Contradiction.
- It's obvious now that A_2 is defended$_i$ by E'.

□

The following theorem follows directly from the Fundamental Lemma.

Theorem 4.1. *Let BAF be a bipolar argumentation framework, for $i \in \{0,1,2,3\}$:*

- *The set of all admissible$_i$ sets of BAF forms a complete partial order with respect to set inclusion.*
- *For each admissible$_i$ set S of BAF, there exists a preferred$_i$ extension E of BAF such that $S \subseteq E$.*

Note that the empty set is always admissible$_i$, we have the following Corollary for $i \in \{0, 1, 2, 3\}$:

Corollary 1. *There exists at least one preferred$_i$ extension in any bipolar argumentation framework for $i \in \{0, 1, 2, 3\}$.*

Proposition 4.1. *For $i \in \{0, 1, 2, 3\}$, we have the following:*

- *Every complete$_i$ extension is also admissible$_i$.*
- *Every preferred$_i$ extension is also complete$_i$.*
- *Every stable$_i$ extension is also preferred$_i$.*

Proof.

- **Every complete$_i$ extension is also admissible$_i$:** Obvious
- **Every preferred$_i$ extension is also complete$_i$:** We know that every preferred$_i$ extension E in BAF is admissible$_i$, we just need to prove that $d_i(\mathscr{F}, E) \subseteq E$. For any argument $a \in d_i(\mathscr{F}, E)$, from the Fundamental Lemma 4.1, we have that $E' = E \cup \{a\}$ is also admissible$_i$, and $E \subseteq E'$. Since E is a preferred$_i$ extension of BAF, then E is a maximal admissible$_i$ set (with respect to set inclusion) in BAF. So having $E \subseteq E'$, means that $E = E'$. So $a \in E$, and it follows that $d_i(\mathscr{F}, E) \subseteq E$. So E is a complete$_i$ extension of BAF.
- **Every stable$_i$ extension is also preferred$_i$:** Let E be a stable$_i$ extension in BAF, we know that E is admissible$_i$, we need to prove that E is a maximal admissible$_i$ set (with respect to set inclusion) in BAF. We do it by contradiction. Suppose that there exists an admissible$_i$ set E' such that $E \subseteq E'$. Since E is stable$_i$ then E attacks all of its outside, so E' is not conflict-free$_i$. But E' is admissible$_i$. Contradiction.

□

Proposition 4.2. *The characteristic function $d_i(\mathscr{F}, E)$ is monotonic (with respect to set inclusion) for $i \in \{0, 1, 2, 3\}$.*

Proposition 4.3. *The characteristic function $d_i(\mathscr{F}, E)$ is monotonic (with respect to set inclusion) for $i \in \{0, 1, 2, 3\}$.*

Proof. Let BAF $\mathscr{F} = \langle Ar, att, sup \rangle$ be a bipolar argumentation framework, and X_1, X_2 be two sets of arguments such that $X_1 \subseteq X_2$. Then it follows that $d_i(\mathscr{F}, X_1) \subseteq d_i(\mathscr{F}, X_2)$, for $i \in \{0, 1, 2, 3\}$.

- If $X_1 \subseteq X_2$, then $d_0(\mathscr{F}, X_1) \subseteq d_0(\mathscr{F}, X_2)$. This is because for any argument $z \in d_0(\mathscr{F}, X_1)$, z is defended$_0$ by X_1, so for any argument $y \in Ar$ such that y attacks z, there exists $x \in X_1$ such that x attacks y. $X_1 \subseteq X_2$, so $x \in X_2$, so z is defended$_0$ by X_2, $z \in d_0(\mathscr{F}, X_2)$.
- If $X_1 \subseteq X_2$, then $d_1(\mathscr{F}, X_1) \subseteq d_1(\mathscr{F}, X_2)$. This is because for any argument $z \in d_1(\mathscr{F}, X_1)$, z is defended$_1$ by X_1, so for any argument $y \in Ar$ such that y attacks z, there exists $x \in X_1$ such that x attacks y and x supports z. $X_1 \subseteq X_2$, so $x \in X_2$, so z is defended$_1$ by X_2, $z \in d_1(\mathscr{F}, X_2)$.

- If $X_1 \subseteq X_2$, then $d_2(\mathscr{F}, X_1) \subseteq d_2(\mathscr{F}, X_2)$. This is because for any argument $z \in d_2(\mathscr{F}, X_1)$, z is defended$_2$ by X_1, so for any argument $y \in Ar$ such that y attacks z, there exists $x, w \in X_1$ such that x attacks y and w supports x. $X_1 \subseteq X_2$, so $x, w \in X_2$, so z is defended$_2$ by X_2, $z \in d_2(\mathscr{F}, X_2)$.

- If $X_1 \subseteq X_2$, then $d_3(\mathscr{F}, X_1) \subseteq d_3(\mathscr{F}, X_2))$. This is because for any argument $z \in d_3(\mathscr{F}, X_1)$, z is defended$_3$ by X_1, so for any argument $y, b \in Ar$ such that y attacks z and b supports y, there exists $x, w \in X_1$ such that x attacks y and w attacks b. $X_1 \subseteq X_2$, so $x, w \in X_2$, so z is defended$_3$ by X_2, $z \in d_3(\mathscr{F}, X_2)$.

□

Proposition 4.4. *The grounded$_i$ extension of BAF for $i \in \{0, 1, 2, 3\}$ is the minimal (with respect to set inclusion) complete$_i$ extension of BAF.*

Proof. The grounded$_i$ extension is the least fixed point of the characteristic function $d_i(\mathscr{F}, E)$ for $i \in \{0, 1, 2, 3\}$. From Corollary 1, we have proved that there exists at least one preferred$_i$ extension in BAF and from Proposition 4.1, every preferred$_i$ extension is complete$_i$, so there exists at least one complete$_i$ extension in BAF, for $i \in \{0, 1, 2, 3\}$. Then, there exists at least one conflict-free$_i$ fixed point. The grounded$_i$ extension is the intersection of all the fixed points, so the grounded$_i$ extension is conflict-free$_i$, so the grounded$_i$ extension is complete$_i$. So the grounded$_i$ extension is the minimal complete$_i$ extension of BAF. □

We now give a real legal example to illustrate the intuition behind semantics$_1$. This example deals with a neighbor's quarrel over a row of conifers and was used to explain how the judge defends the claimant's interest [248].

Example 4.3 (Neighbours' quarrel over conifers)**.** *(...) Defendant argues that the conifers have been planted to reduce draught in his house, but this argument is absolutely unsound since most of the window posts are closed and the window that does open, is located on a point higher than the tops of the conifers and has not been fitted with any anti-draught facilities. (...) Whereas the defendant has no considerable interest in these conifers, removal is of significant concern to the claimant since they block his view and take away the light. (...) (2981. Country court Enschede 6 October 1988)*

The judge defends the standpoint that the claimant's interest in the removal of the conifers is greater than the defendant's interest in leaving them untouched. In the judge's preceding remarks, he mentions the defendant's argument: he does have a considerable interest in the conifers since they reduce draught in his house, thus he wants to keep the conifers. To support the standpoint of the claimant and against the defendant, the judge argues that the conifers block the view and take away the light, most of the window posts are closed and the opening window, which has no anti-draught facilities whatsoever, is located higher than the tops of the conifers.

As stated by Plug: "the judge's argumentation consists of a pro-argument and the refutation of a counter-argument which, in conjunction, form sufficient support for his standpoint." This type of defence inspires semantics$_1$.

We now give an example to illustrate the intuition behind semantics$_2$.

Example 4.4 (Twelve Angry Men play using Semantics$_2$)**.** *We consider an example extracted from the NoDE benchmark [67], which consists of annotated datasets extracted from a variety of sources (Debatepedia, Procon, Wikipedia web pages and the script of "Twelve Angry Men" play), where the aim of this benchmark is to analyse the support and attack relations between the arguments. We explore the Twelve Angry Men dataset, this play is about a jury consisting of twelve men who must decide whether a young man is guilty or not for murdering his father. Consider the following arguments extracted from this dataset.*

- ***Argument id= "55-j8":*** *I think we proved that the old man couldn't have heard the boy say, "I'm going to kill you" but supposing he really did hear it? This phrase: how many times has each of you used it? Probably hundreds. "If you do that once more, Junior, I'm going to murder you." "Come on, Rocky, kill him!" We say it every day. This doesn't mean that we're going to kill someone.*
- ***Argument id= "56-j3":*** *The phrase was "I'm going to kill you" and the kid screamed it out at the top of his lungs. Anybody says a thing like that the way he said it—they mean it.*
- ***Argument id= "58-j8":*** *Do you really think the boy would shout out a ["I'm going to kill you"] so the whole neighbourhood would hear it? I don't think so. He's much too bright for that.*

*The example above is shown in Figure 4.3. In this example, argument **"58-j8"** attacks argument **"56-j3"** by raising some doubt about it. In the same manner, argument **"56-j3"** attacks argument **"55-j8"**. We can see that the argument **"58-j8"** defends argument **"55-j8"** in Dung's sense. Just because argument **"58-j8"** is not attacked, argument **"55-j8"** is accepted.*

*In a legal case, any given argument must be evaluated based on the evidence provided to support it. In the absence of such evidence, the presence of at least a support, even if it is challenged, seems necessary. Therefore, one can ask themselves whether Dung's notion of defence seems enough, in this case, to say that the argument **"58-j8"** defends the argument **"55-j8"**. Hence, for this kind of application, one might want to use a stronger notion of defence. An example of such a notion is our semantics$_2$, where an argument must be supported in order to be able to defend another argument. The idea behind this semantics is to provide a stronger and more restrictive defence notion than Dung's defence notion, by taking into account the support relation.*

We consider now the following arguments extracted from the same dataset, to illustrate semantics$_2$.

- ***Argument id= "48-j2":*** *Maybe he didn't hear [the boy yelling "I'm going to kill you"]. I mean with the el noise.*
- ***Argument id= "49-j3":*** *[The old man cannot be a liar, he must have heard the boy yelling "I'm going to kill you"].*
- ***Argument id= "50-j5":*** *It stands to reason, [the old man can be a liar].*
- ***Argument id= "51-j9":*** *Attention, maybe [the old man is a liar].*

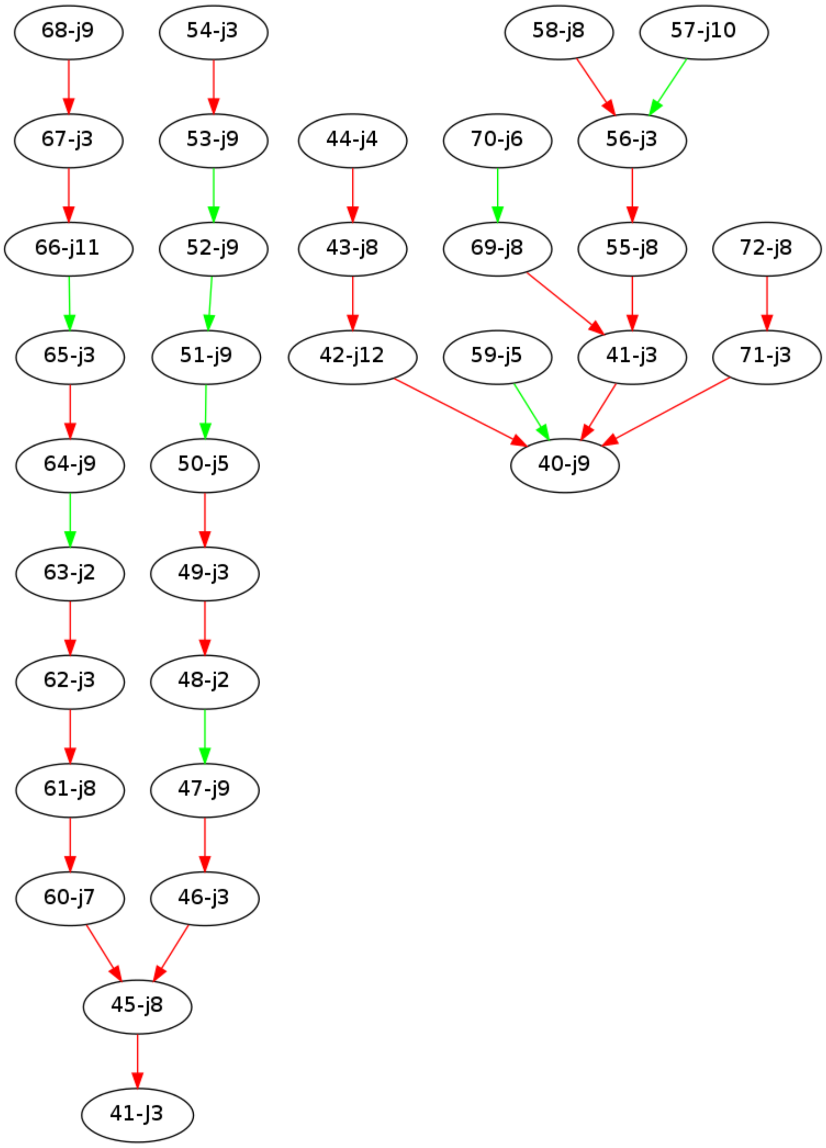


Figure 4.3: The BAF illustrating the Twelve Angry Men dataset - Act 2. The arrows in red represent the relations of attack and the arrows in green represent the relations of support between the arguments.

Contrary to the previous example, we see that argument "51-j9" is supported by another one, hence it might be seen as having a better capacity to defend "48-j2". Formally, the set of arguments ***{"51-j9", "50-j5"}*** *defends*$_2$ *the argument* ***"48-j2"****.*

Example 4.5 (Recruitment using semantics$_3$)**.** *Consider the following arguments.*

- a*: Alice should be hired as a professor.*
- b*: Alice lacks many essential qualifications to become a professor.*
- c*: Alice has few number of publications.*
- d*: Alice has recently got her PhD, she does not have enough teaching experience.*
- e*: All of Alice's publications are in top conferences. When it comes to publications, quality beats quantity.*

- *f: Alice has taught 64 hours of practical works during every year of her PhD, which is considered enough as teaching experience.*
- *g: Alice is good at team work, she also has an excellent academic carrier, these two enable her to become a professor.*

This example can be represented in Figure 4.4. g fails to reinstate a because g does not attack b's supporters c and d. The set of arguments $\{e, g, f\}$ reinstates a because it attacks all the supporters of b. $\sigma_3^{c,g,p,s}(\mathscr{F}) = \{\{a, e, g, f\}\}$.

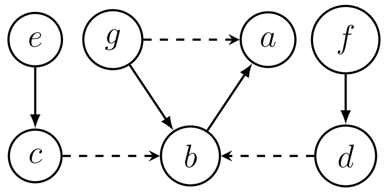


Figure 4.4: A BAF illustrating recruitment case

Defended$_5$ and conflict-free$_5$ compares the number of attacks and supports an argument has. Intuitively, an attack can be neutralised by a support [6, 136]. We consider a set to be conflict-free$_5$ if each attacked argument receives at least as many supports as attacks.

Definition 4.6 (Conflict-Free$_5$ and Defended$_5$)**.** *Let $\mathscr{F} = \langle Ar, att, sup\rangle$ be a BAF. Argument a in Ar is active w.r.t. a set E iff $|\ \{b \in E \mid (b, a) \in att\}\ | \leqslant |\ \{b \in E \mid (b, a) \in sup\}|$. Then, E is conflict-free$_5$, written as $cf_5(\mathscr{F}, E)$ iff all the arguments a in E are active w.r.t. E. Let $Batt(a, E) = \{b \in Ar \mid b$ attacks a and b is an active argument w.r.t. $E\}$. Let $Bsup(a, E) = \{b \in Ar \mid b$ supports a and b is an active argument w.r.t. $E\}$. The set of arguments defended$_5$ by E, written as $d_5(\mathscr{F}, E)$, is the set of arguments a such that $|\ Batt(a, E)\ | \leqslant |\ Bsup(a, E)\ |$ (active-supporters defence).*

Another defence-based approach is to interpret support as compensation for an attack, as illustrated by the following example.

Example 4.6 (Conflict-Free$_5$ and Defended$_5$)**.** *Consider the BAF in Figure 4.2 with both red and blue support relation. Let $S = \{a, c\}$. All the arguments in S_1 are active w.r.t S_1 because they receive as many supports as attacks.*

4.4 Selection-based Semantics

Support can be used in the post-processing step for Dung's theory of abstract argumentation [140]. Semantics$_6$ and semantics$_7$ are two selection-based approaches, i.e. they select extensions from semantics$_0$. Semantics$_6$ selects the extensions that have the largest number of internal supports, reflecting the idea that for a coalition, the more internal supports they have, the more cohesive they are. Semantics$_7$ selects the extensions that receive the most support from outside, reflecting the idea that the more support a coalition receives, the stronger it is.

We say that argument b in E is internally supported if b receives support from arguments in E. Argument b in E is externally supported if b receives support from arguments that are outside E.

Definition 4.7 (Number of Internal and External Supports). *Let $\mathscr{F} = \langle Ar, att, sup\rangle$ be a BAF. For an extension $E \subseteq \sigma_0^x$, the number of internal supports is written as NS_I, such that $NS_I(\mathscr{F}, E) = |a \text{ supports } b \mid a, b \in E\}|$, and the number of external supports is written as NS_O, such that $NS_O(\mathscr{F}, E) = |a \text{ supports } b \mid b \in E, a \in Ar \smallsetminus E\}|$.*

Definition 4.8 (Semantics$_{6-7}$). *For each $\mathscr{F} = \langle Ar, att, sup\rangle$, for $x \in \{c, g, p, s\}$:*

- $\sigma_6^x(\mathscr{F}) = \arg\max_{E \in \sigma_0^x(\mathscr{F})} \{NS_I(\mathscr{F}, E)\}$*; and*
- $\sigma_7^x(\mathscr{F}) = \arg\max_{E \in \sigma_0^x(\mathscr{F})} \{NS_O(\mathscr{F}, E)\}$.

We use Example 4.7 to illustrate the difference between semantics$_6$ and semantics$_7$.

Example 4.7 (Semantics$_{6-7}$). *Consider the bipolar argumentation framework in Figure 4.2. $\sigma_6^{cps}(\mathscr{F}) = \sigma_7^{cps} = \{\{b, d\}\}$, because $\{b, d\}$ has most internal support and receives the most external support.*

4.5 Reduction-based Semantics

Reduction-based approaches have been studied exclusively in the literature [80, 81, 83]. Semantics$_8$ and semantics$_9$ are two reduction-based approaches where support is used as pre-processing for Dung semantics. The corresponding abstract argumentation frameworks are reduced by adding indirect attacks from the interaction of attack and support with different interpretations, i.e. deductive support and necessary support. So-called supported attack and mediated attack come from the interplay between attack and deductive support, while secondary attack and extended attack come from the interplay between attack and necessary support.

Definition 4.9. *(Four Indirect Attacks [83]) Let $\mathscr{F} = \langle Ar, att, sup\rangle$ be a BAF, and let arguments $a, b, c \in Ar$. There is:*

- *a supported attack from a to b in $\mathscr{F}$ iff there exists an argument c such that there is a sequence of supports from a to c and c attacks b, represented as $(a, b) \in att^{supp}$;*
- *a mediated attack from a to b in $\mathscr{F}$ iff there exists an argument c such that there is a sequence of supports from b to c and a attacks c, represented as $(a, b) \in att^{med}$;*
- *a super-mediated attack from a to b in $\mathscr{F}$ iff there exists an argument c such that there is a sequence of supports from b to c and a directly or supported-attacks c, represented as $(a, b) \in att^{med}_{att^{supp}}$;*
- *a secondary attack from a to b in $\mathscr{F}$ iff there exists an argument c such that there is a sequence of supports from c to b and a attacks c, so that $(a, b) \in att^{sec}$;*
- *an extended attack from a to b in $\mathscr{F}$ iff there exists an argument c such that there is a sequence of supports from c to a and c attacks b, so that $(a, b) \in att^{ext}$.*

Definition 4.10 (Semantics$_{8-9}$ [83]). *Let $\mathscr{F} = \langle Ar, att, sup\rangle$ be a BAF:*

- *let $att' = \{att^{supp}, att^{med}_{att^{supp}}\}$ be the collection of supported and super-mediated attacks in $\mathscr{F}$, and we have $RD(\mathscr{F}) = (Ar, att \cup \bigcup att')$, and $\sigma_8^x(\mathscr{F}) = \sigma_0^x(RD(\mathscr{F}))$;*
- *let $att' = \{att^{sec}, att^{ext}\}$ be the collection of secondary and extended attacks in $\mathscr{F}$, and we have $RN(\mathscr{F}) = (Ar, att \cup \bigcup att')$, and $\sigma_9^x(\mathscr{F}) = \sigma_0^x(RN(\mathscr{F}))$.*

4.6 Principle-based Analysis

In this section, we present ten principles. Due to the space limitation, we only present some interesting proofs, others can be found in additional supplement.

The first principle concerns support relation alone. It expresses transitivity of support.

Principle 4.1 (Transitivity). *A semantics σ_i^x for $BAFs$ satisfies transitivity principle iff for all $BAFs$ $\mathscr{F} = \langle Ar, att, sup \rangle$, if a supports b, and b supports c, then $\sigma_i^x \langle Ar, att, sup \rangle = \sigma_i^x \langle Ar, att, sup \cup \{a, c\} \rangle$.*

Principle 4.2 states that supports can be used to select extensions.

Principle 4.2 (Extension Selection). *A semantics σ_i^x for $BAFs$ satisfies the extension selection principle iff for all BAFs where $\mathscr{F} = \langle Ar, att, sup \rangle$, that $\sigma_i^x(Ar, att, sup) \subseteq \sigma_i^x(Ar, att, \varnothing)$.*

Principle 4.3 and Principle 4.4 are robustness principles that distinguish between semantics$_4$ and semantics$_5$. The set of robustness principles were proposed by Rienstra et al. [277]. Here, we adapt their idea to bipolar argumentation in order to investigate the robustness of bipolar argumentation semantics when removing and adding support. Principle 4.3 states that if two arguments a and b are in an extension E such that a supports b, then E is still an extension after we remove the support from a to b.

Principle 4.3 (Internal Support Removal Robustness). *A semantics σ_i^x for BAFs satisfies the internal support removal robustness principle iff for all BAFs $\mathscr{F} = \langle Ar, att, sup \rangle$, for every extension $E \in \sigma_i^x(\mathscr{F})$, if arguments $a, b \in E$ and a supports b, then $E \in \sigma_i^x(Ar, att, sup \setminus \{(a, b)\})$.*

Proposition 4.5. *σ_4^x do not satisfy Principle 4.3.*

Proof. We use a counterexample from Figure 10.30. We have the initial BAF $\mathscr{F}$ on the left, $\sigma_4^x(\mathscr{F}) = \{\{a.b.d\}\}$. However, if we remove the support from b to a and the support from b to d, we have a new BAF $\mathscr{F}'$ on the right, which is $\sigma_4^x(\mathscr{F}') = \{\{b, c\}\}$. □

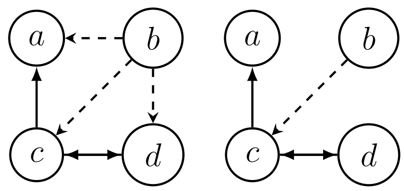


Figure 4.5: σ_4^x do not satisfy Principle 4.3.

Principle 4.4 states that if argument a is not in an extension E and argument b is in this extension E such that a supports b, then E is still an extension after we remove the support from a to b.

Principle 4.4 (External Support Removal Robustness). *A semantics σ_i^x for BAFs satisfies the external support removal robustness principle iff for all BAFs $\mathscr{F} = \langle Ar, att, sup \rangle$, for every extension $E \in \sigma_i^x(\mathscr{F})$, if argument $a \in Ar \setminus E$ supports argument $b \in E$, then $E \in \sigma_i^x(Ar, att, sup \setminus \{(a, b)\})$.*

Principle 4.5 and Principle 4.6 both concern the closure under support relation. Closure says that if an argument is in an extension, the arguments it supports are also in the extension, while inverse closure says the opposite, i.e. if an argument is in an extension, the arguments supporting it should also be in the extension [57, 83, 250].

Principle 4.5 (Closure). *A semantics σ_i^x for $BAFs$ satisfies the closure principle iff for all $BAFs$ $\mathscr{F} = \langle Ar, att, sup \rangle$, for every extension $E \in \sigma_i^x(\mathscr{F})$, if $(a, b) \in sup$ and $a \in E$, then $b \in E$.*

Principle 4.6 (Inverse Closure). *A semantics σ_i^x for $BAFs$ satisfies the inverse closure principle iff for all $BAFs$ $\mathscr{F} = \langle Ar, att, sup \rangle$, for every extension $E \in \sigma_i^x(\mathscr{F})$, if $(a, b) \in sup$ and $b \in E$, then $a \in E$.*

Principle 4.7 reflects the idea that if there is no support relation, the extensions under semantics σ_i^x are equivalent to the ones in Dung semantics.

Principle 4.7 (Extension Equivalence). *A semantics σ_i^x for $BAFs$ satisfies the extension equivalence principle iff for all $BAFs$ $\mathscr{F} = \langle Ar, att, sup \rangle$, that $\sigma_i^x(Ar, att, \varnothing) = \sigma_0^x(Ar, att, \varnothing)$.*

Principle 4.8 and Principle 4.9 both state the positive effect of supports on the supported arguments. We first present the definition of the status of arguments as introduced by Baroni and Giacomin [35]. Extension-based semantics classifies arguments into three statuses, namely sceptically accepted, credulously accepted, and rejected.

Definition 4.11. ***(Status of an Argument [35])*** *Let $\mathscr{F} = \langle Ar, att, sup \rangle$ be a BAF. If the set of extensions is empty, all the arguments are declared to be rejected. Otherwise, we say that an argument is: (1) sceptically accepted if it belongs to all extensions; (2) credulously accepted if it is not sceptically accepted and it belongs to at least one extension; (3) rejected if it does not belong to any extension.*

Gargouri et al. [140] write Status$(a, \mathscr{F}) = sk(resp.\ cr, rej)$, and they define the order $\leqslant$ on the set of statuses as expected: $sk > cr > rej$. We denote the set of sceptically accepted (resp. credulously accepted, rejected) arguments of a BAF by $Sk(Ar, att, sup)$ (resp. $Cr(Ar, att, sup)$, $Rej(Ar, att, sup)$. Principle 4.8 states that adding supports to arguments does not change their status into a lower order. Gargouri et al. [140] call this monotony, but we prefer to use a more specific name (i.e. monotony of status) to make it more precise and avoid ambiguity.

Principle 4.8 (Monotony of Status). *A semantics σ_i^x for $BAFs$ satisfies the monotony of status principle iff for all $BAFs$ $\mathscr{F} = \langle Ar, att, sup \rangle$, for every extension $E \in \sigma_i^x(\mathscr{F})$, for all $a, b \in Ar$, we have Status $(a, \langle Ar, att, sup \rangle) \leqslant$ and Status$(a, \langle Ar, att, sup \cup \{(b, a)\} \rangle)$.*

Principle 4.9 shows a skeptically accepted argument stays skeptically when supports are added [174].

Principle 4.9 (Extension Growth). *A semantics σ_i^x for $BAFs$ satisfies the extension growth principle iff for all $BAFs$ $\mathscr{F} = \langle Ar, att, sup \rangle$, for every extension $E \in \sigma_i^x(\mathscr{F})$, it holds that $Sk(Ar, att, sup) \subseteq Sk(Ar, att, sup \cup sup\prime)$.*

Directionality is introduced by Baroni, Giacomin, and Guida [36]. It reflects the idea that we can decompose an argumentation framework into sub-frameworks so that the semantics can be defined locally. For the directionality principle, they first introduce the definition of an unattacked and unsupported set.

Definition 4.12 (Unattacked and unsupported arguments in BAF)**.** *Given an* $BAF = \langle Ar, att, sup\rangle$*, a set* U *is unattacked and unsupported if and only if there exists no* $a \in Ar \backslash U$ *such that* a *attacks* U *or* a *supports* U*. The unattacked and unsupported sets in BAF is denoted* $US(BAF)$ *(*U *for short).*

Principle 4.10 (BAF Directionality)**.** *A BAF semantics* σ *satisfies the BAF directionality principle iff for every BAF, for every* $U \in US(BAF)$*, it holds that* $\sigma\ (BAF_{\downarrow_U}) = \{E \cap U | E \in \sigma(BAF)\}$*, where for* $\mathscr{F} = \langle Ar, att, sup\rangle$*,* $BAF_{\downarrow_U} = (U, \mathscr{R} \cap U \times U, sup \cap U \times U)$ *is a projection, and* $\sigma\ (BAF_{\downarrow_U})$ *are the extensions of the projection.*

Lemma 4.2. *If a set of arguments* $E' \subseteq U$ *is admissible*$_i$ *in* $BAF_{\downarrow_U}$*, then* E' *is admissible*$_i$ *in* BAF*, for* $i \in \{0, 1, 2, 3\}$*.*

Proposition 4.6. σ_1^c *satisfy Principle 4.10.*

Proof.

- **Part 1:** Suppose that $E' = E \cap U$ such that $E \in \sigma_1^c(BAF)$, we need to prove that $E' \in \sigma_1^c(BAF_{\downarrow_U})$.
  - From Lemma 10.15, if E is admissible$_1$ in BAF, then E' is admissible$_1$ in $BAF_{\downarrow_U}$.
  - We need to prove that there does not exist $x \in U \setminus E'$ such that E' defends$_1$ x. We prove this by contradiction. Suppose that E' defends$_1$ arguments in $U \setminus E'$. We have two cases.
    1. Case 1: Suppose that there exists $x \in U \setminus E'$ such that x is not attacked. Then x is defended$_1$ by E which is not possible because E is a complete$_i$ extension in BAF.
    2. Case 2: Suppose that there exists $x, y \in U \setminus E'$, such that y attacks x. Suppose that E' defends$_1$ x, so there exists $z \in E'$ such that z attacks y, and z supports x. Since $z \in E'$, so $z \in E$, so E defends$_1$ an argument outside of E, but $E \in \sigma_1^c(BAF)$. Contradiction.

- **Part 2:** Suppose that $E' \in \sigma_1^c(BAF_{\downarrow_U})$, we need to prove that there exists $E \in \sigma_1^c(BAF)$ such that $E' = E \cap U$.
  For each set of arguments S in BAF, we denote by Def$_i(S)$ the set of arguments defended$_i$ by S. Let $d_1(\mathscr{F}, E') = E_1$, $d_1(\mathscr{F}, E_1) = E_2$, ..., $d_1(\mathscr{F}, E_{i-1}) = E_i$. $(E_1 \cap U) = E'$ because $E' \in \sigma_1^c(BAF_{\downarrow_U})$, which means that E' cannot defend$_1$ arguments in $(U \setminus E')$. Thus, $(E_1 \setminus E') \subseteq (Ar \setminus U)$.
  Also, for each i, $E_i = d_1(\mathscr{F}, E_{i-1})$, E_{i-1} cannot defend$_1$ the arguments in U, so $(E_i \setminus E') \subseteq (Ar \setminus U)$. That is because if U is attacked, it is going to be attacked from the arguments in U, U cannot be defended$_1$ from the arguments outside of U.
  Let us now prove by induction, that for each i, $E_i = d_1(\mathscr{F}, E_{i-1})$ is admissible$_1$ in BAF.

– **Base:** Let us prove that E' is admissible$_1$ in BAF.
From Lemma 10.14, if E' is admissible$_1$ in $BAF_{\downarrow U}$, then E' is admissible$_1$ in BAF.

– **Step:** Let us suppose that it holds that E_i is admissible$_1$ in BAF, and let us prove that $E_{i+1} = d_1(\mathscr{F}, E_i)$ is admissible$_1$ in BAF.

1. We prove that E_{i+1} is conflict-free$_1$.
   - E_i is conflict-free$_1$ because E_i is admissible$_1$ in BAF.
   - E_i cannot attack $(E_{i+1} \setminus E_i)$, because since E_i defends$_1$ E_{i+1}, if an arguments $x \in E_i$ attacks E_{i+1}, there exists an argument $y \in E_i$ attacking x, which is not possible since E_i is conflict-free$_1$.
   - $(E_{i+1} \setminus E_i)$ is conflict-free$_1$ because if an argument $x \in (E_{i+1} \setminus E_i)$ attacks an argument $y \in (E_{i+1} \setminus E_i)$, E_i must attack x, which is not possible because E_i defends$_1$ E_{i+1}.
   - $(E_{i+1} \setminus E_i)$ cannot attack E_i because since E_i is admissible$_1$ in BAF, if $(E_{i+1} \setminus E_i)$ attacks E_i, E_i defends$_1$ itself so E_i must attack $(E_{i+1} \setminus E_i)$ which is not possible.
2. We prove that E_{i+1} is admissible$_1$ in BAF.
   - E_i is admissible$_1$ in BAF.
   - E_i defends$_1$ E_{i+1} in BAF.

Hence, by induction, we conclude that for each i, $E_i = d_1(\mathscr{F}, E_{i-1})$ is admissible$_1$ in BAF.

Having that at each step i, $d_1(\mathscr{F}, E_i)$ is admissible$_1$, also having a finite number of arguments in Ar, we conclude that there is a fix point of the above defined sequence, that is, there exists j such that $d_1(\mathscr{F}, E_j) = d_1(\mathscr{F}, d_1(\mathscr{F}, d_1(\mathscr{F}, d_1(\mathscr{F}, ...d_1(\mathscr{F}, E'))))) = E_j$. Hence, E_j is a complete$_i$ extension in BAF such that $E_j \cap U = E'$.

□

Table 4.1: Comparison of semantics and principles. We refer to the semantics as follows: complete ($\mathbb{C}$), grounded ($\mathbb{G}$), preferred ($\mathbb{P}$) and stable ($\mathbb{S}$). When a principle is never satisfied by a certain reduction for all semantics, we use the × symbol. P1 refers to Principle 1, and the same holds for the others.

	P4.1	P4.2	P4.3	P4.4	P4.5	P4.6	P4.7	P4.8	P4.9	P4.10
σ_0^x	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$
σ_1^x	×	×	×	$\mathbb{CGPS}$	×	×	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$
σ_2^x	×	×	×	$\mathbb{CGPS}$	×	×	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGP}$
σ_3^x	×	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGP}$
σ_4^x	×	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	×	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×
σ_5^x	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	×	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×
σ_6^x	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	×
σ_7^x	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	$\mathbb{CGPS}$	$\mathbb{CGPS}$	×	×	$\mathbb{CGP}$

Table 4.1 compares the semantics with respect to the principles. All the proofs can be found in the appendix. For the defence-based semantics, semantics$_1$ and semantics$_2$ can be classified by the same principles, and they can be distinguished from semantics$_3$ by Principle 4.3, 4.7 and 4.9. Semantics$_4$ and semantics$_5$ are selected from semantics$_0$, they can be distinguished by Principle 4.3 and Principle 4.4. However, Table 4.1 indicates it is not the case that if semantics$_0$ satisfies a principle implies semantics$_4$ and semantics$_5$ also satisfy it, e.g. the results regarding Principle 4.10. Reduction-based semantics can be distinguished from others by Principle 4.1, 4.5, 4.6 and 4.8. More precisely, themselves can be further distinguished by Principle 4.5 and 4.6, and surprisingly, only semantics$_7$ satisfies Principle 4.8. One thing worth being noticed is that, in the literature, there are two other reductions based on necessary interpretation of support, i.e. one introduces only secondary attacks and the other introduces only extended attacks. Both of them do not satisfy directionality [338]. However, the result shows when the necessary reduction induces both secondary and extended attacks, semantics$_7$ (except for stable$_7$) satisfy directionality.

4.7 Legal Child Custody Case Modeling

We now use a custody case to illustrate reduction-based semantics. Consider the bipolar framework visualized below. The figure should be read as follows. A normal arrow visualizes attack, a dashed arrow visualizes support, a double box visualizes a prima facie argument which is self-supporting and single box visualizes a standard argument that does not support itself.

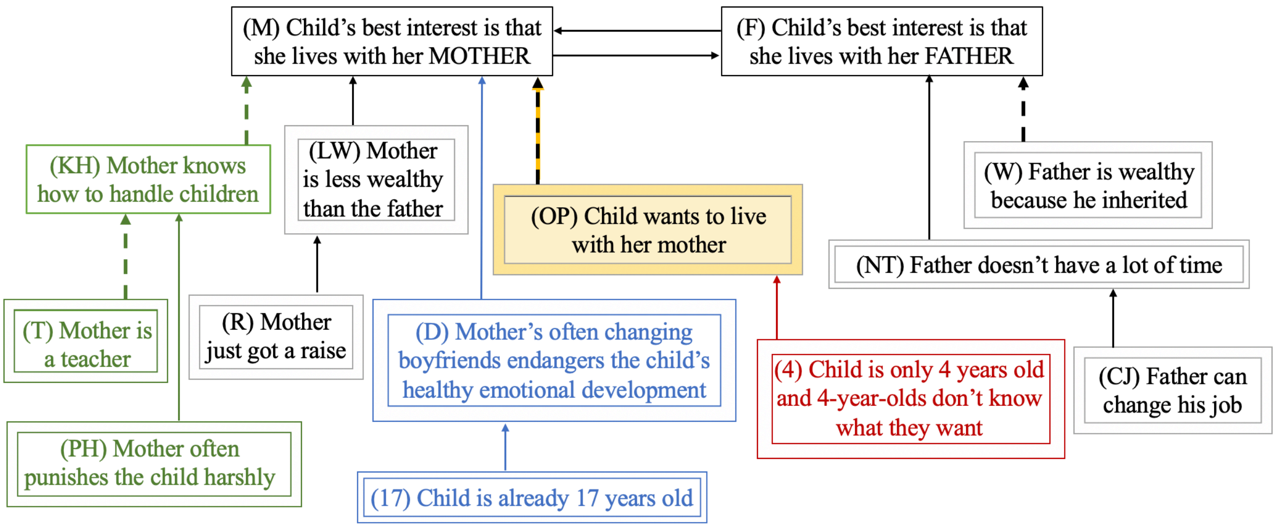


Figure 4.6: Divorce action

We first consider the graph with only the arguments in black and orange and the relations among them. The basic dilemma is represented by two arguments attacking each other, stating respectively that the child's best interest is that she lives with her mother (M) or that the child's best interest is that she lives with her father (F). Obviously, there may be additional reasons for these conclusions which we do not make explicit here, and in order to illustrate the dilemma-nature of the situation the judge has to deal with, we consider a well-balanced case. There are arguments attacking both (M) and (F), and there are arguments attacking those attackers. If we do not consider support, then the grounded extension is $\{CJ, W, OP, R\}$, and for all other Dung semantics, there are two preferred extensions $E_1 = \{R, M, W, OP, CJ\}$ and $E_2 = \{F, CJ, W, OP, R\}$. The judge cannot make any decision based on these two extensions, thus, further investigation is needed in this case.

The first support relation we interpret is the support of the father being wealthy (W) to the argument that it is in the child's best interest that she lives with her father (F). We leave the support of mother side in this following paragraph. There are different options for the interpretation of the support between (W) and (F). If the interpretation is deductive, then we add the supported attacks from (W) to (M) i.e., using RS, we have only one preferred extension $\{W, CJ, OP, R, F\}$. If we add the mediated attack from (M) to (W) with the same interpretation but under RM, then we have two preferred extensions: $E_1 = \{M, R, OP, CJ\}$ and $E_2 = \{R, F, W, CJ, OP\}$. If we choose RD, we still have the two preferred extensions containing (M) and (F) separately. When we consider the interpretation of this support as necessary, saying, for example, that raising a child does take a lot of money, it means adding secondary or extended attack, since there is no attack coming from or going towards (W), we have the same preferred extensions under these two reduction: $E_1 = \{R, M, W, OP, CJ\}$ and $E_2 = \{F, CJ, W, OP, R\}$. We see this as a clear case of deductive support, but as we saw, this doesn't solve the case either.

A similar analysis can be given using the P 4.5 and P 4.6 instead of the reductions, which as Table 2.1 shows are characteristic for deductive and necessary support respectively. P2 represents Closure and is characteristic for deductive support. It says that if (W) is accepted then (F) is accepted. So by contraposition, it means that if (F) is not accepted, then also (W) is not accepted. This implies that for the extension containing (M), where (F) is not accepted, by accepting closure, (W) cannot be in the extension. Based on the above, the extensions show the preferred semantics under deductive reduction satisfies P 4.5. However, in this scenario, the extensions still do not give decisive influence to the decision of the case. It may seem counterintuitive that under the deductive interpretation, a mediated attack is added from (M) to (W), as there does not seem to be a reason to question the wealth of the father. This surprising indirect attack is partly explained by P5, which shows that mediated attack does not satisfy BAF directionality. This reflects that the direction of the indirect attack goes against the direction of the attacks and supports in the framework.

If we consider the attacks and support of the (M) argument, first we need to note that, according to the judicial practice and the public opinion, for decades, (M) was taken for granted: judges automatically gave the custody for the mother, that is, (M) was a *prima facie argument* and (M) was the only argument being accepted. Thus, traditionally, (M) could have been modeled as a self-supporting argument. However, the judicial practice and the public opinion have been changing, so in the figure above, we modeled (M) as a standard argument requiring evidentiary support. While the argument structure on the mother's side seems to be the same as the father's side, there is a difference coming from the law. The supporting argument might have a special status because of the rules of the Civil Code: *the judge has to take the child's opinion into consideration when deciding about custody*. The variants of the support interpretations and their relation to the interpretation of law can be shown with the analysis of this rule. We assume that the child wants to live with her mother (OP). What does this mean? One can say that the obligation of taking an argument into consideration means that the (OP) is prima facie and has to be accepted. If it is a prima facie argument, (M) receives the evidentiary support it needs.

But this in itself doesn't decide how argument (OP) affects the extension. The extension depends on how we interpret the support relation between (OP) and (M): deductive or necessary. It seems to be very intuitive to interpret the support relation deductive: the obligation of taking the opinion into consideration is apparently very much in align with what deductive support means: if we accept the opinion (which is prima facie) then we have to accept (M) too. But if we interpret the relation between (OP) and (M) deductive, under RS we have the only preferred extension $\{OP, CJ, R, W\}$. Under RM we have the two preferred extensions $\{F, W, CJ, R\}$ and $\{M, R, CJ, OP\}$. These results in this scenario, on one hand, reflect the RS satisfies P8, because supported attack is directional. On the other hand, this result means that even the deductive support between the prima facie (OP) and (M), the fact that the child wants to live with her mother won't decide the case in the favor of her if there is some support on the father's side too. However, especially in such a well-balanced case, the judge's obligation to take the child's opinion into consideration might mean that it should be decisive. In order to show what that legal interpretation would mean formally, we need another approach. There is also a way to add the supported attack from (OP) to (F), and mediated attack from (M) to (W), by doing so we have the only preferred extension $\{R, OP, M, CJ\}$. That is, in order to give the opinion a decisive nature, considering both relations between (OP) and (M) and (W) and (F) still deductive support, we do so under different reduction: using RS for (OP) and (M), and RM for the other. This context-dependent solution is needed to represent the given legal interpretation.

We now consider a scenario visualized in red and black in which (OP) is attacked by (4): the child is only 4 years old and 4-year-olds don't know what they want. Argument (4) impairs that the child can form a reliable opinion at all, that is, (4) attacks (OP). If the support between (OP) and (M) is deductive, under RS the unique preferred extension is $\{R, F, 4, CJ, W\}$, while changes to $\{R, F, 4, CJ, W\}$ and $\{R, 4, CJ, M, OP\}$ under RM. If the interpretation of the support is necessary, under R2 the only extension should be $\{R, 4, F, CJ, W\}$, while under RE, the extension is the same as the framework without considering support. The P3 Inverse Closure says that if (M) is accepted then (OP) is accepted. So by contraposition, it means that if (OP) is not accepted, then also (M) is not accepted. This implies that for the extension containing (F), where (M) is not accepted, by accepting Inverse Closure, (OP) cannot be in the extension. Based on the above, the extensions show the preferred semantics under necessary interpretation satisfies P 4.6.

Let's consider another scenario, as visualized in blue and black. In a Hungarian case, the court emphasized that the child's opinion is decisive concerning the custody, *unless* the child's healthy development would be endangered in the environment she would choose. This can be translated as the deductive nature of the support depends on whether there is a specific argument (of being endangered) attacking (M). The child wants to live with her mother, but the mother often changes her boyfriends, and according to the judge, this would endanger the child's healthy emotional development (D). If the interpretation of support is deductive and under RS, the only preferred extension is $\{R, 17, OP, CJ, W\}$, the results under RM and RD are not decisive, either.

Finally, we consider a scenario visualized in green and black. The mother is a teacher, which supports that she knows how to handle children, and this again clearly supports that the child's best interest is to live with her mom. However, we also have the argument that mother often punishes the child harshly attacking (KH). While the first support relation between (T) and (KH) is deductive, it seems reasonable to say the one between (KH) and

(M) is necessary: it is difficult to defend a view as it is fine to give custody to someone who cannot handle children. (M) receives secondary attack from (PH) with the interpretation of necessary, if we still consider (OP) supported attacks (F), and the same for (W) to (M), both (M) and (F) should not be accepted. If we consider (OP) mediated attacks (F), and the same for (W) to (M), (F) is accepted in the only preferred semantics.

4.8 Related Work

The notion of support has drawn the attention of many scholars in argumentation theory, including the role of support in argumentation, whether attack and support should be treated as equals, the link between the abstract approaches and ASPIC+, and also higher-order abstract bipolar argumentation frameworks [263, 264, 147, 77]. We now review and comment on the three approaches to define semantics studied in this paper. For the defence-based approach, we adapted the core notions in Dung's theory. There are other variants of semantics that adapt these notions, such as weak defence for weak admissibility semantics [39, 101], but it is not related to the notion of support. For selection-based approach, semantics$_4$ and semantics$_5$ select extensions based on the number of internal (or external) supports received respectively. Such an approach has already been used in some previous work, and most of them are based on preference [10, 174] or weight of arguments and relations [97, 179]. More recently, Gargouri et al. proposed an approach to select the best extensions to BAFs by comparing the number of received supports with scores for each extension [140]. The reduction-based approach allows a BAF to be transformed into an argumentation graph that has been already discussed a lot in the literature [84, 234, 263, 77]. There is a striking similarity at the abstract level between support in bipolar argumentation and preference-based argumentation, as both can be seen as reductions, as well as both can be used to select extensions [174]. For other approaches to bipolar argumentation semantics, Cayrol et al. proposed some properties of gradual semantics for bipolar argumentation [78], after which Evripidou and Toni provided a concrete definition of gradual semantics for bipolar argumentation [123] and introduced the quantitative argumentation debate (QuAD) framework [38]. Concerning aggregating bipolar argumentation frameworks, Chen considered how to cope with different opinions on support relations and analyse which properties can be preserved by desirable aggregation rules during aggregation of support relations [88]. Lauren et al. also considered aggregating bipolar assumption-based argumentation frameworks under the assumption that agents propose the same set of arguments, different sets of attacks and different interpretations of supporting arguments [188].

Baroni and Giacomin are the first to adopt a principle-based approach for classifying argumentation semantics [35], which was followed by other papers axiomatising abstract argumentation [315], preference-based argumentation [174] and agent argumentation [337]. There are papers that propose principles for bipolar ranking-based/gradual semantics [6], and their generalisations [37]. However, there is a lack of such work for extension-based semantics. Cayrol et al. compared bipolar argumentation semantics, they discussed the semantics based on deductive and necessary interpretations, and provided a few properties, e.g. closure, coherence and safe [84]. Inspired by this work, Yu and van der Torre analysed reduction-based semantics with more properties [338], however, they have only considered reduction-based semantics, without comparing them with others.

4.9 Summary

In this paper, we gave an axiomatic analysis of bipolar argumentation semantics. We considered three approaches, namely defence-based, selection-based, and reduction-based approaches. In total, we introduced seven different types of semantics and studied them together with Dung semantics, which is the baseline and does not take into account supports. Semantics$_{1-3}$ are defence-based, i.e. they are defined by generalising the new notions of defence. Such approach allows us to treat attack and support at the same level. Semantics$_4$ and semantics$_5$ are not only based on admissibility, but also borrow the idea from another field, i.e. social voting, to use the number of support as a way of voting or selecting to derive extensions. Semantics$_6$ and semantics$_7$ are based on the notions of necessary and deductive support respectively. We evaluated those semantics against the set of ten principles. The results are shown in Table 4.1. Given the diversity of interpretations of support, such axiomatic analysis can provide us an overview and systematic assessment of different approaches. It can help us to choose a semantics for a given task or a particular application in function of the desirable properties. One can look at the table and see if there exists a semantics that satisfies the given desiderata.

An interesting question for future work is how to relate semantics defined by various approaches, e.g. can we define a new defence with attacks and supports indicating the deductive, necessary or evidential interpretation of support? We have semantics$_2$ stating that only a supported argument can defend others, which also reflects the idea of evidential support [234, 251]. In this paper, we use dynamic properties, e.g. the robustness of semantics when adding and removing support. This could be further developed by analyzing labelling-based semantics of bipolar argumentation. The distinction between arguments labelled out and undecided makes the principles more precise. We also consider that the approaches to the dynamics of argumentation can be used as a source for principles [59, 174]. Another possible direction is to study the relation between the principles, for example, to verify whether one principle implies another one, or if there is a set of principles such that no semantics satisfies all of them.

Chapter 5

Intelligent Human Input Blockchain Oracle (iHiBO)

This chapter delves into the realm of distributed argumentation technology, bringing to light an experimental platform that harnesses blockchain technology. Specifically, we utilize an abstract agent argumentation framework to model decision-making processes in fund management, which we then implement on the blockchain. This implementation could be conceptualized as a pre-trading phase in the securities market, following which securities transactions could also be carried out on-chain using blockchain technologies. In light of this, the subsequent chapter offers a comprehensive overview of Distributed Ledger Technology (DLT)-based securities markets, accompanied by an in-depth analysis of the associated risks, benefits, and legal considerations.

5.1 Introduction

In situations where trust plays a significant role, the decision-making process can be considered the pinnacle of engagement between parties. For instance, in the case of funds management, investors select managers based both on the prediction of future performance and factors such as reliability and trust [183]. Indeed, in these so-called "trust services", fund managers are in the position of a trustee representing the client, acting in the best interest of the client. Fund managers mainly investigate and determine the optimal securities like bonds and stocks to fit fund strategies, followed by buying and selling them. The decisions made by managers affect principals directly, thus legislators have a duty to inspect the relevant activities of trustees by declaring the rights of clients to highlight this duty through its intended controllability. However, this may not be so straightforward, because these activities such as securities transactions are increasingly executed as a collaborative process involving not only fund managers but also other managers, analysts, and external entities that maintain business relationships. The beliefs and assumptions of different participants may be affected by a variety of background knowledge, thus shaping the decisions that lead to the execution of fund activity. The fund management decision-making process has the characteristics of multiple decision-makers, changing and uncertain information, multiple goals, interdependence between projects, strategic considerations, as well as dynamic opportunities [93]. This requires a collaborative process that all participants consider trustworthy and reliable. It is able to support distributed and iterative boundaries beyond traditional fund management, enable auditability and traceability to maintain accountability, and protect sensitive information at any time.

Financial markets often prompt the image of a hierarchical paradigm of centrally organized institutions. With the advent of Distributed Ledger Technologies (DLTs) in finance, however, is gradually dismantling this stigma of centrality and central counterparties (CCPs), which is only apparently immutable [267, 126], key issues like security, transparency and accountability can be addressed. Decentralized Financial Market Infrastructures (dFMI) [126] are being built in a peer-to-peer fashion to decentralize financial counterparties so that investor confidence can be met with the right alignment of interests. For example, smart contracts can play a similar role to that CCPs played previously in some applications, such as undertaking the task of transferring collateral as a margin calculation agent. In addition, smart contracts can be also applied to dealing with disputes in the case of non-compliance with a payment in a different way [221]. Specifically, Decentralised Finance (DeFi) is a novel P2P financial infrastructure on the basis of smart contracts, providing composable, publicly verifiable, permissionless, and non-custodial operations [328]. However, the involvement of DLTs in fund management fails to address possible trust problems between trustees and clients. In other words, clients still have no access to the cause of the given transaction and whether it occurred in their best interest. DLTs are actually only used to trace the output of such a decision-making process. In this case, a reasoning system for making these decisions can enhance traceability, transparency, and auditability to provide interpretability.

To address these challenges, we propose a novel system to leverage DLTs, smart contracts, negotiation and formal argumentation in a multi-agent setting. It is believed that formal argumentation helps to explain why claims or decisions are made, in terms of justification, dialogue, and dispute trees [99, 339]. At the same time, a negotiation for resolving conflicts can be applied to determining the investment time, quantity, etc. In [339], the theory of our system is first proposed to mainly discuss how the aggregation of argumentation and DLTs can increase trust. On the other hand, the contribution of this chapter includes the implementation of our new system and the evaluation of the feasibility of our proposal. The utility of the proposed system is demonstrated by using a private Ethereum Blockchain to achieve a proof-of-concept for resolving conflicts. As far as we know, our contribution is the first to include formal argumentation implemented using smart contracts together with a multi-agent system. This leads to the impossibility of comparison with related works in performance evaluation.

The remainder of this chapter is structured as below. Section 5.2 gives the motivation of our work, Section 5.3 describes the significant components of our system, including formal argumentation, negotiation, blockchain and smart contracts. Section 5.4 illustrates three possible DLT-based systems for decision-making as well as the analysis of each. Section 5.5 introduces the IHiBO framework in details and its architecture. Section 5.6 presents the experimental results. Section 5.7 conducts a discussion of IHiBO from the perspective of trust and legal consideration. Related work is discussed in Section 5.8 and Section 5.9 concludes.

5.2 Motivation

In this section, we generally talk about the procedure of fund management (at the securities market) and the roles of the parties and their relation, in order to show that the decision-making process can be suited into argumentation modeling.

In portfolio management, the core duties of fund managers under AIFMD[1] and UCITSD[2] is to perform portfolio management and risk management on behalf of their investors[3]. The fund can be managed by a team of three or more people. Fund managers primarily research and determine the best stocks, bonds, or other securities to fit the strategy of the fund, then buy and sell them. Since the fund managers are responsible for the success of the fund, they must also research companies, and study the financial industry and the economy. Keeping up to date on trends in the industry helps the fund managers make key decisions that are consistent with the fund's goals [87]. The main characteristic of investing in a fund is trusting the investment management decisions to the professionals.

The process of portfolio management on the manager side is formally defined as follows [93][94]: *Portfolio management is a dynamic decision process, whereby a business's list of active new product (and development) projects is constantly up-dated and revised. The portfolio decision process encompasses or overlaps a number of decision-making processes within the business, making Go/Kill decisions on individual projects on an on-going basis, and developing a new product strategy for the business.*

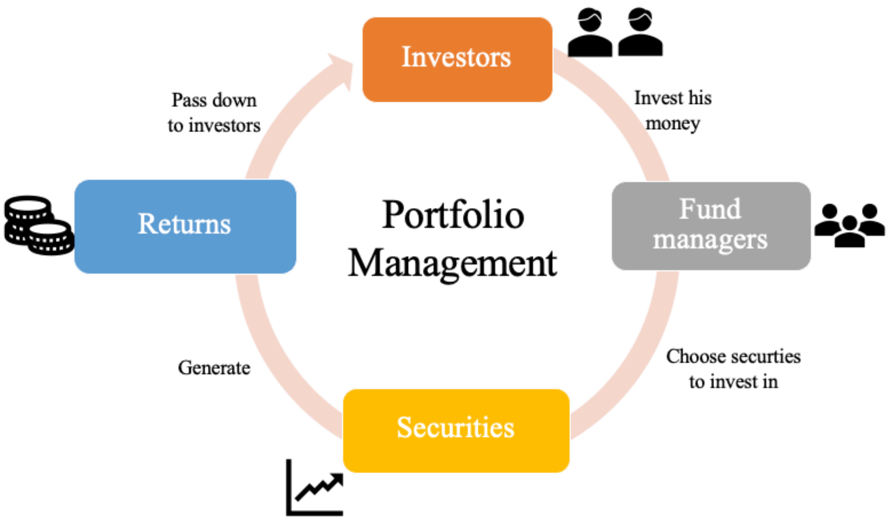


Figure 5.1: Fund Investment Process

In Figure 5.1, investors first pool their money together and then fund managers conduct investment research, prepare the specific plan for the investment portfolio. According to their research and the final decision of the investment plan, fund managers invest securities

[1]Directive 2011/61/EU of the European Parliament and of the Council of 8 June 2011 on Alternative Investment Fund Managers (AIFMD.http://data.europa.eu/eli/dir/2011/61/oj)

[2]Directive 2009/65/EC of the European Parliament and of the Council of 13 July 2009 on the coordination of laws, regulations and administrative provisions relating to undertakings for collective investment in transferable securities (UCITS). http://data.europa.eu/eli/dir/2009/65/oj

[3]A portfolio is a collection of financial investments like stocks, bonds, commodities, cash, and cash equivalents, as well as their fund counterparts. In this chapter, in order to give a concise explanation, we only consider stocks.

on behalf of their clients (investors). The investment generates returns and the returns would be passed down to investors.

5.3 Distributed Ledger Technologies and Smart Contract

In this section, we introduce the components to build our framework IHiBO in Section 5.5 except for formal argumentation which is introduced in previous sections, i.e. distributed ledger technologies, autonomous agents and negotiation.

5.3.1 Distributed Ledger Technology

Distributed Ledger Technologies (DLTs) consist of networks of nodes that maintain a single ledger and follow the same protocol for appending information to it. The blockchain is a type of DLTs where the ledger is organized into blocks and where each block is sequentially linked to the previous one [224]. The execution of the same protocol, i.e. source code, guarantees (most of the time) the property of being tamper-proof and not forgeable. This allows a trust mechanism to be created without the need for third-party intermediaries [40].

In the architecture of a traditional centralized database, information is positioned, stored, and kept in a single location and under the control of a central administrator, who guarantees integrity. The data is often kept in a raw form, while some security perimeters keep the database from outer attacks, while a distributed database can give an alternative configuration, by which the database (or its copies) runs through a network and can be kept at different physical positions. Although the database can be shared by all the nodes, its control is often centralized, while the database integration is decided by a central administrator. A centralized application can facilitate data synchronization and database management. The nodes are within the security perimeter and can be trusted, while a network administrator regulates the access. Although the terms blockchain and DLT are usually used interchangeably, they are different. Blockchain is a particular Distributed ledger (DL) architecture in which data is divided into a sequence of batched blocks that are mutually linked by cryptographic tools. Various blocks of data highly rely on the previous blocks so it is extremely difficult to change the database in a retroactive manner. Thus, a perpetual chain with immutable records is formed by the blockchain. Rights of validation and access - There are diversified kinds of DLs based on the accessibility limits and the type of employed consensus processes employed so that the integrity of a ledger can be ensured. In terms of a specific object with access to data, DLs may be:

Public, wherein a user can view/read a ledger, or

Private, data is accessed by only participants approved.

In a similar way, according to the object allowed to realize ledger verification, DLs can be:

Permitted, wherein ledger verification or modification of ledger entry is only allowed by a specific group of trustworthy users;

Permission-less, wherein everyone is allowed to verify and build the ledger. As central authority is not needed for entry verification of the ledger, it is resistant to censorship; namely, no actor can avoid a transaction's adding to the ledger [126].

Accessibility is closely correlated with confidentiality and privacy. In a public ledger, such as the Bitcoin blockchain: for example, everyone can view that a user A has made payment of amount X to user B (despite unknown true identities of user A and user B), and security transactions often need to be confidential. With trade repositories and regulators as the exception, in most cases, only those parties legally asked to be informed must view the details of trading. It is still difficult to keep a DLT implementation confidential, which is worthy of research. Moreover, financial markets could be supervised, so a challenge may be caused by permission-less structures, in view of the potential opacity. Because of these aspects, the settlement industry mainly focuses on private and permitted ledgers, wherein validation rights are only granted to a trusted group of participants, while nodes can be given different roles carrying diversified access rights.

Consensus mechanisms concerning any database, to ensure a ledger's integrity, and modifications thereof must be validated. For each new transaction: for instance, somebody has to check if counterparties have sufficient cash or assets in accounts and guarantee that these assets or cash are used over one transaction (i.e., it is free of dual-spending). With a central authority of trade validation, for the solution of this problem, DL protocols are mainly dependent on a consensus mechanism, wherein all the nodes or a selected subset agree on the validity of a trade as well as the most latest version of a case. The consensus mechanism will be designed according to whether the DL is permitted or not.

As for permission-less DLs, participants' identity is unknown and thus there must be a consensus mechanism, which does not depend on the trust of participants or identification of their identities. Such protocol of validation cannot be decided by a simple majority (one vote for each node), as an attacker may have the chance to change the ledger by controlling over 50 percent of votes or creating nodes. Hence, one of the problems is about how participants share validation rights without making a significant compromise about the system security due to the risk brought by denial-of-service and Sybil attacks, i.e., a Sybil attack is a traditional way of undermining a P2P network by creating large amounts of fake identities in order to gain a disproportional amount of influence or votes in the network. Bitcoin solution was a mechanism of consensus based on "Proof of Work" (PoW) nakamoto2008bitcoin; the so-called "proof" is realised through the solution of a computationally-intensive math question which greatly increases the economic cost for a single attacker, so he fails to alter a ledger by enough computing capability. Theoretically, under the one-CPU-one-vote rule, an attacker needs to have control of at least 51 percent of the computing capability for altering a ledger; nevertheless, under certain cases, control of even 30 percent of the capability might be enough for the ledger alteration eyal2018majority. The consensus process of "Proof of Stake" (PoS) is another system of validation. It assigns a share of validation rights to the participant according to stakes hold from the system. The measurement standards for stacks held by the validators from the system may be different. They can be measured by the forms of off-ledger assets or internal tokens or pledged as collateral, or according to the fame of the validator in a network (in case of the identity being known). As requested by some processes of PoS, voters must put "bets" on a ledger's true state, so would-be attackers would be likely to claim some false things, who will become a minority losing the bets and carrying the prediction14 castro1999practical. Permitted systems do not need to tackle the problem concerning the untrusted network. Thus, they can process transactions faster at a lower cost to maintain compared with capital-intense permission-less ones. In addition, arguably speaking, permitted systems would be further open towards reversibility and censorship. On the difference between deploying permissionless and permissioned DLT systems in this domain: see (Priem, 2020). Overall, permissioned systems are more suitable to be implemented in securities markets. They also lower the risk of money laundering. However, in a permissioned system there must be an entity (institution) that acts as a gatekeeper, which means the idea of a central institution is reintroduced.

5.3.2 Smart Contracts

In addition to transaction data, distributed ledgers may involve computer code, which is known as “smart contracts”. After satisfying the pre-defined conditions, these contracts automatically execute and deal with a transaction on the ledger. The concept of smart contracts was first introduced by Nick Szabo in 1993, and verification and contractual obligations are executed through self-enforced computer code in this type of cryptographic [302]. Such contracts need to be drafted, authenticated, and executed in an unbiased environment which makes distributed consensus-oriented ledger networks ideal platforms. A pertinent example is Ethereum smart contracts. As a second-generation blockchain technology, Ethereum has been specifically designed as a decentralized smart contract platform. In addition to being implemented with the Turing-complete scripting language, Ethereum has a built-in virtual machine (EVM) that works as a decentralized computer. The smart contracts in Ethereum can be written in Serpent and Solidity and other high-level languages. The transactions in DLT are essentially smart contracts, it is only the complexity that varies. For example, a smart contract in a security trading scenario refers to an automatically executed contract in which the buyer-seller agreement’s terms are written in a few lines of code. Its code and agreements are in a distributed and decentralized network. The transaction is traceable and irreversible and the code controls the execution.

Some DLTs allow smart contracts to be executed. These consist of instructions that, once distributed on the ledger, cannot be altered. Thus, the result of their execution will always be the same for all DLT nodes running the same protocol. Smart contracts enable a wide range of applications far beyond cryptocurrency transactions, especially in the Ethereum [184, 344, 348] blockchain. However, smart contracts are usually “closed” to the outside world, e.g. they cannot contact a website, to ensure that execution is more resistant to attacks with a higher degree of certainty [344]. This obviously limits the possibilities of using these technologies, given that the great majority of possible applications would require real-time information from the world outside the network.

In this context, Oracles assist DLTs in enabling smart contracts to operate in the real world by flowing any kind of data from any kind of service external to the DLT. They act as a bridge, providing the ability to retrieve, verify and digest data into smart contracts. An oracle can be implemented as [46]: (i) software, interacting with the necessary information from online sources; (ii) hardware, retrieving data from the physical world directly through sensing devices; (iii) humans, interacting with individuals. In all cases their off-chain execution can be centralized, i.e. from a single source, or decentralized, based on the consensus of a multitude of sources.

5.3.3 Autonomous Agents and Negotiation

Multi-agent systems are distributed platforms composed of a set of agents that interact with each other through an appropriate organization, with properties of self-organization, self-adaptation, self-maintenance. An agent can be defined as a computer system that is located in some environment and is able to act autonomously to achieve its goals. Agents are designed to be related to individual perspectives [301], but this also includes aspects of autonomy, context, responsiveness and pro-activity (rationality) [331]. Agents are good candidates for representing the subjectivity and nuances of different expert opinions. Moreover, they are able, using well-established conflict resolution mechanisms (e.g. negotiation), to help different stakeholders find agreements that satisfy their often conflicting interests.

In this work we focus on negotiation. It is the process by which a joint decision is made by two or more parties, agents in this case. The agents firstly verbalize contradictory demands and then move towards agreement by a process of concession making or search for new alternatives [270]. The problem being negotiated, or the topic under discussion (e.g. stock price) can be usually divided into issues (also called attributes). Automated negotiation is one taking place among autonomous agents through a shared protocol, i.e. a set of rules that governs their interactions during a negotiation session (also called a thread). Whilst the protocol is shared among agents, the decision model for acting in the negotiation is unique for each agent [125]. It allows the agent to (i) evaluate the value of an offer received from the opponent (e.g., using a utility function), (ii) decide whether it is acceptable, and (iii) determine what to do next (known as the negotiation strategy).

Automated negotiation has been applied to solve conflicts and reach agreements in several domains including cloud and service provisioning [223], smart grid and power distribution [309], and trading and stock market [327]. Compared with human negotiation, autonomous agent negotiation is efficient in contexts where the number of issues under negotiation is intractable for human users, or in one-to-many [204] or many-to-many negotiation [12] settings in which the numbers of negotiators makes it difficult for humans to keep track of the evolution of the negotiation process. Therefore, autonomous agents can offload these tasks from the human expert shoulders, assist them in formulating their preferences, and help reach optimal solutions that can be otherwise inaccessible to human negotiators with the agent assistance.

5.4 Analysis of Architectural Designs for DLT-based Systems for Decision Making

In recent years, the adoption of DLTs and smart contracts has been taken into account because of the immutability property of DLTs, which allows a favorable environment for storing information that can be subsequently audited. When it comes to decision-based systems to implement a process on DLT, most solutions using these technologies only store the final output of the decision process (which then triggers a process, e.g. via a smart contract). In our work, on the other hand, we pay attention to the decision phase. Particularly, in a fund management context, this is the pre-trading phase where the investment decisions are made. Moreover, we deal with a case in which the final decision enacts a process on a “mainchain”, i.e. an on-chain process. This mainchain is a public permissionless ledger accessible to everyone, such as the Bitcoin or Ethereum blockchains [224, 66]. For instance, the outcome of a negotiation, i.e. the decision, can be given in input to a smart contract in the Ethereum public blockchain that will enact an action, e.g. security transaction, in order to buy a stock. As stated earlier, current smart contract implementations need oracles in order to perform operations based on data that is found outside the ledger, i.e. off-chain. In the use case we address in our work the input comes from a human, hence we are dealing with human-input-based oracles.

Based on this premise, in the following, we provide an analysis on the possible architectural designs of a DLT-based system that takes into account formal argumentation and negotiation for a decision-making process.

We argue that three different generalized versions of architectures can take the form:

Case1: Centralized human-input-based oracle - in this case, argumentation and negotiation phases do not involve any DLT process, neither a smart contract execution. These phases are executed in a "centralized" environment, e.g. a web platform or an internal firm application, after a human expert has provided an input. Then, a single oracle, depending on the result of the argumentation and negotiation steps performed in the centralized environment, directly stores the result in the mainchain, enacting an on-chain process[4].

Case 2: Mainchain smart contract for argumentation and negotiation - a second case would be to implement the argumentation and negotiation processes directly as a separate smart contract in the mainchain. Human experts directly interact with the blockchain for giving in input the data for the two processes. The execution of these processes never leave the mainchain thanks to the smart contract implementation and the result of the decision is directly given in input to another smart contract that implements the desired behavior (e.g. buy a stock).

Case 3: Decentralized human-input-based oracle - finally, as the third case, we consider a network of independent nodes that execute a distributed oracle software. This network does not take part in the mainchain protocol and limits the interactions with the ledger to only a few, necessary to provide trustworthiness to the oracle. The implementation of such a network consists in a so-called "layer two" solution [155], where the same principle of decentralization of DLTs is applied. The execution of argumentation and negotiation processes is distributed among the network participants and the result of the decision is directly given in input to the mainchain, enacting an on-chain process.

We take as reference Table 5.1 for comparing the three architectures and we refer to them as case 1, 2 and 3 for centralized, smart contract and decentralized respectively.

The first discriminator for these cases is where the argumentation graph (and all the remaining data needed to the execution) is stored. Such data enables the execution of the argumentation and negotiation process, hence is constantly updated. This information is needed to be stored on-chain, i.e. on the mainchain's ledger, only in case 2. The drawbacks of storing large quantities of data on-chain are many, above all, the elevated transaction fee cost [184] and that scalability and response latency is often compromised for some features, such as smart contracts execution [288, 348]. Furthermore, the fact that instruction execution would be unfavorable compared to the other two, again in terms of fees and latency between operations, should also be taken into account for case 2. However, the advantage of this architecture is that the negotiation execution would be completely traced

[4]The reference to "on-chain" will always be to the mainchain through the text.

Table 5.1: Comparison between the three architectural choices: centralized vs. decentralized oracle vs. mainchain smart contract argumentation and negotiation execution. Columns indicate the presence or not of some features or properties. The × symbol indicates no presence, while the ✓presence. The ✓✓indicates that the feature or property is intrinsic to the chosen solution (i.e. it cannot be otherwise).

	Argumentation Graph Off-chain	Negotiation Execution Tracing	No Single Point of Failure	No Execution Overhead Costs	Tamper-resistant	Verifiability	Privacy
Case 1 **Centralized**	✓	×	×	✓	×	×	✓✓
Case 2 **Smart Contract**	×	✓✓	✓	×	✓✓	✓✓	×
Case 3 **Decentralized**	✓	✓	✓	×	✓✓	✓✓	✓

and verifiable, since the smart contracts execution is completely traced in the mainchain. Indeed, a public permissionless DLT solution is that it usually offers a high level of security and decentralization [102], needed to completely trace and verify processes with trust.

On the other hand, the complete execution of the decision process would be poorly verifiable using a centralized oracle (case 1), because only the results would be stored on-chain. This case would also be highly susceptible to single point of failures.

A good compromise between the two cases would be the use of a decentralized oracle, case 3. The data needed for carrying out these processes, such as the argumentation graph, would be stored in a layer two technology that preserves data immutability and that is cheaper in terms of fees and latency costs. The negotiation execution can happen outside the mainchain, i.e. off-chain, and then be "committed" on the mainchain using a hash function in order to be verifiable [155]. This case would not be susceptible to a single point of failure and the execution overhead cost would be favorable in respect to case 2.

On the basis of this analysis, the solution we will propose in the next sections will be based on case 3, i.e. it is a decentralized human-input-based oracle architecture. We implement this architecture as a layer two solution. In fact, we refer to the layer two solution of using a sidechain in support of the mainchain [155]: (i) the first layer includes a public permissionless DLT, i.e. mainchain, while (ii) the second layer consists of a private permissioned DLT, i.e. sidechain. Private permissioned DLTs solve the public permissionless issues of: (i) the publicity of information that would clash with trade secrets and privacy, as only allowed actors can read from the ledger; (ii) expensiveness and scalability, as permissioned DLTs protocols can be designed ad-hoc to specifically address these issues. However, the level of security of private permissioned solutions decreases in respect to public permissionless ones, due to the fact that they generally are less decentralized and that usually use more efficient but less secure consensus mechanisms [288, 102].

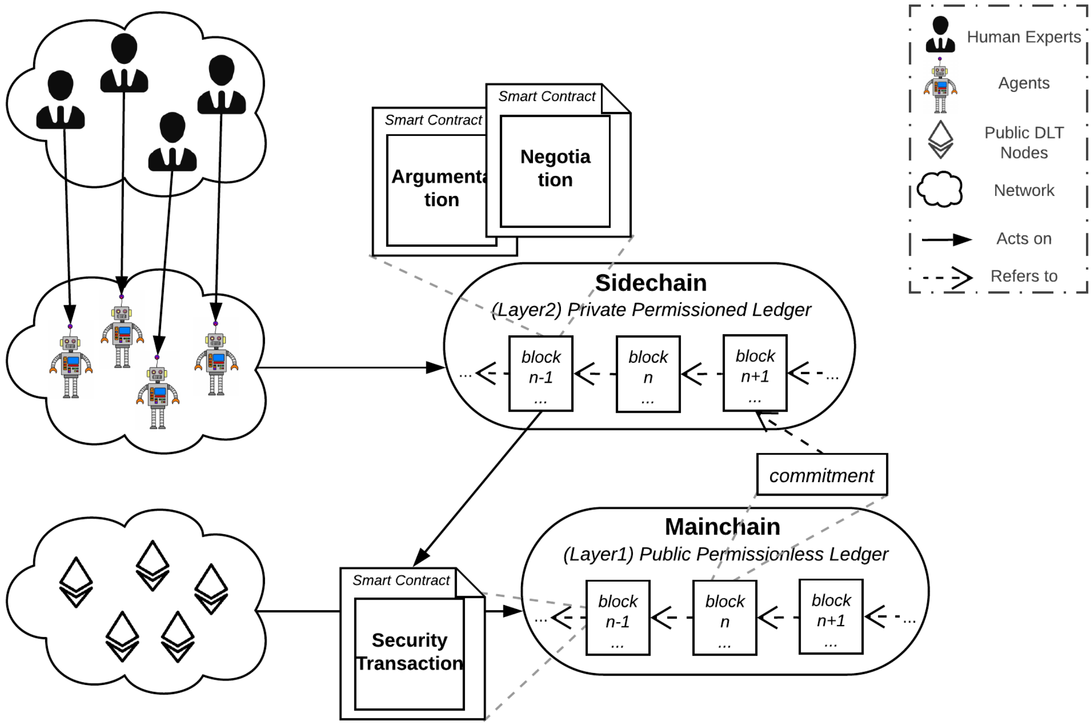


Figure 5.2: IHiBO Framework Architecture

5.5 Intelligent Human-input-based Blockchain Oracle (IHiBO) Framework

In this section we describe in detail the Intelligent Human-input-based Blockchain Oracle Framework that we intend to propose. IHiBO is a decentralized human-input-based oracle that enables the execution and traceability of argumentation and negotiation processes involving human experts. We focus first on describing the architecture of the framework and then proceed to address the individual implementation aspects.

Before going into the architecture details we describe the roles of the actors involved in the architecture, with reference to Figure 5.2:

Human Expert, the one who takes most of the decisions and that gives inputs to the agent.

Agent, the one that can assist human experts in executing the argumentation process and that negotiate with other agents; agents are also the ones that directly interact with the sidechain.

Public DLT Node, the one that takes part to the mainchain consensus mechanism and that is external to the sidechain; this actor receives transactions to be stored in the mainchain, i.e. a DLT full-node [224].

5.5.1 Architecture

Most of the processes executed within the framework run on a sidechain. Specifically, the sidechain consists of a DLT where smart contracts are executed and data stored, whereas, the mainchain is the DLT where "commitments" for the security of the framework are periodically stored. Moreover, the mainchain is where the result of the decision process is stored and enacted, e.g. execution of a DeFi contract for transferring funds.

Figure 5.2 shows a diagram describing the whole architecture. A network of agents and/or other nodes maintain the sidechain that acts as a decentralized oracle. We refer to a private permissioned DLT for the framework sidechain, where only agents have the permission to read and write to the ledger. Furthermore, the mainchain and sidechain are tied together in the framework by the use of periodical commitments. A commitment consists of storing in the mainchain the result of a hash function applied to the state of the sidechain at a certain point in time. This would allow it to store data that cannot be tampered with in the mainchain and to allow its verification. At the same time, thanks to the hash function, the privacy of information stored in the sidechain is maintained, while assuring that any data corruption will be detected [155], i.e. the hash result will change. Indeed, through the use of commitments, once the nodes operating the sidechain reveal part of (or all of) the information stored in the sidechain to possible auditors, the latter can apply the hash function to the data received and check that the obtained result is equal to the hash stored in the mainchain [296].

5.5.2 Mainchain

For the IHiBO Framework implementation we refer to Ethereum and to its smart contract specification [66].

In particular for the mainchain, we leverage the Ethereum public blockchain and the functions exposed by its network nodes for creating and/or interacting with smart contracts. In the Ethereum blockchain some applications built through the use of smart contracts, i.e. decentralized applications (dApps), are already being developed for the execution of securities transactions [259]. The Ethereum protocol allows smart contracts inter-communication, thus the framework we present is meant to include a dedicated smart contract that "bridges" the output of the execution in the sidechain, e.g. a conflict resolution, to a smart contract deployed to the mainchain. For this we refer to the ChainBridge implementation. ChainBridge [86] is a multi-directional modular blockchain bridge to enable the transfer of data and value between any number of Ethereum-based blockchains. This solution allows agents to specify a destination blockchain, i.e. the mainchain, from their source chain, i.e. the sidechain, and send data to that blockchain for consumption on the destination chain[5]. This could be an operation that is initiated on the source chain, i.e. the argumentation and negotiation process, and then finished in a smart contract deployed to the mainchain. We refer to this final smart contract as the "SecurityTransaction", and its implementation mostly depends on the on-chain business process it interacts with. For instance, Decentralized Finance (DeFi)[6] protocols such as Decentralized Exchanges (DEX) have already been provided in the Ethereum blockchain for enabling anyone to engage in non-custodial exchange of on-chain digital assets, e.g. tokens [328]. Smart contracts that implement such DEXes can be directly invoked for swapping tokens and cryptocurrencies depending on their value [170]. An instance of a SecurityTransaction would be a smart contract that includes a method that directly invokes a DEX smart contract for executing

[5]In this work we are not interested in the detailed description of the blockchain bridge, and we leave it as future work the optimization of such component.

[6]DeFi is a term that refers to smart contract based financial infrastructures that are non-custodial, permissionless, openly verifiable and composable [328].

a token swap. This can be seen as the direct selling/buying of traditional stocks that have been “tokenized” [170, 53].

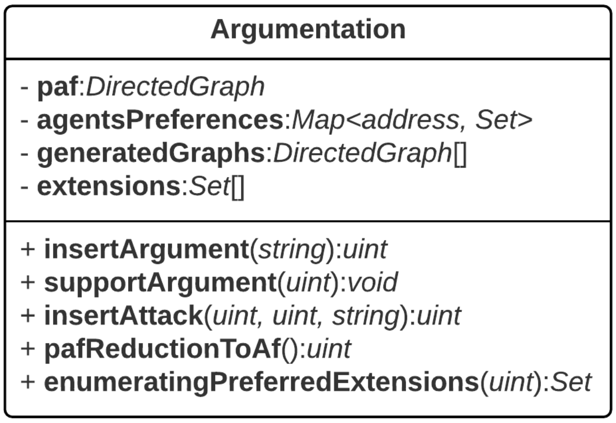


Figure 5.3: Argumentation smart contract class diagram

5.5.3 Sidechain

For what concerns the sidechain, any implementation of a permissioned smart contract enabled DLT is suitable for the framework we proposed. In our implementation, we make use of a permissioned Ethereum blockchain distributed among nodes in a private network. In this case, the consensus algorithm adopted by the network does not necessarily have to be the Proof-of-Work [224, 66], but, in order to provide a faster service, the Proof-of-Authority (PoA) consensus is used [312]. PoA, indeed, does not depend on solving mathematical problems, and to issue a new block this one must be signed by the majority of the authorities, i.e. the nodes that are explicitly authorized to create new blocks and secure the blockchain.

The main purpose of this sidechain is to support smart contracts whose execution log can be later audited. Thus, we implemented two smart contract specifications for executing argumentation and negotiation processes, however many others can be implemented following the Ethereum smart contracts specification [66].

Argumentation Smart Contract

We implemented a smart contract for providing a PAF (Section 5.3) to the agents that operate in the sidechain.

- A data structure within the smart contract allows to create and manage a directed graph, where nodes are arguments and edges are attack relations. Each agent can add a set of arguments (*insertArgument()*) and attacks among arguments (*insertAttacks()*) to the graph or support an argument already add by another agent (*supportArgument()*).

- Arguments are handled through their id and the metadata associated with it, i.e. the actual argument text, can be stored directly on the ledger or outside and referenced through a hash pointer.

- After a predefined time period needed for completing the PAF, reductions of PAF to AF can be invoked and executed directly by the smart contract (*pafReductionToAfPr()*). The result of invoking this method is a new directed graph representing the corresponding AF.

- Finally, the semantics can be found for the previously obtained AF, by invoking another method (*calculatePreferredExtensions()*). The implementation of this method is based on the algorithm found in [226] for listing all preferred extensions of an AF (Algorithm 1). This possibly provides a set of arguments that leads to a final decision.

Negotiation Smart Contract

We implemented a smart contract that concludes the conflict resolution (Section 5.6.1) with a negotiation on the arguments provided by the argumentation process.

- A data structure within the smart contract holds the data needed during a negotiation thread. A list of such structures enables agents to interact for automated negotiations on several issues. Each agent can start a new negotiation with another agent for a specific set of issues (*newNegotiation()*).

- Each agent has its own decision model executed off-chain, that allows this to evaluate the value of an offer received from another agent, e.g. a time dependent tactic [125].

- Based on the evaluation, the agent can invoke the smart contract to make a new offer (*newOffer()*) providing a new set of values related to the issues, accepting (*accept()*) the other agent's offer, or refusing it (by not providing input to the smart contract).

- The invocation of the smart contract method for accepting the offer can directly enact the process of interaction with the SecurityTransaction smart contract on the mainchain [281].

5.6 Validation

In this section we present a use case and an experimental evaluation that can help to validate our proposed IHiBO framework. We provide a conflict resolution use case described through an agent argumentation framework and a negotiation thread. Then we provide and discuss the results obtained by evaluating the use case smart contract implementation in terms of gas usage.

5.6.1 Conflict Resolution Use Case

In this subsection we use a simplified use case to illustrate how we use abstract agent argumentation and autonomous negotiation for dealing with conflicting information raised by agents in IHiBO.

Argumentation Framework

The process of decision-making in fund management fits well with argumentation theory. The decision can be seen as being based on arguments and counter-arguments. Argumentation, as the result, can be useful for deriving decisions and explaining a choice already made. Managers provide their arguments from their own research to identify promising stocks with different levels of accuracy and thereby make different portfolio choices which are likely to be incomplete and inconsistent. The fictitious simple example (the real life cases would be much more complex) is as follows.

- Manager α and β hold the arguments a: *To buy the stocks, since the company just donated to charities that is beneficial to good commercial reputation*
- They also hold an argument c: *To buy the stocks, since the company has started to use a new promising technology which will develop the sales performance.*
- Another manager γ at the same time is against buying the stocks, he holds the argument b: *To sell the stocks, since there is evidence that the leader is under accusations of charity fraud*
- And also holds d: *To sell the stocks, since the company now has poor sales performance.*

Based on the above, we can build an agent argumentation framework on the left side of Figure 5.4, $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ where $\mathscr{A} = \{a, b, c, d\}$, $\rightarrow = \{(a, b), (b, a), (a, d), (d, a), (b, c), (c, b), (c, d), (d, c)\}$,S=$\{\alpha, \beta, \gamma\}$, $\sqsubset$= {(a,α),(a,β),(c,α),(c,β),(b,γ),(d,γ)}. Since $|\mathscr{S}_a| > |\mathscr{S}_b|$, $|\mathscr{S}_a| > |\mathscr{S}_d|$, $|\mathscr{S}_c| > |\mathscr{S}_b|$, $|\mathscr{S}_c| > |\mathscr{S}_d|$, $a > b$, $a > d$, $c > b$ and $c > d$, we get the corresponding PAF showing in the middle of Figure 5.4, and giving the four reductions from PAF to AF, we have the AF on the right side of Figure 1. Then we can calculate the only acceptable set $\{a, c\}$ which is the only grounded, complete, preferred and stable extension. The set tells the final decision is to buy the stocks.

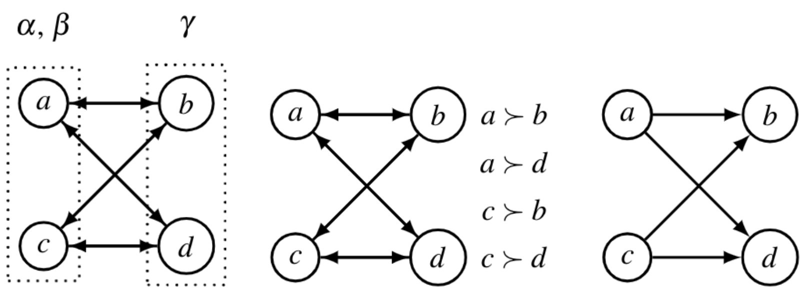


Figure 5.4: Social Reduction that translates AAF (in the left) to AF (in the right), where PAF (in the middle) is an intermediate step, the unique extension of this AAF is $\{a, c\}$.

Agents' Negotiation

One thing needs to be noticed: argumentation does not always provide a unique outcome. People need to select the desired semantics based on various reasoning flavor [33]. On the other hand, depending on the decision making process, different protocols can be specified in advance for such cases: e.g. to roll back or to assign weights to the arguments and the relation among them. After reducing to AF and calculating the acceptable set, indeed, when the outcome results in the decision to buy the stocks, the next problem could become the numbers of stocks to buy and the buy timing. Here the computational automated negotiation comes into play.

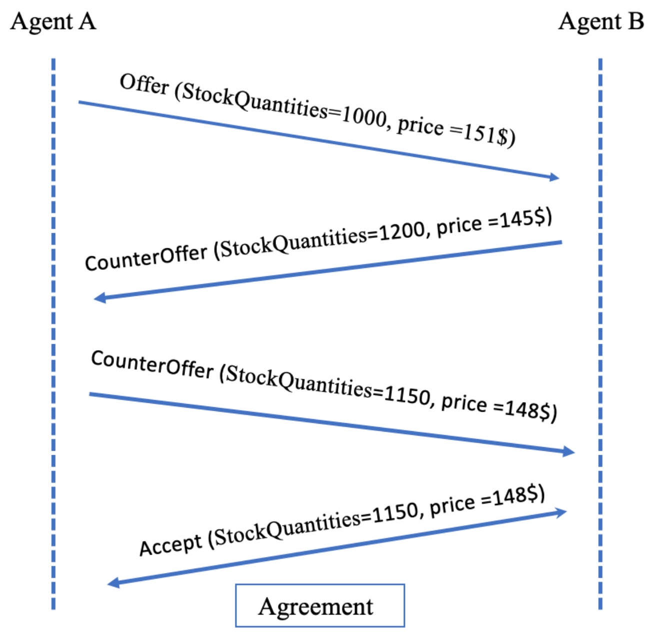


Figure 5.5: Negotiation Sequence to Decide The Quantities and The Price

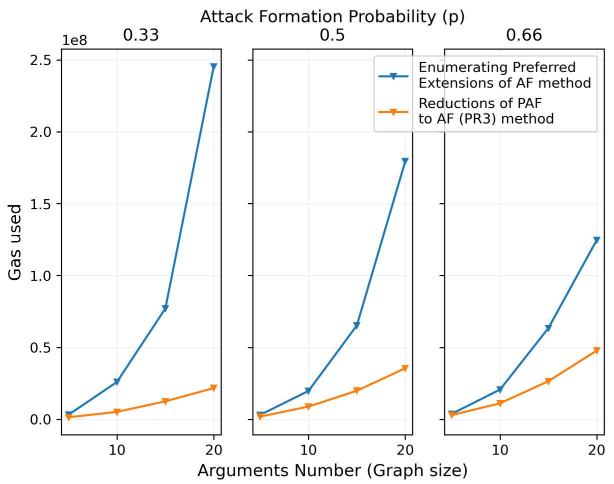


Figure 5.6: *pafReductionToAfPr()* and *calculatePreferredExtensions()* methods' gas usage when varying the arguments number (i.e. the number of graph nodes). Each plot shows the results for a different attack (i.e. graph edge) formation probability, namely 0.33, 0.5, 0.66.

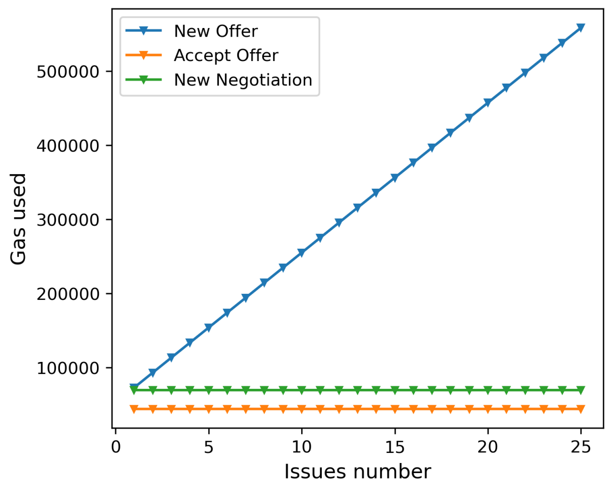


Figure 5.7: Negotiation gas usage for the newNegotiation(), newOffer() and accept() methods.

To illustrate how it works, we give an example of the negotiation sequence based on the quantities of stocks to buy. The negotiation process is based on the alternating offer protocol [282]. Agents can bid new offers to the opponent ($Offer()$ function). When receiving an offer, an agent can accept it using the $accept()$ function or reject it and propose a counter-offer (with the $CounterOffer()$ function). In the example shown in Figure 5.5, we have manager A, i.e. agent A, and manager B, i.e. agent B. Agent A proposes to buy 1000 stocks at the price of 151$, while agent B counteroffers to buy 1200 stocks at the price of 145$, then agent A proposes to buy 1150 stocks at the price of 148$. The final offer given by A is accepted by both parties which means they come to an agreement.

5.6.2 Experiments

We tested the feasibility of the use of smart contracts for conflict resolution and here we present the results of some experiments based on two assumptions. Firstly, we are not interested in testing out the performances in terms of transaction per seconds and scalability for public permissionless DLTs, since these have already been studied in literature for similar use cases [288, 102, 184, 347]. Indeed, these results have already impacted the IHiBO framework design by limiting the issuing of transactions to the mainchain only for periodic commitments [339]. Secondly, for what regards the sidechain, performance depends on the specific implementation used by the actors in a specific use case. In our implementation we used an Ethereum private network using PoA and it has been shown that, with optimal configuration, it can reach up to 1000 transactions per second [312].

Therefore, our focus is on the execution of the smart contracts (that we described in the implementation section 5.5.2), applied to the conflict resolution use case. We measure our experiments in terms of gas usage, following the Ethereum protocol [66]. Gas is a unit that measures the amount of computational effort that takes to execute operations in Ethereum smart contracts. Thus, the higher the gas usage for a method, the more intense the computation of a blockchain node to execute the method's instructions. The complete experiments dataset and the reference software can be found in [346], following the FAIR data principles for access and reuse of models [329].

Table 5.2: Gas Usage

Smart Contract	Method	Occurrency	Gas Usage
Argumentation	*insertArgument()*	a	157 470
	supportArgument()	$\leq a \times [n-1]$	80 491
	insertAttack()	$\leq a \times [a-1]$	215 011
	pafReductionToAfPr()	1	1 877 277
	calculatePreferredExtensions()	1	1 412 065
Negotiation	*newNegotiation()*	1	104 961
	newOffer()	t	52 438
	accept()	1	64 211

Results

Table 5.2 shows the gas usages for the execution of the Argumentation and Negotiation smart contracts methods, taking as input the data of the example in Section 5.6.1. These results give an indicative idea relative to the different method executions, since their latency (i.e. the time between submitting a transaction that invokes such methods and the actual insertion to the blockchain) depends heavily on the blockchain's consensus mechanism. For instance, considering a as the arguments number and n as the agents number, the *supportArgument()* method is much less expensive than the *calculatePreferredExtensions()*, but it is executed up to $\leq a \times (n-1)$ times while the latter only 1 time.

In Figure 5.6 it is shown the increase of the gas usage while varying the AF. For each argument number a taken into consideration, i.e. $5, 10, 15, 20$, some graphs representing a different AF have been created randomly. In these graphs, the edge connecting any two nodes, i.e. an attack in the AF, was firstly formed with a probability of 0.33, then 0.5 and finally 0.66. For each probability value, 20 random graphs were created and the average of gas usage for invoking the methods *pafReductionToAfPr()* and *calculatePreferredExtensions()* was computed. For the latter method, results show that the gas usage depends heavily on the arguments number a, as with the increase of a the gas usage increases exponentially. At the same time, however, incrementing the edge formation probability p leads to a decrease of the gas usage. Results for the *pafReductionToAfPr()* method show a much less dramatic increment of gas usage with the increase of a, but here the increase of edge number leads to an increase of gas usage instead of a decrease. The minimum value for the *calculatePreferredExtensions()* method gas usage is ~ 1.2 million gas units, while the maximum is ~ 528 million gas units. For what concerns the other method, ~ 0.6 million gas units is the minimum and ~ 51 million the maximum.

Finally, we provide the results of the measurement of the gas usage for the *newOffer()* method of the Negotiation smart contract. In this case, we implemented two agents negotiating using a time dependent tactic, as in [125], with two different sets of starting conditions and maximum values. The number of new offers t proposed by each agent cannot be known a priori because it depends on the specific strategy of the agent. For this reason we measured the impact of the issue number j on the gas usage. Figure 5.7, indeed, shows that the latter increases linearly with the former, due to the increasing storage demand. On the other hand, the *newNegotiation()* and *accept()* methods have a constant gas usage because these are used only to respectively open and close the negotiation thread.

Generally speaking, we experienced a strong dependence on the argument number for the increase of the gas usage. This was expected, as more arguments means a more complex argumentation framework to deal with. The use of a private Ethereum PoA network allows to limit the latency based on the results obtained in [312]. Assuming one invocation per transaction, methods such as *insertArgument()* or *insertAttack()* easily fall into the 1000 transactions per second range. However, *pafReductionToAfPr()* and *calculatePreferredExtensions()* methods require more computation and might limit the transactions per second number. Regarding the Negotiation contract, the *newOffer()* method might highly influence performances when the number of issues is > 25.

The use of sidechain allows agents to operate without too many performance limitations, while maintaining a level of traceability that allows full auditing by an inspector. These results would not have been possible in a permissionless DLT. In fact, for example, in Ethereum the limit of gas usage per block is currently (at the time of writing this chapter) 15 million gas units. This means that, not only some transactions could not be executed (e.g. *calculatePreferredExtensions()* with an AF with > 20 arguments), but also that the latency between operations would be very high because currently, in the Ethereum network, a block is created every 10/15 seconds on average.

5.7 Discussion

In this section, we discuss how the IHiBO we proposed might have particular relevance in cases where the decision making process about what data should be fed in the smart contract needs to be transparent: for fund management, the investors don't know what exactly happens to their money, and especially why, so the question whether the fund managers do fulfill their legal and ethical commitment of acting in the best interest of the investor might remain unanswered.

In general, the transparency that can be gained due to the proposed intelligent oracle architecture could be highly valuable in any trust services. The concept of the fiduciary is based on—as the name of these services shows—trust. It requires fiduciaries to be bound both legally and ethically to operate and use their expertise in the best interest of the investor. It also requires trust on the part of the investor to believe that the fiduciary has done so and will do so. This trust can be, to some extent, replaced by intelligent, decentralized solutions providing full transparency of, for instance, fund management: not only the transactions can be fully traced but the expert opinion input and the decision mechanism too. By implementing argumentation and negotiation phases through oracles into smart contracts or making them on a side-chain can generate more transparency for investors: investors can know how the final decision is made at the end of reasoning. This could be highly relevant for the investor practicing his right to check the fiduciary's activities in the case of an asset management contract.

5.7.1 Trust

IHiBO can enhance trust in several ways.

As argued by Walton, it seems to be more generally acknowledged now that we do have to rely on experts, and that such sources of evidence should be given at least some weight in deciding what to do in practical matters [325]. In our case study, managers play the role of experts and the professional certificate of them as well as their past creditable experience could be part of the backup of trustworthiness of the source information, and we calculate the weight of the arguments in the parallel of voting theory, i.e. to count the number of supporting managers.

Another way to gain and restore trust from investors is to make the resources and decision-making process explicit, our case can be considered as a good example of the use of argumentation for favoring trust. Being skillful and sophisticated could be not enough for the requirements of managers. Especially when they are in a corporation, other problems may arise to obtain trust, like reliability and agency problems. For instance, problems arising from managers' unwillingness and lack of incentives to act in the principal's best interests, rather than from a lack of expertise. In our case design, investors have the advantages to audit the resources of the information, thus such risks could be mitigated. Falcone and Castelfranchi relate trust explicitly to the goals of agents, and consider trust to be concerned with whether another agent can and will perform an action that will enable the first agent to achieve its goals [124]. In the case of fund management, fund managers are sharing the same goal——gain interests for the investors. In the study case, agents must coordinate and communicate with their own information to reach an agreement. In this scenario, the requirement to reach trust is to ensure and audit the trustworthiness of a source of information within an argument which is then to be decided to be accepted or not. We ensure the trustworthiness of the information by counting the values or the numbers of support from agent to arguments to ensure the resources based on somehow voting theory.

The adoption of blockchain and DLT has been under consideration for several years both from economic and legal aspects [160, 267]. However, most of them only consider the transaction process, i.e. how to use these technologies for clearing and settlement, and some propose to use smart contracts to conduct the functions of CCP or central securities depository (CSD) [7] [233]. In our work, we pay attention to the pre-trading phase, where the investment decisions made by the trust services are extremely crucial to investors. As discussed above, the decision-making process is traceable and immutable on blockchain. As a result, the entire reasoning decision and transaction process are transparent and investors can gain maximum confidence and thus trust for the trust services.

5.7.2 Legal Considerations

Next to the technical and financial aspects, legal considerations should also be taken into account when comparing the different architectures. While our motivation is to provide transparency regarding the decision-making process to the principal to gain some insights whether the work of the fiduciary indeed happens according to his best interest, the transparency one should gain with using DLTs is subject to serious limitations.

On one hand, the the principal's right to check is not limitless, it concerns strictly the processes of managing his assets, but more importantly, given the characteristics of DLTs, a(n unwantedly) broader audience would be involved in the disclosure of information if one chose not the appropriate architecture, threatening trade secrets and involving privacy problems.

On the other hand, once the application of DLTs becomes widespread in the securities market, mandatory disclosure rules motivated by anti-tax avoidance should be aligned with the new technology [307]. Indeed, DLT-based automated disclosure may lead to the release of information that is too fast, limiting the ability of investors to properly speculate. Thus, mandatory disclosure requirements would still be necessary, but the enforcement of such provisions and detection of violations redesigned using DLTs and smart contracts would have to deal with the necessity of stakeholders.

The final IHiBO architecture (in contrast to the other two discussed in comparison in Section 5.5.1) seems to be the best option from these point of view too: in contrast to the public, permissionless verification that DLTs usually employ while smart contracts are executed, layer two solutions usually move this process off-chain. This definitely poses

[7] CSDs operate the infrastructure that enables the securities settlement, allow the registration and safekeeping of securities, allow the settlement of securities in exchange for cash, track how many securities have been issued and by whom, track each change in the ownership of these securities

security issues compared to a protocol executed completely on-chain, however there are currently some viable solutions proposed that address this issue [126].When using a layer two technology, information that would clash with trade secrets and privacy can be stored on that permissioned chain and maintained by the participants who have been nominated for this, e.g. joint data controllers as permissioned blockchain operators [201]. Through the use of commitments on the main chain [155], i.e. the permissionless one, the necessary steps for verification are implemented, and once the fiduciaries operating the sidechain reveal part of the information to the principals, the latter can verify its validity on-chain [296].

We argue that IHiBO's final architecture provides the proper middle ground in terms of cost of execution, for latency and fees, and verifiability of the complete process. Indeed, there might be use cases where some data should not be disclosed, and an argumentation and negotiation architecture based on a full execution on smart contracts would not allow it. In the other extreme case, for a centralized oracle, the entire process behind a decision could be concealed or its log could be altered. Members of the management body[8] shall have adequate access to information and documents which are needed to oversee and monitor management decision-making[9]. In IHiBO, each execution of all the smart contracts can be audited, validated and maintained by every participant, thus reducing the time and fee of extra work of surveillance, which will in turn reduce potential corruption or conflicts of interests.

5.8 Related Work

To the best of our knowledge, there is no mature work on adopting argumentation in the financial world, apart from the work of Palmieri [239], where argumentation is provided as a tool to gain the stakeholders' support and trust. Moreover, in literature we do not find any methodology that is a hybrid of decision-making based on formal argumentation, autonomous negotiation, DLTs, smart contracts, and oracles. For this reason, we approach the related work from multiple perspectives.

Regarding the use of formal argumentation with the aim of entrusting processes, we find different works [300, 305]. In the earlier works on argumentation theory, researchers only focused on the relation among arguments [272]. Since then, many scholars have included the trust component in the evaluation process of arguments [206, 244, 322]. Parsons et al. [244] suggest that argumentation might play a role which tracks the origin of information used in reasoning, thus it can provide provenance in trust. Later the same authors develop a general system of argumentation that can represent trust information, and be used in combination with a trust network, using the trustworthiness of the information sources as a measure of the probability that information is true [304].

On the other side, we find the new concept of "trustless trust" that, first the Bitcoin blockchain and then other DLTs, have brought in finance [267]. Different papers have provided ways to construct securities based on DLT, however, they only address the trans-

[8]Art. 4(8) MiFID II: 'management body' means the body or bodies of an investment firm, market operator or data reporting services provider, which are appointed in accordance with national law, which oversee and monitor management decision-making and include persons who effectively direct the business of the entity.

[9]DIRECTIVE 2014/65/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU

action process on the clearing and settlement side, which consists of a process after securities are negotiated [233, 324]. Thus, not paying attention to the pre-trading phase as in our case. Also infrastructures, such as the Decentralized Financial Market Infrastructures (dFMI) [126], have been proposed. These are built with a vision of a peer-to-peer (P2P) consortia whose members are many financial markets' participants and where the organization is managed through the P2P infrastructure itself, rather than through a central intermediary. In this case, smart contracts can take on a role similar to the one previously played by central counterparties (CCPs) [267]. For instance, smart contracts can be used to act as a margin calculation agent and to take on the task of transferring collateral. Although in a different way, the smart contract can also be used to resolve disputes in the event of non-compliance with payment [221]. Alternatively, instead of replacing CCPs, smart contracts can support them to calculate and update collateral as well as manage funds. A concrete application of dFMI is Lianjiaorong, a blockchain AssetBacked securitization platform, built by the Bank of Communications in China [240]. The blockchain is maintained by original stock holders, such as trust companies, investors, rating agencies, and it links funds and assets on the ledger, realizing the credit penetration of the securities business system. dFMI can be rendered completely decentralized and therefore not based on the traditional financial sector. This is the case of the Decentralized Finance (DeFi) [328]. With DeFi protocols such as Decentralized Exchanges (DEX), anyone can engage in a non-custodial exchange of on-chain digital assets, e.g. tokens. This new decentralized paradigm can pave the way for new approaches in finance, but at the same time provide a risk to its users due to its unregulated nature [30].

Lastly, with reference to DLT-based solutions with the application of oracles, many works and commercial solutions have been proposed in past years. Between the DeFi oracles being used in practice, we find MakerDAO [197] and Compound [191] cryptocurrency lending platforms. MakerDAO is the first and most popular one, in which each type of cryptocurrency collateral has a corresponding oracle that feeds real-time price to smart contracts. Both solutions collect the prices from multiple sources and aggregate them as an average price. Aside from finance, oracles are usually implemented for general-purpose data. Provable [269], is a blockchain agnostic oracle service that provides a data transport-layer for smart contracts to fetch external data from Web APIs. Chainlink [121], instead, offers a similar service but implemented as a decentralized system for Ethereum. Human oracles, i.e. the ones requiring an input which involves human intervention, are rarely applied [100]. The rare existing ones are deployed in applications with binary inputs, i.e. they only take input by one of two possibilities, typically "yes" or "no" [225], such as ASTRAEA [2] that leverages human actions through a voting game. Augur [46] is a decentralized oracle that needs specific human users obligated by *Reputation Tokens* to report outcomes at specific times. Gnosis [144] approach is different, as it derives information from centralized oracle services, but enables the users to challenge those results.

5.9 Summary

In this chapter, we have proposed IHiBO, a framework which integrates formal argumentation and negotiation within a DLT environment. These techniques have distinctive features that complement each other. They together can make a decision-making process more transparent and traceable. We focused on the concept of trust both from an implementation and a legal point of view. Our motivation came from trust services that let us describe our system proposal through a fund management scenario. However, IHiBO is not bound to this domain and can be implemented in different use cases.

We argue that our methodology can help enhance principals' trust towards trust services. This is because, when knowing how the fund management makes decisions sufficiently well, the behavior of managers can be understood and predicted more accurately. Moreover, we dealt with the research on oracles, that is still in its infant stage for what concerns implementing transparent and decentralized decision processes. Thus, there are multiple pressing questions and challenges for future work. To the best of our knowledge, this is the first study where such a framework that incorporates argumentation and negotiation is implemented using a cross-chain oracle and smart contracts. The results of our experiments shows that the use of a two layer blockchain architecture, allows us to securely operate without too many performance limitations, while maintaining a high level of traceability that allows us to audit trust services operations.

For future work we plan many extensions of IHiBO. A possible follow-up is to provide and adapt to a high level of adaptability in the decisions of the fund management, e.g. to define different investment scenarios considering attitude (aggressive or moderate) and the financial environment (e.g. bull or bear market) Another possible extension could be to integrate consensus mechanisms with argumentation as well as negotiation, since there is no specialized DLT, yet, that integrates reasoning in its protocol. For instance, if there is a blockchain based on Proof of Stake, validators need to vote to validate a transaction based on a reasoning process where each validator has a different set of knowledge data. Lastly, we also plan to rely on the recent advances of the domain of Explainable AI, in order to render the decision-making process presented in this chapter explainable for different types of users (experts, non-experts, etc.) and for different purposes (e.g. transparency, debugging, etc.).

Chapter 6

Risk Analysis and Regulatory Compliance of DLTs for Financial Securities

In Chapter 5, we introduced the IHiBO architecture as a robust framework for storing decision-making processes in fund management, leveraging the power of DLTs. However, the transformative potential of DLTs extends well beyond the pre-trading phase. In this chapter, we venture into the realm of securities trading facilitated by DLTs, an area that garners considerable attention and exhibits greater activity compared to DLTs integrated with AI. Our specific focus lies on the critical aspects of risk analysis and regulatory compliance intertwined with the adoption of DLTs for securities. By meticulously exploring these pivotal dimensions, our objective is to foster a comprehensive understanding of the multifaceted challenges and compelling opportunities that arise from employing DLTs in the securities trading domain. Notably, this chapter also serves as a future-oriented exploration for IHiBO, as the transactional phase naturally follows the decision-making phase in the investment process.

6.1 Securities Market without DLTs

To ensure this chapter is self-contained, we will initially provide a succinct introduction to the functioning of securities markets, explore the life cycle of securities transactions, and discuss the drawbacks.

6.1.1 Investigating Securities: From Definition to Transaction

The term ”security” means a negotiable and fungible tool that carries a monetary value of certain times. It can manifest that a position of ownership in a public trade corporation is presented in the form of stocks; a creditor’s relation with a governmental organization, or a firm represented by holding of entity’s bonds; or ownership rights in the form of options [127]. Securities of public trade are listed on stock exchanges. There, issuers seek security listings and draw attention from investors by guaranteeing a regulated and current market for trading. Informal systems of electronic trading have been further generalized in recent years, and at present, securities are often subject to over-the-counter (OTC) trade or conducted between investors via cell phones or the Internet [109]. An initial public offering (IPO) refers to the first major sale delivered by a company of equity securities to a public market. Based on the IPO, all stocks issued and still sold in a primary market can be called a secondary offering [168]. Or, securities may be privately offered to a qualified and strict group known as a private placement. It represents a significant distinction concerning securities regulation and laws. Sometimes companies put stocks on sale through the supplementation of a private and public placement [280].

In the secondary market, which can be called an aftermarket as well, securities are just taken as assets for transfer between investors: shareholders may sell other investors the securities in the form of capital gains and/or cash. Hence, the primary market is supplemented by the secondary market which has lower liquidity for securities under private placement because they are only applicable to qualified investors' transfer and cannot be traded in the public [117].

6.1.2 Life-cycle of Securities Transaction

The whole life cycle of securities transactions is rather long, involving trading, clearing, and settlement, with numerous intermediaries. In the trading phase, both the seller and the buyer of the financial instrument need to inform their brokers with demanded access to and knowledge about the stock exchange and seek counterparties on a trading venue, e.g. a regulated exchange, Regulated Market (RM), Multilateral Trading Facility (MTF) and Organised Trading Facility (OTF) are three main categories of trading venue [28]. Besides, this agent will charge fees as a commission. Alternatively, the trade can take place OTC [220]. Following the trade, the post-trade process workflow encompasses all processes, intermediaries, and infrastructure from the agreement of a financial security transaction to its final settlement [192]. It is comprised broadly of three key functions: settlement (final asset transfer), clearing (counterparty obligation calculation), and order management (including trade validation). To be more specific, in the clearing, behaviors between the trade date and settlement date are managed. It can be carried out formally by a central clearing counterparty (CCP) or can be conducted directly in an informal manner by a buyer and a seller [11]. In CCP clearing, the CCP becomes a buyer to any seller or a seller to any buyer. Therefore the counterparty's risks are transferred to CCP, namely to the trade from actual parties. In clearing fields, EMIR, namely the European Market Infrastructure Regulation of 2012, stipulates obligations of clearing, CCP regulations, and obligations for reporting towards Trade Repositories. At present, the regulations on Level 2 are implemented (Regulation (EU) No 648/2012 of the European Parliament and of the Council of 4 July 2012 on OTC derivatives, central counterparties, and trade repositories Text with EEA relevance). In settlement during a post-trade flow, a buyer receives securities bought, and a seller gains corresponding cash for the securities. As intermediaries of investors, brokers, and banks join the process of trade settlement of securities in the form of book entry and also enable access to central security depositories (CSD). In the securities settlement field, the CSD Regulation (CSDR) takes charge of securities settlement improvement especially (such as decreasing the settlement cycle to T+2) as well as CSD regulation [318]. Asset service and custody constitute asset safekeeping conducted by CSDS, brokers, and intermediary banks on behalf of investors, besides execution of asset service supply, including proxy voting services, tax reclamation, firm action processing, and income collection.

As defined, custody services are auxiliary functions in the Markets in Financial Instruments Directive 2 / Regulation (MiFID 2/MiFIR) [1], while alterations to the Shareholders Rights Directive, which was proposed by the Commission in April 2014, in particular, expand the responsibilities of intermediaries. Thanks to the practice of private sectors, optimal market practices have been generated based on market norms of company actions processing, by which general meetings can better execute rights flowing from the held securities. At present, all these practices are being carried out in all markets of Europe. Measures taken by the investment fund field (exchange-traded funds, unit trusts, mutual funds, etc.) also have an important influence on post-trade processes. Specifically, the European Fund and Asset Management Association (EFAMA), the European Fund and Asset Management Association (EFAMA), legislation, notably the Undertakings for Collective Investment in Transferable Securities, UCITS [2], and the Alternative Investment Fund Managers Directive, AIFMD [3] are carried out. Remarkably, swapping assets for cash creates a risk for a party to breach the agreement after trading is partially underway. CCP helps the mitigation of the risk for both a buyer and a seller and calls for another system, which transfers the risk instead. This institution controls the funds and assets which are not yet transferred to both sides.

6.1.3 Drawbacks of Traditional Securities Transaction

The current structure of the financial industry is dominated by centralization. Trade chains and life cycles of custody may be very long, while many intermediaries organize their own dedicated databases and overlapping transaction information so that significant duplication is generated [267]. Redundant and time-consuming processes due to securities being managed independently by each counterparty. Some studies have pointed out problems in the industry of post-trade settlement, which is concluded as follows.

- Lack of interoperability between siloed database systems. The current structure of the financial industry structure is dominated by institution centralization. The custody chains and trade life cycle may be very long, while a lot of intermediaries have their own dedicated databases providing overlapping information for trading, resulting in a lot of duplications. Those who participate in a post-trade value chain often have to update digital records for coordination with any change taking place to counterparties' records in a holding chain at a different level, so that great operational risk can be generated [247, 285].

- Unnecessary complexity and inefficiency of trading. A lot of scholars have criticized the low efficiency of most financing instruments that are applicable in the market,

[1]Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU Text with EEA relevance

[2]Directive 2009/65/EC of the European Parliament and of the Council of 13 July 2009 on the coordination of laws, regulations and administrative provisions relating to undertakings for collective investment in transferable securities (UCITS) (Text with EEA relevance

[3]Directive 2011/61/EU of the European Parliament and of the Council of 8 June 2011 on Alternative Investment Fund Managers and amending Directives 2003/41/EC and 2009/65/EC and Regulations (EC) No 1060/2009 and (EU) No 1095/2010 Text with EEA relevance

which is generated due to high trading fees and deviation from the attributes performances manifested by infrastructure assets [98]. Investment funds of infrastructures are often related to high friction (trading) fees concerning highly customized direct investments, cost of fund management (e.g., carried interest), and other expenses or fees concerning management and operations. A large number of intermediaries (such as swap suppliers of interests, insurance agencies, rating agencies, and banks) can also explain why the transaction efficiency decreases when traditional financing tools are used because it takes several days to settle a single transaction under the engagement of multiple intermediaries in general. In addition, the heterogeneous and unique development of infrastructure needs a complex and comprehensive legal arrangement to ensure high operation efficiency [91]. Legal arrangements can also effectively clarify the responsibilities of both parties, ensuring proper distributions of salary, and building an appropriate mechanism for risk-sharing. However, legal arrangements need a lot of intermediaries. Trading of listed infrastructure asset securities can spend as long as 2 days (T + 2) [313].

- Lack of transparency. There is always information asymmetry in infrastructure investment, from the procurement process to the stages of maintenance and operation [122]. The information investors need for understanding a project is very scattered and opaque. Moreover, the highly customized and complex nature of infrastructure investments spends a lot of resources and time for investors to control and understand those risks. Conventional tools of infrastructure financing fail to cope with these problems. Information asymmetry is also related to risks concerning unfair competition, inadequate governance, and widespread corruption, all of which can greatly reduce the investors' interest obtained by infrastructure assets.

6.2 DLT-based Securities Market

At the beginning of the advent of DLT-based digital finance solutions, regulators were focusing on the primary market phase because most regulation was spurred by the onset of Initial Coin Offering (ICO) [130]. As of now, however, significant attention is paid also to the "secondary market" aspect, as a way to fully enable the development of a primary market: if the secondary market does not have liquidity and does not support DLTs and DLT-based instruments, then the primary market cannot develop. Despite ICO frauds in 2017 and 2018 [341], DLT-based tokens have been used increasingly in the market of finance, and the tokenization of assets has become one of the most typical cases of application for DLTs in the market of finance. Tokenized assets include securities (e.g., stocks and bonds), but also commodities (e.g., gold), and other non-financial assets (e.g., real estate).

Asset tokenization may have integrated implications for participants and practices in the financial market, market regulators, and infrastructure in diversified financial directives and types of assets. Within the fintech community where many proponents believe in the potential for DLT to facilitate the development of certain securities markets without trusted intermediaries by market participants. Settlement and clearing as the first step for studying potential impacts brought by DLT to payment behaviors, are carried out. Thus, before we go further into applying DLT for securities trade, the clarification of tokenized securities which are the liquidity subjects is first conducted. Compared with traditional databases which are distributed, because of the three features, DLT becomes an attractive candidate for post-trade industrial application. Those distributed databases can provide synchronized and shared copies of a ledger, which can reduce the number of reconciliations, while decentralization would allow those nodes to better control relevant information of their own, and could bring down costs further by facilitating the post-trade process coordination. In the end, traditional distributed databases depend on trusted nodes and maintain the ledger's copies in a security perimeter the central administrator controls, the ability of safe operation in a competing context with un-trusted third parties plays a critical role in any potential application of the financial industry.

Financial institutions have not yet demonstrated DLT is a sustainable and valuable solution. It could be unclear based on pilot cases if trading and after-trading segments will be further intertwined [267]. Also as of now, institutions are generally focusing on applying DLT to a part of the process, because there is not enough regulatory clarity yet in terms of the regulatory treatment of tokenized securities to implement a wider solution. "No full-scale DLT system is completely the same. Hence, the wider application of DLT concerning business model influences is not yet known" [267]. In general, scholars and industry experts have put forward that "it is widely known in the industry that DLT operationalization can be a gradual other than a booming revolution. In fact, market participants may not intend to write off investments in current technologies at a quick pace. In other words, market participants might focus on segments for the realization of the most efficiencies (i.e. segments with a lot of manual interventions, timelines that are long and expensive, and/or there are obvious potential errors). Afterward, they might focus on the whole life cycle of trade" [267]. As a result, it is unclear which DLT system is best suited for the clearing and settlement of securities in terms of operational functionality. Since most of the inefficiencies are in the post-trade segment, which DLT system is best suited for the settlement and clearing of securities is still uncertain in terms of operational functionality.

6.2.1 A non-exhaustive List of Application

In 2015, Nasdaq launched Linq to make the issuance of private securities possible [245]; a platform based on DLT for trading and issuance of private company shares [345]. Private firms adopted the Nasdaq Linq blockchain to present their share ownership digitally via DLTs, or record and complete private securities transactions by this blockchain. In 2016, Overstock.com initiated a closed-system platform of trading for its proprietary blockchain [284]; Overstock.com Inc., a company with a public listing on Nasdaq issued public securities of a new class directly on the blockchain, which exist only there and using a transfer agent. Blockchain Voting Series A Preferred Shares were completely consistent with regulatory requirements, raising total gross income by about USD10.9 million. The shares provided the same-day settlement [185]. Switzerland's stock exchange, managed and owned by SIX, is establishing all-integrated custody, settlement, and trading infrastructure for digital assets, namely SDX [330]. It will provide a safe surrounding for trading and issuance of digital assets, which makes tokenization of non-bankable assets and existing securities 'make previously untradeable assets be traded. London Stock Exchange tested the foundation, admission, and issuance for equities trading, with 20|30 becoming the first British company which completes the issuance and tokenisation of their equities as part of the UK Financial Conduct Authority's (FCA) Sandbox22 [3]. Australian Stock Exchange

(ASX) The ASX has tried overhauling the Clearing House Electronic Sub-register System (CHESS) as of 2015 [175]. The Australian securities market gains a market value of AUD 1.6 trillion, making it the 15th largest market in the world [133]. ASX cooperated with Digital Assets in DLT application for settlement and clearing of equity transactions [213]; They drew a roadmap of implementation to replace the CHESS settlement system with a clearing system based on DLT and being executed in 2021 [85]. This action will realize efficient reconciliation and trade settlement and while reducing operational costs. The French central securities depository (CSD) 'ID2S' applies the DLT in the issuance of French commercial paper [268]. Canadian Securities Exchange builds a DLT securities settlement and clearing platform to allow company issuance of equities and fixed-income securities based on security token offerings. Similar projects are able to test the benefits and feasibility of DLT application in securities settlement have been found in other jurisdictions [214], like the Bank of Japan (BOJ) joint research project STELLA and European Central Bank (ECB) which presented conceptual analysis and experimental results of DvP success for singular and crossed ledger(s) [178].

6.3 Analysis of DLT-based Securities Market

In the ensuing section, we present a detailed analysis of the Distributed Ledger Technology (DLT)-based securities market. This emerging landscape is reshaping the financial industry, offering a new paradigm for conducting securities transactions. However, like all transformative technologies, DLT is not without its challenges. Our analysis thus provides a balanced view, highlighting the potential advantages offered by DLT in increasing efficiency and transparency, while also addressing the risks and challenges that come with its implementation. This exploration will contribute to a comprehensive understanding of the DLT-based securities market, providing valuable insights for stakeholders navigating this innovative territory.

6.3.1 Advantages

The use of DLT can accelerate and compress transaction settlement and settlement to almost real-time, reduce the risk of counterparties and release collateral, and potentially generate capital efficiency for transaction participants [323]. It can simplify the multi-step process after the transaction and significantly reduce the back-office management burden. However, the experimental application of DLT in clearing and settlement has produced mixed results, there is a need to overcome the obstacles in the development of technology to make the application reach the stage of providing better performance than the current system [132].

Increased efficiency in trading. Through the combination of smart contracts, dividends, trading, and interest distribution, information storage and management can be automatically executed in the blockchain through triggers encoded in contracts. Value proposition requires the least involvement of regulatory, financial legal, and other intermediaries, thus reducing the relevant transaction costs. Since token trading is done through Ethereum on a point-to-point basis, the transaction time is reduced. The token transaction based on Ethereum will be determined in a few seconds. Therefore, the use of DLT can speed up and compress the short time of trade settlement and settlement, reduce the risk of counterparties and release collateral, and generate capital efficiency for trading participants. It simplifies the multi-step process after the transaction and greatly reduces the burden of background management. However, the experimental application of DLT in clearing and settlement has produced mixed results, so we need to overcome the obstacles in the development of technology to make the application reach the stage of providing better performance than the current system.

Improved transparency. Distributed databases, democratically owned by nodes around the world in permission-free or licensed systems. In the former, no entity or individual can arbitrarily change the distributed ledger. In addition, the consensus among communities provides a second layer of protection against data manipulation. As a result, token transactions that take place on the blockchain are considered immutable. The legal rights and ownership of the tokenized infrastructure assets are directly embedded in the security tokens, and the owners of these tokens can participate in the decision-making process by exercising the voting rights they acquire along with the tokens. The identity of the seller and buyer of the token, as well as the full transaction history stored in the blockchain, is immutable and easy to track.

Fulfillment of Know Your Customer (KYC). While some argue that DLT systems increase anonymity and may increase criminal activity, the opposite may still be true. In a licensed blockchain, an administrator with a well-established startup process can easily meet KYC needs. At present, the KYC process chain involves exchanges, clearinghouse members, and brokers. DLT's implementation of corporate information releases will increase speed and transparency in disseminating information to the wider public. This is done through public-key cryptography and distributed networks.

6.3.2 Risks and Challenges

The introduction of DLT could bring a lot of benefits when it is used in settlement and clearing, yet, many academic research and business use cases show the applications still face challenges.

Privacy breach. For higher trust and transparency of a DLT system, all system participants typically observe all information about transactions in the ledger and copy it to their ledger. When applied to financial markets, such transparency may raise privacy or competition issues, thereby contravention of applicable laws, such as the General Data Protection Regulation [4]. All participants would know every transaction and existing as well as their details, including volumes and values of the assets integrated [285]. Certain solutions including encryption techniques and advanced obfuscation are now tested to enhance the privacy of participants together through economic avatars and tokenized identity [54].

Settlement Finality. In post-transaction clearing and settlement, the settlement deadline is currently a legally defined moment, usually supported by a regulatory, legal, and/or contractual framework based on a given financial transaction [326]. When the trading parties and their intermediaries update their account books to implement the settlement, determine asset ownership, and measure and monitor various risks [290], they rely on the

[4] GDPR: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

definition and timing of the final results. On the contrary, in some versions of DLT arrangements, multiple parties are allowed to update the shared ledger, and they must agree to the specific status of the ledger through a consensus process. In this world, the final result of settlement depends on the final result of probability, that is, the system participants believe that the longer a transaction is settled, the less likely it will be reversed (or deleted) [308]. This final approach contrasts with the traditional, well-defined, and transparent final moment approach. In such a network, if the final result is determined by the method of probability, the legal liability may be difficult to assign or ambiguous, and this uncertainty has an impact on the balance sheet of the participants and the rights of their customers and creditors

Scalability. Securities settlement requires the ability to handle a large number of transactions safely and reliably [4]. In addition, any settlement platform should be able to cope with the potential growth of trading volume when the market is tight. This further limits the choice of verification protocol and transaction processing speed to a certain extent.

Management of Identity. The private licensing structure seems to be the most popular securities settlement method in the industry [1]. Identity and access management is currently focused on several trusted entities (CSD, CCPs). These same entities can perform similar functions using centralized distributed ledgers. They will verify identity and credentials. However, if the ledger is to be decentralized, it is essential to find processes that can effectively control identity and access management and protect them from attacks. In addition, if authentication is performed by encryption keys, there should be a trusted mechanism for issuing and replacing those keys [292].

Matching of Trade Matching and Management of Errors. As argued, significant obstacles exist in the application of DLTs to the matching of trade [47]. This is because the factor DL does not need to compare different data fields, handle contract mismatches, or handle exceptions. Compared with secure transactions, the case of Bitcoin is less functional. In secure transactions, it may be necessary to match a large number of attributes with complex rules and cross-dependence. As believed by some commentators, central reconciliation will still be needed as a pre-processing of ledgers. In addition, due to the invariance of the ledger and the lack of a central administrator, the already complex aspect of exception management in financial transactions becomes more complex.

Confidentiality. Despite rules for consensus realization, multiple participants are involved in the validation typically, so there is conflict against confidentiality requested in security trading, therein, transaction contents must be invisible to everyone, other than the sides involved in the transaction [20].

6.4 Legal Constraint and Regulatory Challenges

Financial markets with the adoption of DLT should obey the regulation-related requirements that enhance financial consumption, investor protection, and an integrated and competitive market and prevent the increase of systemic risks [55]. From the perspective of securities market application, the risks and challenges brought by blockchain to regulation are mainly reflected in the following three aspects [237]. The first aspect is related to the adaptation of DLT-based securities issuance, trading as well as the current regulatory system. The application of DLT will shift the issuance and trading of securities from a "centralized model" to a "decentralized model" distributed in a variety of cyberspace. The changes and innovations formed in this process are likely to conflict with the existing regulatory framework. It will influence the traditional form of securities rights representation, leading to the dilemma of securities rights proof regulation. Using DLT to issue securities, traditional securities will be replaced by digital assets as the subject of issuance. The trading of digital assets is also different from the process of traditional securities trading. It is difficult to apply the current regulatory system, including investor suitability, separation and custody of securities accounts and fund accounts, information disclosure, and lock-up period effectively.

Furthermore, it is about the impact of "disintermediation" and "de-trust" on the structure of the securities market [215]. The "decentralization" will reduce the role of securities registrars and settlement institutions in the operation of securities markets. If the securities registry and settlement mechanism are changed, the securities trading regulatory model, which is based on the traditional securities registry and settlement infrastructure also needs to be changed accordingly [293]. With the application of DLT, intermediaries, which are created to eliminate information asymmetry, will gradually lose the need to exist under the mechanism of "de-trusted" operation, and the issuance, trading, registration, and settlement of securities can be completed directly, and the service functions of intermediaries in underwriting, brokerage, registration, clearing, and settlement will be weakened. As a result, the "gatekeeper mechanism" of securities intermediaries is missing, which increases the pressure of administrative supervision and self-regulation to a certain extent.

A proper regulatory scheme is put in place to deal with it. When securities are identified and traded in the form of digital assets, investors may suffer significant losses in the event of theft of digital assets, implying that the securities and the rights to the identity and property they represent will be lost [332]. If the use of blockchain technology does not find an effective balance between increased transparency and investor protection, and the relatively mechanical technical processing cannot adapt to the specific requirements of different environments during the operation of the securities market, then it may provide room for moral hazard and fraud to breed [297]. On the one hand, the tamper-evident nature of DLT ensures the security of transactions; on the other hand, it greatly increases the difficulty of modifying transaction information. In addition, the loss of the private key may prevent the securities investors from operating the assets under their accounts, and the potential risk of loopholes in the process and the applicability of smart contracts in practice may hinder the application of DLT in the securities market. Regulators should guarantee that the securities markets based on DLT are in line with the regulatory objectives, including the improvement of financial stability, protection of financial customers, and ensuring of market integrity. With respect to a few jurisdictions that adopt a technology-neutral regulatory approach, it may be necessary for the existing regulations to be applicable to new products and actors, and for regulators to establish new requirements so as to solve the new risks arising from the new feature of a few business processes and models [25]. It is mandatory for DLT-based markets to obey regulatory requirements with the aim of reinforcing financial stability, protecting financial customers, ensuring market integrity, and enhancing competition. The digital technology which is employed for the security custody, i.e., electronic book entries in securities registries of central securities depositories is possibly substituted with another one, i.e., the cryptography-enabled dematerialized security on the basis of a DLT-enabled network; consequently, it will not be a problem in jurisdictions that adopt a technology-neutral regulatory approach.

In the regulatory handling of tokenization and DLT, potential gaps may cause opportunities for regulatory arbitrage. The existing oversight needs to be applicable to new actors, for example, the trusted third party that ensures the information accuracy at the asset on-chain onboarding and protects the asset, and/or the introduction of new requirements are necessary, for example, the requirement involving the interoperability between DLTs or the interaction or gateways connecting the on-chain environment and the off-chain environment. In addition, it is also necessary to properly monitor the new risks potentially caused by the application of DLT technologies, for example, the related operational risks and digital identity. Since investors are confident about the proper protection of their legal rights, the perceived uncertainty in terms of the crypto assets' legal status and the enforceability of the smart contract based on the private law possibly hinder the broader use and transaction of these assets. Especially, the uncertainty about whether crypto assets are qualified as property according to private law, and whether contracts that are written in codes cause the binding legal responsibilities can be perceived. These questions matter because crypto assets cannot be owned if they are not recognized as property as per private law. Likewise, if smart contracts do not lead to binding legal obligations, the transacting side's rights cannot be implemented if a technology failure occurs.

6.4.1 Legal Questions Raised by Applying DLT to Securities Markets

The integration of DLT and smart contracts within our legal and financial systems surface a myriad of legal queries and concerns. As these technological innovations intersect with legal precedents and regulations, they prompt us to rethink our existing legal frameworks and challenge our understanding of contractual obligations. The following are the principal legal questions that arise in the context of DLTs and smart contracts.

Data breach. To improve transparency and trust in a DLTs system, all system participants typically observe all transaction information in the ledger and copy it to their ledger. When such transparency is applied to the financial market, it might raise privacy or competition that would violate applicable laws, such as the General Data Protection Regulation [238]. All participants should recognize the existing deals and detailed information like the corresponding value and quantity of assets.

Regulatory or legal ambiguity is raised by definitions of new emerging entities. The definition of assets tokenization can introduce risks and uncertainties to participants in the token market, which undermines the smooth operation of such markets, potentially indirectly influencing the traditional off-chain market [157]. In terms of the safe development and use of tokenized assets and markets by market participants, it will be a stepping stone to improving the transparency of the regulatory or oversight framework in the application of tokenized assets and markets. International collaboration is required in the transactions of tokenized assets across borders so as to constrain the regulatory arbitrage and guarantee the functioning of tokenized markets.

The legal status of smart contracts. Although contract law is a fairly developed field of law and the freedom of contract is broadly recognized in most of jurisdictions, the self-executing software automatically and autonomously executing the terms of the contract on a peer-to-peer and immutable basis still has some legal problems, namely, whether and how the contract law should be revised to make smart contracts valid and enforceable, considering their deterministic and automatic features. Whether smart contracts will eventually substitute the existing legal contracts entirely, or whether they can only be employed to automate the executions specified in legal contracts needs to be further reflected [13].

Legal ownership. Once securities and cash change hands, the settlement requires a formal or legally defined sign of ownership transfer. Therefore, an important question to be addressed is whether a DLT entry can legitimately constitute proof of ownership.

Since DLT has been adopted on a large scale in the market of securities, a number of technical challenges including settlement finality, scalability and cyber risks, governance risks, including Anti Money Laundry (AML) or Financing of Terrorism (CFT) [182], and the problems of digital identity, privacy, and data protection and the questions about smart contracts' legal status will appear.

The large-scale adoption of DLT would bring a series of challenges related to the fundamental technology itself. Given the global financial market requires significant throughput, appropriate levels of privacy, settlement results, inter-platform interoperability, and buffers against network risks and hackers, it is necessary to solve the resolution of technical challenges around scalability, faced by the widespread adoption of DLTs. The smart contract's legal status has yet to be determined, and until it is clear whether the contract law is applicable to smart contracts, the issues of financial protection and enforceability will remain. The potential transformation of financial products and markets into a tokenized DLT-enabled environment is not hoped to occur anytime soon even by the most significant advocates of blockchain technology. It is easier to envisage a gradual shift to a DLT-based market, giving priority to those processes that are most likely to improve efficiency. The tokenized market is likely to flourish as a supplementation to the current traditional market, for some processes accompanying the security lifecycle like post-trade.

Given the above, a discussion on the high-level risk analysis of DLT and smart contracts in general for the securities market is warranted. By carefully considering the possible effects of DLT and the use of asset tokenization, policymakers can predict potential risks related to the broader use of DLT. The technical feasibility, business savings, and cost efficiency brought by securities market disintermediation still need to be fully evaluated and quantified by means of practical application. An understanding of the theoretical obstacles to the potential cost-effectiveness of DLT-based clearing settlement may include, for example, the fact that the application of DLT in the post-trade process may not be complete and comprehensive for the entire process: back-end settlement is still required in post-trade clearing and settlement. Full efficiency cannot be achieved, if other activities affecting the position, payment, or delivery of securities, for example, securities lending or derivatives, are not based on the same kind of technology. In other words, in this case, DLT would be more of a complement to instead of a replacement for, the current traditional market for the same assets, at least in the beginning stage of market development, for some processes or parts of the life cycle of the security. In this way, market participants would be allowed to validate the DLT's capability and enjoy some benefits.

In the end, the legacy and DLT-enabled systems could be eventually integrated as a hybrid version of the interface with traditional elements of infrastructure with the combination of DLT-based applications and automation in the fields of settlement, clearing, and others, in which the gains of efficiency are sufficiently high to prove (gradually) that the shift to a decentralized infrastructure is reasonable. Standardization of protocols and coordination among the participants in the market will also contribute to the faster adoption of DLT-based technologies and a wider and more quickly shift into such networks. Policy-makers can promote such coordination in the fields in which the significance and benefits brought by the application of DLTs to financial markets and their participants have been proven.

6.4.2 Legal Barrier for Introducing DLT to Securities Markets

Regulatory authorities have not approved the tokenized securities on the DLT network yet. The transaction between fiat currency and blockchain cryptocurrencies is not allowed in China, Egypt and India [167], and other countries. In addition, the regulators in other major countries still hold a vague view of blockchain tokens. Currently, many governmental permits and approvals are not permanent. Potential investors should then proceed, assuming that their tokens may not be publicly listed or sold without a definite period. Tokenization brings many important benefits, one of which is the cross-border transfer; while the harmonization of regulations at home and abroad will not support tokenization in the coming years. Fraud and hacking in the token market without regulation also impede public participation in tokenization, thus hampering the long-term development and maturity of the token market. However, in the case of too strict regulations, the benefits of tokenization will be diluted. Under strict centralized management, the key values of tokenization, including financial independence, decentralization, and democratization can be weakened.

- According to EMIR, all standardized OTC derivative contracts are required to be cleared through a central counterparty; therefore, a DLT system that does not have a CCP would not satisfy the requirements of EMIR.
- In legal terms, there might not exist a double-entry account in a DLT system, and that a DLT provider being able to obtain a CSD license in Europe is not likely to happen. Therefore, Member States can provide issuance services or comply with asset isolation requirements.
- Due to the difficulties with the concept of a transfer order in a DLT environment, determining the finality is difficult, which may fail to follow the settlement finality requirement of CSDR.

6.4.3 Legal Adaption to DLT-based Securities Markets

The LawTech Delivery Team's UK Jurisdiction Task Force (UKJT) brings the Law Commission for Wales and England, the Judiciary, and technical and legal professionals together and invites the Financial Conduct Authority to be the technical consultant. In May 2019, a public consultation was launched by the UKJT in order to identify and analyze the raised crucial legal questions to offer a reliable basis for the major adoption of crypto assets and smart legal contracts in Wales and England. The work group aims to release a legal statement related to the status of smart contracts and crypto assets under British private law [177]. The enhanced digitalization and standardization of legal documentation support the faster development of technical solutions, the achievement of interoperability between systems and services as well as the improvement of transparency and consistency between regulators and market players. However, if new technologies are developed and carried out in the market of derivatives, the potential uncertain areas of law and regulation emphasized in the whitepapers and legal guidelines of ISDA for smart derivatives contracts will be caused [128]. France is one of the first jurisdictions that take into account the application of a regulatory framework for DLT in the financial field. Regulators' actions around DLT are reflected in the 2016 Monetary and Financial Code of Law, the 2017 Blockchain Executive Order well as the innovative framework for token offerings established through the PACTE Action Plan for Business Growth and Transformation bill (published on May 24th, 2019) [89]. The purpose of this legislation is to provide blockchain stakeholders with an all-around legal structure instead of a sandbox so as to further provide answers to all the questions encountered by stakeholders in the new ecosystem, regardless of a regulatory, fiscal, or accounting nature. In October 2019, the Swiss National Bank (SNB) and the Bank for International Settlements (BIS) reached an agreement in terms of the operation on the BIS innovation center in Switzerland [212]. The Centre will take the initiative to study two projects. The first project is to explore the integration of the central bank's digital money into a distributed ledger technology infrastructure. This new form of digital money will aim to facilitate the settlement of tokenized assets between financial organizations. The project, as part of a collaboration between the SNB and the six-person team, will be carried out as concept proof. The second one is to solve the increasing requirements imposed on central banks to track and supervise rapidly developing electronic markets in an effective way. With the increasing automation and fragmentation of financial markets and more adoption of new technologies, these requirements have been raised. In 2018, the Swiss Capital Markets and Technology Association (CMTA) published a blueprint to tokenize the shares of Swiss enterprises with the adoption of DLTs [289], providing detailed guidance on the tokenization of Swiss enterprises' equity securities i.e., the incorporation into a digital token recorded on a blockchain. In September of 2020, the European Commission proposed a set of legislative proposals accompanying its novel digital finance strategy [340]. The overall goal is to foster digital innovation and competition in the financial domain while mitigating the relevant risks. The goal for the framework will take effect by 2024. One proposal in the package is to regulate the market for crypto assets, i.e., a Regulation on Markets in Crypto-Assets: 'MICA' proposal; with the DLT-based Market Infrastructure proposal [340] [5]. It is the first practical action, aiming to properly protect consumers and investors, offers legal certainty for crypto-assets, enables creative enterprises to utilize blockchain, DLT, and crypto-assets, and guarantees financial stability in this field [6].

[5]Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on markets in crypto-assets, and amending Directive (EU) 2019/1937. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020PC0593

[6]Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a pilot regime for market infrastructures based on distributed ledger technology. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020PC0594

Chapter 7

Related work

In each preceding chapter, we have contextualized our methodologies by discussing relevant literature and drawing comparisons to our approaches. In this section, we give a big picture of formal argumentation. First is formal argumentation vs. computational argumentation, then the branches of argumentation as inference and argumentation as dialogue, and we also provide a comprehensive overview concerning the three aspects we explored in this thesis: argument strength, multi-agent argumentation, and support in argumentation. Lastly, we briefly discuss argumentation used in legal AI. This allows us to position our contributions within the broader intellectual landscape of this field, thus underscoring the significance of our work.

Argumentation has recently received significant interest within AI, it is a part of the symbolic approach to AI in the field of knowledge representation and reasoning. As can be observed in all walks of life, arguing is so natural to deal with differences of opinions, whether it is in the form of a monologue, in one's mind, by evaluating pro and con arguments, or in the form of dialogue where arguments are exchanged among multiple agents. It thus mimics human reasoning and decision-making in rich forms, even if with uncertain and possibly contradictory information, individually or collectively. There already have been a number of applications of argumentation, including law [265], medicine [164, 131], health promotion[154], debate [298]and dispute mediation [172] (for a survey, see [24]), have recently made use of formal and computational argumentation methods. The developments of techniques of AI in argumentation theory have led to the design of machines in real-world situations. For instance, recently the autonomous debating system Project Debater has been developed that can perform a debate with a human expert debater [298].

Argumentation is a colorful landscape painting, which is by no means dedicated to people from a specific field but developed by absorbing the essence of different fields. As identified by Gabbay et al. in the Dagstuhl Perspectives Workshop [135], the most important open problems in argumentation are "how to formally represent various kinds of arguments and how to identify sets of postulates on the reasoning activity over arguments in specific contexts", and "the relationship between argumentation and other research fields (e.g. natural language processing, machine learning, human-computer interaction, social choice) was seen to be of major importance, especially to develop more mature applications."

The current state of the art of formal argumentation is nourished by the combined efforts of people from a variety of fields, e.g., philosophers, logicians, lawyers, et al. People from different disciplines bring up different theoretical perspectives and approaches. As witnessed by the handbooks of argumentation and by the plenty number of papers, the big picture of formal argumentation consists of three interrelated components. Argumentation is deeply rooted in philosophy, i.e., defeasible reasoning, and mathematics, i.e., non-monotonic logics [274, 138, 139]. As distinguished by Eemeren, Garssen, et al. [317] in the Handbook of Argumentation Theory: on the one hand, the program has a philosophical component, in which a philosophy of reasonableness is developed, and a theoretical component, in which, starting from this philosophy, a model for argumentative discourse is designed. On the other hand, the program has an empirical component, in which argumentative reality as it manifests itself in communicative and interactional exchanges is investigated. Finally, in the practical component, the problems that occur in the various kinds of argumentative practices are identified, and methods are developed to tackle these problems. Different concepts lead to different theoretical models, while the concepts are dependent on the content and argumentative situations where the argumentation is happening[316]. The various components of the complex research program of argumentation theory and their relationships are depicted in Figure 7.1 adapted from [316].

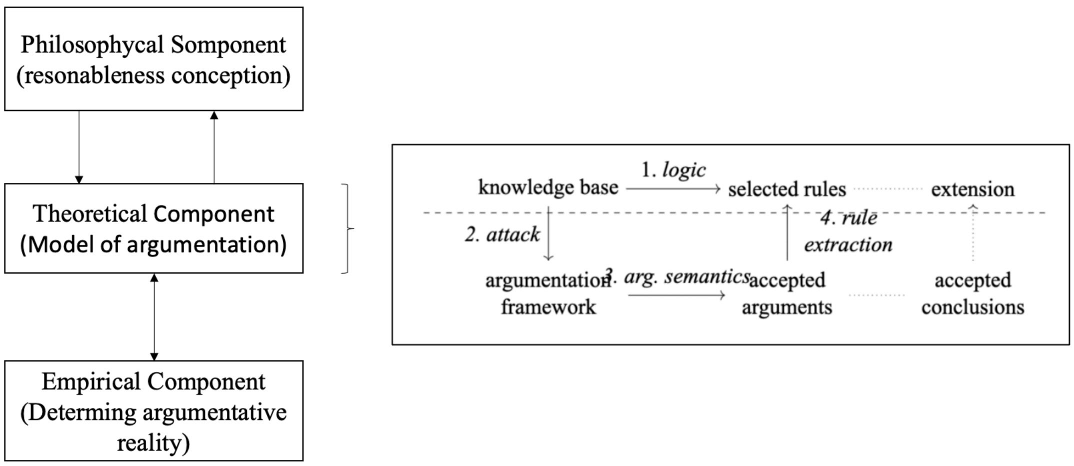


Figure 7.1: Components of argumentation theory

7.1 Formal Argumentation vs. Computational Argumentation

Formal argumentation is a subfield of AI and Computer Science that evolved from Knowledge Representation and Reasoning, and as such, formal argumentation is closely related to computational argumentation but not identical to it. In fact, to distinguish itself from traditional fields of argumentation, the journal of the community is called "Argumentation and Computation", it has provided a dedicated venue for papers in the field of computational argumentation. The conference called "Computational Models of Argumentation" (COMMA) has been a regular forum for the exchange of the results of computational argumentation, since 2006. The two volumes of Handbook of Formal Argumentation, though, are called Handbook of Formal Argumentation. Computational argumentation incorporates argument mining, a subfield of natural language processing (NLP), making it more inclusive than formal argumentation. With the advent of text reasoning based on foundational models, also known as large language models (LLMs), the distinction between logic and language is becoming less distinct.

7.2 Argumentation as Inference and Argumentation as Dialogue

Argumentation as inference One of the strengths of the abstract argumentation framework is its powerful generality, its process that transforms a knowledge base into an argumentation graph and obtains a set of acceptable conclusions for this knowledge base has been ”dubbed the argumentation pipeline” [159]. In more detail, the argumentation pipeline takes as input from a knowledge base formulated in a formal language, specifying how arguments are constructed relative to a premise set and a number of inference rules. Premises are formulas in a given formal language. They represent the evidence or information on the basis of which we build arguments. Rules are used for inferring new formulas from others. Arguments thus are considered the result of applying inference rules to the given premises and, possibly, of chaining such applications. As a second step, attack relations are established between the arguments, taking various considerations about the arguments into account (such as the syntactic form of the arguments, the strength of the arguments, and so on). Argumentation semantics are then used to obtain sets of acceptable arguments based on the argumentation graph constructed in the previous step. Finally, sets of acceptable conclusions are obtained on the basis of the sets of acceptable arguments. Such a knowledge base can be used to model, for example, default reasoning [335], logic programming with negation as failure[74], autoepistemic reasoning [58], causal inference [56]. The way defeasibility is represented comes in various flavors. For example, In ABA, all inference rules are strict but premises can be defeasible. This type of defeasible reasoning has been labeled plausible reasoning by Rescher [275].

There are three central approaches that correspond to this line of research: logic-based deductive methods [49, 16, 52], assumption-based argumentation systems [58, 310], and ASPIC systems [218]. A well-known and popular family of frameworks for structured argumentation is ASPIC (see [217, 218, 76]). In ASPIC+, both premises and inference rules may be defeasible. Arguments are constructed using an argumentation system. An argumentation system allows for a distinction between strict (i.e. deductive) and defeasible rules. Acceptable arguments are determined using the semantics of abstract argumentation. To be able to apply these semantics, an argumentation graph has to be constructed on the basis of the argumentation system. This is done by building arguments using the premises and rules of the argumentation system and specifying attacks between these arguments on the basis of the contrariness function and the preference order of the argumentation system. Indeed, in many situations, it might be opportune to take into account other criteria, i.e., argument strength. In most ASPIC alike formalisms, argument strength is taken into account by distinguishing between attack and defeat. In those formalisms, a defeat is an attack that is successful in view of considerations related to argument strength. In more detail, when two arguments conflict, one of the arguments may defeat the other due to its higher priority. In order to compare arguments with respect to their strength, we first need to specify how to obtain a (relative) measure of argument strength based on the strength of the defeasible rules used in an argument. In structured argumentation, this is done by defining so-called lifting principles. One of the most-studied and best-behaved lifting principles is based on the intuition that an argument is as strong as its weakest link [218], which can be dated back to Pollock [256].

An important development is the study of rationality postulates introduced by Caminada and Amgoud [71, 72], and later extended by Caminada et al. [75] and Wu and Podlaszewski [333]. They proposed several properties that any argumentation system should fullfil. These properties are meant to ensure that argumentation-based inferences make sense from a logical point of view, i.e. that the graph based selection is sensible from the perspective of the logical language that was used to construct the argument graph. The choice of attack relation (e.g. unrestricted versus restricted rebut) can have a major impact on the satisfaction of the rationality postulate.

Argumentation as dialogue Argumentation dialogues have been significantly applied in the fields of AI and law and multi-agent systems since the 90s, see the first chapter of the first volume Handbook of Formal Argumentation. where the role of agents is on the central stage. In the early days, Lorenzen and Lorenz developed formal dialogue systems for argumentation by using a game formulation of disputes among agents in argumentation [199]. The acceptance of an argument given by an agent is depending on many aspects. For instance, trust [273, 166], voting in social choice [96, 190, 26, 73]. In 2011, Rienstra et al., proposed multi-sorted argumentation, where each agent owns a part of the framework and may locally adopt a different semantics [276]. Multi-agent systems can be roughly grouped into two categories: cooperative and non-cooperative [120]. In cooperative systems, agents share a common goal and fully cooperate in order to achieve it. Agents can form coalitions in order to improve their performance, i.e., pooling their efforts and resources so as to achieve particular tasks at hand in a more efficient way [119]. In a non-cooperative system, each agent has its own desires and preferences, which may conflict with those of other agents. Multi-agent argumentation takes inspiration from several disciplines, such as game theory, and it can be further developed towards coalitional game theory by introducing the notion of the coalition and associate arguments of (sets of) agents. An alternative approach to multi-agent argumentation takes its inspiration from voting theory, and more generally from social choice.

Dung [110] shows how his abstract theory can also be applied to reasoning in game theory. Game theory is a branch of mathematical economics that models and analyzed the behavior of entities that have preferences over possible outcomes, and have to choose actions in order to implement these outcomes, thus it is suitable to provide a theoretical foundation for the analysis of multi-agent systems that are composed of agents with self-interests. Dialogue-based semantics has been developed, and a comprehensive overview of the work carried out in this direction is provided by Caminada [70]. In these dialogues, two players exchange arguments from a given argumentation framework in order to prove or disprove the acceptability of a particular argument.

Argumentation as inference and argumentation as dialogue approaches are distinct but not incompatible. Some definitions amalgamate these two perspectives. For example, when there is only one agent, that agent engages in reasoning with a single inconsistent knowledge base. Conversely, in scenarios with multiple agents, each one may have its own knowledge base, whether consistent or not, necessitating the study of how to construct a reasoning process among these agents, for example, see [195]. They are both approaches to managing some form of conflict. Argumentation as dialogue approach encompasses the presence of different actors and emphasizes the dynamic and aggregation aspects of the process, while the argumentation-as-inference approach is more focused on arguments and relation construction and the evaluation of argument acceptability. From the view of combining these two approaches, agents can use argumentation as a process based on the exchange and valuation of arguments for and against opinions, proposals, claims, and decisions. In a sense, monological argumentation is a static form of argumentation. It captures the net result of collating and analyzing some conflicting information. In contrast, dialogical argumentation is a dynamic form of argumentation that captures the intermediate stages of exchanges in the dialogue(s) between the agents and/or entities involved [51]. These dialogues serve as a proof theory for the acceptability status (skeptical, credulous) of arguments with respect to various semantics, allowing one to prove the acceptability of an argument without requiring the computation of all the extensions. In its essence, argumentation can be seen as a particularly useful and intuitive paradigm for doing nonmonotonic reasoning [72].

7.3 Three aspects of formal argumentation

7.3.1 Argument Strength

The concept of preference plays a prominent role in common sense reasoning, with a decidedly non-monotonic flavor. In epistemic reasoning, argument preferences are often based on probabilistic considerations, degrees of belief, or credibility estimates of information sources [198]. In argumentation as decision-making, they have been based on preferences for decision outcomes. In normative reasoning, for example, it is traditional in legal reasoning to order laws hierarchically, using criteria such as time, source, chronology, and speciality. There are close parallels between changes in norms and changes in beliefs. In order to apply a norm system with conflicting norms to a particular situation, some of the norms may have to be ignored. The problem of how to prioritize among conflicting norms is similar to the selection of sentences for removal in belief contraction [156].

There is a large variety in how to deal with preference among logic-based AI, e.g., some approaches take a preference ordering as expressing a “desirability” that property be adopted while in others the ordering expresses the order in which properties (or whatever) are to be considered, while some approaches conflate the notion of inheritance of properties with the general notion of preference. Preferences are added in such a manner to default logic [63, 103], autoepistemic logic [180, 279], circumscription [211, 196], and logic programming [343, 62]. For a systematic survey and classification of preference handling mechanisms in non-monotonic the interested reader is referred to Delgrande et al. [103] and to Beirlaen et al. [41]. In argumentation modeling, preference is embedded argumentation logic, providing argumentation-based characterisations of non-monotonic formalisms augmented with preference, then it needs to provide an account of how these priority orderings can be ‘lifted’ to preferences over arguments. Pollock’s system formulated strength in terms of numerical degrees of belief, as his style, were meant for epistemic reasoning. He used the notion of weakest link in 1995 [256], as a way to compare the strength of arguments. It has been used as a general framework for instantiating (prioritised) default logic, etc. These logics can be formalised in structured argumentation (e.g. ASPIC+) to generate abstract argumentation frameworks. In ASPIC+, the attack relation is defined by argument strength based on weakest link or last link. Pollock's work and the distinction between weakest and last link in particular played a central role in formal models of structured argumentation. In fact, it is the need to represent the distinction between weakest and last link that necessitates the possibility to represent default rules in structured argumentation. However, given the long history of the discussion of these two principles, it may come as a surprise that the philosophical notions of them and when to use them are not clear. There have been few developments characterising how to use them to instantiate abstract argumentation frameworks. Modgil and Prakken suggest that the right way to use preferences may depend on the kind of content of arguments, for example, on whether the reasoning is epistemic, normative or about decision making [218]. Dung compares them with a set of rationality postulates. For example, the postulate called attack monotonicity informally says that strengthening an argument cannot eliminate an attack of that argument on another, and credulous cumulativity informally means that changing a conclusion of an argument in some extension to a necessary fact cannot eliminate that extension. Dung then identifies several sets of conditions under which one or both of these postulates and/or the original postulates of [72] are satisfied. Dung then continues by investigating several definitions of defeat in terms of the preference relation on defeasible rules on whether they satisfy these various postulates. His results actually characterize last link, while weakest link does not satisfy, for example, credulous cumulativity and context independent. However, there is even a debate over whether these properties are desired or not. According to Prakken and Vreeswijk, it is positive for attack monotonicity but, but negative for credulous cumulativity. They point out that strengthening a defeasible conclusion to an indisputable fact may make arguments stronger than before, which can give them the power to defeat other arguments that they did not defeat before. This may in turn result in the loss of the extension from which the conclusion was promoted to an indisputable fact.

Orders Since arguments are often based on multiple (defeasible) premisses (resp. applications of defeasible rules), in order to compare the strength of arguments, one needs to lift the quantitative or qualitative comparisons from premisses (resp.rules) to sets of premisses (resp.rules). According to [41], such strength could stem from probability (e.g., the credence an agent has in the argument being acceptable), the quantity of the available evidence that supports it, and the degree of specificity. There are several lifting principles, lexicographic liftings [64, 45], the weakest link and the last link [256, 218], etc. These many choices may give rise to quite different outcomes. Most authors agree that the appropriateness of a specific lifting principle mainly depends on the context of the application. For instance, Pollock (as well as stated by Prakken) argued that weakest link is appropriate in an epistemic setting [256, 218], while perhaps the last link is more appropriate in a legal setting [218]. Another question that needs to be settled regarding preference is the so-called representational choice: how to order defeasible information. There are choices like linear orders, preorders, total orders, and modular orders.

Translation between frameworks

There is a lot of work in the nonmonotonic logic and logic programming literature on prioritised rules, see e.g. Delgrande et al. [104] for an overview. Pardo and Straßer give an overview of argumentative representations of prioritized default logic, concerning weakest link, they mainly consider dwl [242]. Various authors have discussed the dilemma between weakest link and last link [69, 194, 216, 219]. The analysis of weakest link related to swl indicates that it is more complicated and ambiguous than it seems at first sight. With partial orders, ASPIC+ tries to accommodate both in combination with democratic and elitist orders [216, 219], but neither of them is clearly better than the other. Young et al. [334, 336] show that even for total and modular orders, swl cannot always give intuitive conclusions. They also show the correspondence between the inferences made in prioritised default logic (PDL) and dwl with strict total orders. Then they raise the question of the similarity between weakest link and PDL for modular and partial orders. Moreover, Liao et al. [194] give similar results but use other examples to demonstrate that the approach of Young et al. [334, 336] cannot be extended to preorders [194]. Liao et al. [194] use an order puzzle in the form of Example 2.3 to show that even with modular orders, selecting the correct reasoning procedure is challenging. This leads them to introduce auxiliary arguments and defeats on weakest arguments. Beirlaen et al. [41] point out that weakest link is defined purely in terms of the strength of the defeasible rules used in argument construction. More recently, Lehtonen et al. present novel complexity results for ASPIC+ with preferences that are based on weakest link (swl in this paper) [189], they rephrase stable semantics in terms of subsets of defeasible elements.

7.3.2 Multi-Agent Argumentation

In many computer science fields, such as database integration, and multi-criteria decision-making, there is a need to synthesize information from multiple sources. In the database field, a key issue is to be able to integrate multiple databases into a single database [294]. While the issue is that information from different resources often comes up with conflicts. Systems organized around multiple reasoning agents face the similar problem of resolving conflicts among contradictory knowledge or beliefs held by different agents. For example, inconsistency problems occur when one wants to combine several expert systems. Consider a set of belief bases coding the belief of several human experts. In order to build an expert system it is reasonable to try to combine all these belief bases in a single belief base that expresses the belief of the experts' group. As classified by Amgoud et al. [9], to solve such a problem, there are mainly two different approaches have been proposed:

- The first category of approaches merges the different bases into a unique consistent base, e.g., to take the disjunction of the maximal consistent subsets of the union of the knowledge bases.
- and the second category, by adopting argumentation, accepts inconsistency and copes with it, additionally considering the priorities to solve conflicts.

Very recently, Beishui et al. proposed a multiple-agents ethical advisory component, called Jiminy, based on a theory of normative systems and formal argumentation [195]. Jiminy is a comprehensive framework that somehow incorporates the above two approaches together. In their setting, each agent is with a normative knowledge base. They design a step-wise process to deal with moral dilemmas emanating from multiple normative systems. All of the knowledge bases are treated independently, i.e., without interaction between agents, and the conclusions derived from all the knowledge bases are compared. If there is a dilemma, first, they combine all the arguments of the agents, i.e., combining all the argumentation frameworks into a unique framework consisting of all of the arguments. If the dilemma is not resolved, as a second step, they combine all the knowledge bases into a single normative system. If all the approaches fail, they finally rely on the selection by the agent who has the most expertise that is context-dependent.

The idea of combining all the argumentation frameworks of multiple agents in Jiminy has essentially developed in abstract agent argumentation. Abstract agent argumentation extends Dung's abstract argumentation framework to a set of agents to formalize multi-agent argumentation. In [17], the framework is called-triple A. It uses the theory of input/output argumentation described by Baroni et al. [32], also known as multi-sorted argumentation [276], where the sorts have been interpreted as different agents. This theory allows arguments to be assigned to agents, and with the role of agents, various semantics can be defined. They further apply their framework to legal reasoning, using an example from a court case where two scenarios have been distinguished. First is the dialogue among different agents, e.g., the accused, the lawyers, the witnesses, the prosecutors, and so on. The arguments an agent reveals may depend on the arguments revealed by other agents, at each step an agent can commit to accept or reject some arguments or commit to hide or reveal one of his/her rejected arguments. Additionally, there is a so-called external observer, e.g., the jury or the judge, who has to take into account all the arguments put forward during the deliberation and decide which arguments to accept. Thus, they distinguished the collective argumentation of judges from the individual argumentation of the accused, prosecutors, witnesses, lawyers, and experts.

TripleA demonstrates two of the most important aspects of multi-agent argumentation, namely argumentation as dialogue and justification aggregation. Argumentation dialogues have been significantly applied in the fields of AI and law and multi-agent systems since the 90s, see the first chapter of the first volume Handbook of Formal Argumentation. where the role of agents is on the central stage. In the early days, Lorenzen and Lorenz developed formal dialogue systems for argumentation by using a game formulation of disputes among agents in argumentation [199]. The acceptance of an argument given by an agent is depending on many aspects. For instance, trust [273, 166], voting in social choice [96, 190, 26, 73]. Multi-agent systems can be roughly grouped into two categories: cooperative and non-cooperative [120]. In cooperative systems, agents share a common goal and fully cooperate in order to achieve it. Agents can form coalitions in order to improve their performance, i.e., pooling their efforts and resources so as to achieve particular tasks at hand in a more efficient way [119]. In a non-cooperative system, each agent has its own desires and preferences, which may conflict with those of other agents. Multi-agent argumentation takes inspiration from several disciplines, such as game theory, and it can be further developed towards coalitional game theory by introducing the notion of the coalition and associated arguments of (sets of) agents. An alternative approach to multi-agent argumentation takes its inspiration from voting theory, and more generally from social choice.

Dung [110] shows how his abstract theory can also be applied to reasoning in game theory. Game theory is a branch of mathematical economics that models and analyzed the behavior of entities that have preferences over possible outcomes, and have to choose actions in order to implement these outcomes, thus it is suitable to provide a theoretical foundation for the analysis of multi-agent systems that are composed of agents with self-interests. Dialogue-based semantics has been developed, and a comprehensive overview of the work carried out in this direction is provided by Caminada [70]. In these dialogues, two players exchange arguments from a given argumentation framework in order to prove or disprove the acceptability of a particular argument.

The aggregation of opinions is one of the most general and inspiring motivations for the study of argument aggregation. This motivation is shared in part by social choice theory [22]. Social choice theory concerns the design and formal analysis of methods for aggregating the preferences of multiple agents, such methods include voting procedures, which are used to aggregate the preferences of voters over a set of candidates. Social choice theory is one of the fundamental tools for multiagent systems, witnessed by its origination in economics and political science. As stated in [61], " if we view a multiagent system as a "society" of autonomous software agents, each of which has different objectives, is endowed with different capabilities and possesses different information, then we require clearly defined and well-understood mechanisms for aggregating their views so as to be able to make collective decisions in such a multiagent system". This has been devoted to the study of preference aggregation in collective argumentation, which is concerned with the explicit or implicit aggregation of individual preferences among arguments in order to find collective opinions based on collectively supported reasons. Deliberative democracy is clearly reflected in this characterization [246], but models of argument aggregation can potentially be used for a wider range of applications covering, for instance, collective intelligence [26] and prediction markets [232].

7.3.3 Support in Argumentation

An argument can attack another argument, but it can also support another one. This suggests a notion of bipolarity, i.e. the existence of two independent kinds of information which have a diametrically opposed nature and which represent repellent forces. When an argument is aimed at establishing the truth, empirical evidence can be used to support alleged facts. For instance, a witness's testimony can provide evidence for the claim that the suspect was at the scene of a crime, a clinical test can provide evidence against a medical diagnosis, and the outcome of a laboratory experiment can be evidence confirming (or falsifying) a psychological phenomenon [186].

Toulmin proposed a model for the structure of arguments that distinguishes between data, claim, warrant, backing, rebuttal and qualifier, where he identifies support as a relation between data and claims. In decision-making, support is also largely involved. In [7, 108], it has been argued that when making decisions, one generally takes into account some information in favor of the decisions and other pieces of information against those decisions. Similarly, concerning knowledge and preference representation [60, 303, 187, 44]. In [60], two kinds of preferences are distinguished: the positive preferences representing what the agent really wants, and the negative ones referring to what the agent rejects. This distinction has been supported by studies in cognitive psychology which have shown that the two kinds of preferences are completely independent and are processed separately in the mind. The beliefs in Pollock's OSCAR are only justified if they are supported by an argument that cannot be refuted. However, in Dung's theory of abstract argumentation [110], support is implicit, and only the attack relation between arguments is taken into account. In his requirements analysis for formal argumentation, Gordon proposes the following definition covering more clearly argumentation in deliberation as well as persuasion dialogues [147]: "Argumentation is a rational process, typically in dialogues, for making and justifying decisions of various kinds of issues, in which arguments pro and con alternative resolutions of the issues (options or positions) are put forward, evaluated, resolved and balanced." Argumentative support is established by constructing arguments for conclusions from a given set of possible reasons or rules (of inference). Arguments can be constructed step-by-step from knowledge bases by chaining inference rules into trees. Arguments thus contain subarguments, which are the structures that support intermediate conclusions (plus the argument itself and its premises as limiting cases). As such an account, the ASPIC+ framework [218] will be used, which allows modeling of inferential support relations with its notion of a subargument. At an abstract level, it seems that these pro and con arguments can be represented more easily in so-called bipolar argumentation frameworks [79, 82, 81, 83] containing besides attack also a support relation among arguments. A new formalization of coalitions in terms of conflict-free maximal support paths has been proposed [82].

As pointed out in the Handbook of Argumentation Theory [317], support and attack are both important elements in argumentation and can be either modeled as complementary notions, as in bipolar argumentation frameworks or can be handled separately, with support implicit in the construction of structured arguments and attack between arguments handled explicitly. For instance, in structured argumentation systems, the concept of sub-argument already provides a straightforward interpretation of inferential support. Prakken investigates the degree to which abstract argumentation with support may serve as an abstraction of ASPIC+ structured argumentation [263]. Instead of treating support as a connection between arguments and conclusions, in this thesis, we view support as a connection between arguments, we examine bipolarity at the level of argument interaction.

While there is general agreement in the formal argumentation literature on how to interpret attack, even when different kinds of semantics have been defined, there is much less consensus on how to interpret support [92]. There exist very few results and studies about the role of support in abstract argumentation. Moreover, it seems that each variant of support can be used for different applications. There exist different approaches to extending Dung's abstract theory by taking into consideration the support relation. The relation between support and attack has been studied extensively in reduction-based approaches, in the sense that deductive and necessary interpretations of support give rise to various notions of indirect attack [84], thus, they typically give opposite results. Deductive support [57] captures the intuition that if a supports b, then the acceptance of a implies the acceptance of b. This intuition is characterized by the so-called closure principle [84]. Necessary support [228] captures the intuition that if a supports b, then the acceptance of a is necessary to obtain the acceptance of b, or equivalently, the acceptance of b implies the acceptance of a. It has been characterized by the inverse closure principle [250]. Another approach to handling support is the evidence-based approach [235] where the notion of evidential support is introduced. An argument cannot stand unless it is supported by evidential support. Support can also be seen as an inference relation between the premises and the conclusion of the argument itself [261]. Moreover, in selection-based approaches [140], support is used only to select some of the extensions provided in Dung's semantics, and thus does not change the definition of attack, or defense.

Despite this diversity, the study of support in abstract argumentation seems to agree on the following three points.

Relation support and attack The role of support among arguments has been often defined as subordinate to attack, in the sense that in deductive and necessary support, if there are no attacks then there is no effect of support. On the contrary, in the evidential approach, without support, there is no accepted argument even if there is no attack.

Diversity of support Different interpretations for the notion of support can be distinguished, such as deductive [57], necessary [229, 227] and evidential support [48, 234, 252].

Structuring support Whereas attack has been further structured into rebutting attack, undermining attack and undercutting attack, the different kinds of support have not led yet to a structured argumentation theory for bipolar argumentation frameworks.

Given the relevance and significance of all the mentioned approaches, we think that there is still the need to explore other approaches that have not been yet considered for bipolar argumentation frameworks. The aim of our research is not to replace other approaches but rather to point out to the existence of other interesting ones that can be applied depending on the chosen application. Note that our approach is novel in its methodology. On one hand, reduction-based approaches can be seen as a kind of pre-processing step for Dung's theory of abstract argumentation (i.e. adding the complex attacks and then applying Dung's semantics). On the other hand, selection-based approaches can be seen as a kind of post-processing step (i.e. applying Dung's semantics and then applying the approach to select some of the extensions). Differently from those two groups of approaches, our approach (i.e. the defense-based approach) does not affect the concept of attack and conflict-freeness but rather changes the definition of defense.

For the evaluation of legal argumentation and for the evaluation of judicial decisions, one significant role is counter-arguments which make part of the justification of the standpoint [249]. Summers and MacCormick [202], for instance, also emphasize that counter-argumentation may play an important role in the justification of a decision:

> "Recent studies in argumentation theory by, amongst others, Snoeck Herlkemans [158] indicate that the analysis and evaluation of argumentation should be studied in a dialogical context. This dialogical approach to argumentation not only provides an insight into how arguments and counter-arguments result in complex argumentation, it also provides clues for reconstructing the structure of the argumentation. This common, dialogical, ground may be a fruitful starting-point to gain more insight into the various types of complex argumentation that are distinguished in legal theory."

A dialectical approach to argumentation implies that the analysis and assessment of the argumentation are not just aimed at the arguments that are put forward in favor of a standpoint, but at its counterarguments as well. Furthermore, Feteris specifies the judge's role in this context as follows [129]:

> "Since the parties do not themselves solve their dispute in consultation, but rather have the judge, as a neutral third party, decide on the eventual outcome, the parties should be granted insight into the grounds the judge has taken into

> consideration in reaching his verdict. In dialectic terms, stating the considerations underlying his decision amounts to stating the factors which were instrumental in his assessment of the acceptability of the propositional content as well as the justificatory potential of the argumentation of the party asking for a decision."

Carneades [146] proposed by Gordon models arguments as directed graphs consisting of argument nodes connected to statement nodes. The premises and conclusions of an argument graph are represented as statement nodes. A linked argument is one where two or more premises function together to support a conclusion. There can be two kinds of arguments in a graph, a pro (supporting argument) or a con (attacking) arguments. A supporting argument is represented by a plus sign in its argument node, whereas a con argument is represented by a minus sign in the nodes containing argumentation schemes such as modus ponens, argument from expert opinion, and so forth [1]. Conflicts between pro and con arguments can be resolved using proof standards such as the preponderance of the evidence [149].

A second approach, e.g., ASPIC+, does not separate support and attack when combining them. Arguments are constructed from reasons for and against conclusions, which in turn determine whether a conclusion follows or not. In this approach, for instance, conditional sentences are used to express which reasons support or attack which conclusions. An example is Nute's defeasible logic [231, 14], which uses conditional sentences for the representation of strict rules and defeasible rules, and for defeater rules, which can block an inference based on a defeasible rule. Algorithms for defeasible logic have been designed with good computational properties. Another example of the approach is Verheij's DefLog [320], in which a conditional for the representation of support is combined with a negation operator for the representation of attack. In all of these systems, arguments are modeled as graphs containing nodes representing propositions from the logical language and edges from nodes to nodes. In these systems, an argument can be supported or attacked by other arguments, which can themselves be supported or attacked by additional arguments. Another related approach is to extend Dung's abstract argumentation frameworks by expressing both support and attack, e.g., bipolar argumentation [79, 81], Evidential Argumentation Systems (EASs) [236]. Various semantics for such frameworks have been defined, claiming to capture different notions of support.

While there is general agreement in the formal argumentation literature on how to interpret attack, even when different kinds of semantics have been defined, there is much less consensus on how to interpret and handle support [92]. For instance, concerning the burden of proof, Gondon et al. categorized statements using three proof standards [148]:

- SE (Scintilla of Evidence). A statement meets this standard if and only if it is supported by at least one defensible pro argument.
- BA (Best Argument). A statement meets this standard if and only if it is supported by some defensible pro argument with priority over all defensible con arguments.
- DV (Dialectical Validity). A statement meets this standard if and only if it is supported by at least one defensible pro argument and none of its con arguments are defensible.

[1] http://carneades.github.com

These categories reflect different flavors when dealing with pros and cons in this case. For example, consider the following case of a neighbour's quarrel over a row of conifers [249].

> (...) Defendant argues that the conifers have been planted to reduce draught in his house, but this argument is absolutely unsound, since most of the window posts are closed and the window that does open, is located on a point higher than the tops of the conifers and has not been fitted with any antidraught facilities. (...) Whereas the defendant has no considerable interest in these conifers, removal is of significant concern to the claimant since they block his view and take away the light. (...)
> (2981. Country court Enschede 6 October 1988
>
> "The judge defends the standpoint that the claimant's interest in the removal of the conifers is greater than the defendant's interest in leaving them untouched. In support of this standpoint, he argues that the conifers block the view and take away the light. In his preceding remarks the judge mentions the defendant's argument in support of the opposite standpoint: he does have a considerable interest in the conifers since they reduce draught in his house. The judge refutes this argument referring to the fact that most of the window posts are closed and the opening window, which has no antidraught facilities whatsoever, is located higher that the tops of the conifers."
>
> **So the judge's argumentation consists of a pro-argument and the refutation of a counter-argument which, in conjunction, form sufficient support for his standpoint.**

This case calls for a new way to cope with support and attack. In Chapter 4, one of the new semantics (semantics 1) we proposed for bipolar argumentation is suitable for modeling this kind of judicial decision.

7.4 Argumentation Used in Legal AI

One significant work on Legal AI is the first volume of Handbook of Legal AI [319]. It presents a comprehensive overview of the state-of-the-art and trends in the research field of legal AI. The handbook provides a solid introduction to the essentials of the field for newcomers and a selection of advanced issues as a base for future research directions. As the law gets more complex, conflicting, and ever-changing, more advanced methods, most of them come from the Artificial Intelligence (AI) field, are required for analyzing, representing, and reasoning on legal knowledge. The discipline that tackles these challenges is now known as "Legal AI". Legal AI is experiencing, in particular, in the latest years growth in activity, also at the industrial level, touching a variety of issues which go from the analysis of the textual content of the law to reasoning about legal interpretation to ethical issues of AI applications in the legal domain (e.g., the artificial judge). This Handbook presents a collection of chapters that evolves around three main topics, namely norm mining (i.e., how to automatically identify, extract, classify, and interlink norms from text), reasoning about norms and regulations (i.e., how to derive new legal knowledge from the existing legal knowledge bases in such a way to address automatic legal decision making), and norm enforcement and compliance (i.e., how to check and ensure the compliance of the systems' requirements with the regulation).

Particularly, we concern the role of formal argumentation in legal reasoning, it has been in particular proposed to identify and explain inconsistencies in legal knowledge [43, 151]. Legal reasoning has many distinctive features, e.g., the facts of a case need to be identified and characterized; any proposed set of rules inevitably contain gaps and conflicts; many decisions express a preference for particular values and purposes; and all its conclusions are reasoned defeasibly [42]. All of these features mean that deductive reasoning is not suitable for legal reasoning instead argumentation as a way of non-monotonic reasoning must take center stage. For example, in the work of Dong et al. [107], formal argumentation is used to design defeasible deontic logics, based on two classical deontic logics, i.e., ASPIC+, the structured argumentation theory to define non-monotonic variants of well-understood monotonic modal logic. They illustrate the ASPIC+-based approach and the resulting defeasible deontic logic using argumentation about strong permission.

On the other hand, Arisaka et al. provide an overview of multi-agent abstract argumentation and dialogue, and its application to the formalization of legal reasoning [17]. The basis of multi-agent abstract argumentation is input/output argumentation, distinguishing between individual acceptance by agents and collective acceptance by the system. The former may also be seen as a kind of conditional reasoning, and the latter may be seen as the reasoning of an external observer. In this chapter, Arisaka et al. introduce dialogue semantics for abstract argumentation by refining agent communication into dialogue steps.

Chapter 8

Future work

This thesis proposes the novel concept of distributed argumentation technology, a computational approach incorporating argumentation reasoning mechanisms within multi-agent systems. In the ever-evolving landscape of computational argumentation, there emerges dialogue technology, for which we also give a preliminary definition. Dialogue technology is formal dialogue theory combined with computational platforms like chatbots and AI tools, it mirrors formal dialogues, for instance, persuasion, negotiations, etc. A prime example of dialogue technology is ChatGPT. In this section, we explore the intersections and divergences between these two emerging technologies. We consider argumentation as dialogue as a strategy to bridge distributed argumentation technology and dialogue technology. We first discuss how to bridge argumentation as dialogue and argumentation as inference via dialogue games. Then we show five examples in the literature of formal argumentation. These examples can be seen as future use case studies that bridge distributed argumentation technology and dialogue technology.

8.1 Bridging Argumentation as Inference and Argumentation as Dialogue

In the realm of formal argumentation, argumentation as inference and argumentation as dialogue are non-incompatible views. In the dialogue-oriented branch, arguments are moves in a game concerning the acceptance of some claim, and in the reasoning-oriented one, arguments are the results of some inference process, conflicts among them may arise due to the limits of the information and/or knowledge used in the process, and the reasoner needs to take a stance on the conflicts themselves. We can take the example of persuasion. At its core, argumentation as inference pertains to the process of drawing conclusions from a set of premises using logical reasoning. This form of argumentation is deeply rooted in the structure, validity, and soundness of arguments. The objective is to construct a persuasive line of reasoning that leads inescapably to a particular conclusion. Persuasion is achieved when the audience recognizes the validity of the inferential chain and is thus convinced of the conclusion's truth or acceptability. However, Persuasion dialogue is not just about presenting a sound argument but also about understanding the other participant's perspective, countering objections, asking questions, and adapting one's line of reasoning in response to the ongoing dialogue. The persuasive force emerges from the interplay of arguments and counterarguments, the mutual exchange of reasons, and the ability to adapt and respond to the evolving discourse.

Works that have the potential to combine two views can also be found in the literature sampling [17, 195]. If there is only one agent, then this agent reasons with one inconsistent information set; if there are several agents then each agent may have one information set, consistent or not, and it will be necessary to study the way to build a reasoning between all these agents. One important question is how to move between individual reasoning (argumentation as inference) and the direction of collective reasoning (argumentation as dialogue). One example is the judge, who takes argumentation as dialogue, the dialogue happening between various agents, e.g., prosecutor, lawyers, and plaintiff, and turns it into argumentation-as-inference, which is a god-like view, stepping outside the interaction, by balancing pros and cons arguments and make a decision. Bridging this gap holds immense potential for advancing the field. This future work aims to explore the integration of these two perspectives by incorporating dialogue games as a decision procedure within the existing framework.

Dialogue games provide a structured approach to modeling interactions between agents, allowing for a more dynamic and interactive form of argumentation. By introducing the role of agents, we can extend the traditional argumentation graph to accommodate multiple labels or colors for arguments, reflecting the diverse viewpoints and perspectives of the participating agents. An exemplary contribution in this direction is the work on Fatio by Mcburner and Parsons [208], where the authors propose a logic-based formalism for modeling dialogues between intelligent and autonomous software agents. Their approach builds upon the theory of abstract dialogue games, which provides a theoretical foundation for analyzing and structuring complex dialogues. The proposed formalism enables the representation of intricate dialogues as sequences of moves within a combination of dialogue games. This allows for nested dialogues, where dialogues can be embedded within one another, enabling the exploration of complex argumentative scenarios involving multiple layers of interactions and negotiations.

By incorporating dialogue games into the framework of formal argumentation, we can enhance our understanding of the dynamics of argumentation as well as develop more effective strategies for reasoning and decision-making in multi-agent systems. The integration of dialogue games offers opportunities for exploring novel approaches to conflict resolution, consensus building, and negotiation within the context of argumentation. In the following subsection, we identify some challenges for future work and give preliminary ideas on how to define dialogue games based on different frameworks.

8.1.1 Dialogue Games

Dialogue games are decision procedures (coloring graphs) with the role of agents. In this section, we discuss the possible future work on dialogue games.

Dialogue games for structured argumentation One aspect to explore could be how to model the dialogue based on structured argumentation, what constitutes a valid reason? When the opponent questions a claim, they could do so in two different ways: by stating an undercutter, which attacks the relevance or sufficiency of the reasons for a claim, or a rebutter, which directly contradicts the claim. How can these be better incorporated and differentiated in the dialogue game?

Growing knowledge bases. As the dialogue proceeds, the knowledge base, e.g., claims, arguments, and counter-arguments (and their associated reasons and evidence) can change and expand, as can the relationships between them. Future work could develop methods for representing and managing dynamic knowledge bases, where new information can be added, and existing information can be updated or retracted during the dialogue. On the other hand, each participant in the dialogue could have their own individual knowledge base, which represents their beliefs and information. However, there might also be a shared knowledge base, representing information that all participants agree on. How can these individual and shared knowledge bases interact and influence each other during the dialogue? In addition, if the knowledge base can change during the dialogue, this adds a new dimension to the participants' strategies. They not only need to argue effectively based on the current knowledge base but also anticipate how the knowledge base might change and plan accordingly.

Multi-agent dialogue. One possible avenue for future research could be extending dialogue games between proponent and opponent to multi-agent systems. Multi-agent dialogues could entail more complex strategies and provide richer insights into collective decision-making processes.

Defining a dialogue framework to capture dynamic argument acceptance. The acceptance of an argument can be distinguished between weakly accepted and strongly accepted. An argument is weakly accepted iff there is a strategy for an agent to accept it; an argument is strongly accepted iff for all strategies of the agents the argument will be accepted.

Add reasons or explanations. Each argument in the dialogue could be indexed and associated with a set of labels. Use these labels as explanations for the arguments. Each label serves to justify or provide insight into the argument, clarifying its context, its basis, and how it should be evaluated. We can incorporate this labeling system into a reason-based dialogue semantics framework. This would allow the dialogue to be structured around the reasons behind the arguments, with the labels serving to highlight and clarify these reasons.

Recover the extensions of argumentation from dialogue semantics.Recovering the extensions of an argumentation framework from reason-based dialogue semantics is a challenging task, as it requires mapping the dynamic process of dialogue, with its reasons and counterarguments, onto the static structure of argumentation extensions.

8.2 Distributed Argumentation Technology and Dialogue Technology

8.2.1 Relation and Strategy

We first repeat the definitions of distributed argumentation technology and dialogue technology we give in this thesis.

Distributed argumentation technology is a computational approach incorporating argumentation reasoning mechanisms within multi-agent systems.

Dialogue technology is formal dialogue theory combined with computational platforms like chatbots and AI tools, it mirrors formal dialogues, for instance, persuasion, negotiations, etc.

The definitions of distributed argumentation technology and dialogue technology remain in development, which allows for varied interpretations. Upon closer examination, ”distributed argumentation” and ”dialogue” exhibit similarities, if not nearly identical. This is primarily because dialogues intrinsically happen within multi-agent systems, which are by nature, distributed. Consequently, one could argue that dialogue is inherently a component of distributed argumentation.

The distinctions between distributed argumentation technology and dialogue technology are primarily on their scope and the technology adopted in their evolution. Distributed argumentation technology, as explored in this thesis, goes beyond the confines of argumentation as dialogue, encompassing distributed argumentation as inference. It is about ensuring that the logical structures and justifications of arguments are preserved, validated, or communicated in distributed systems. However, dialogue technology is more about real-time interactive communication that employs these arguments. It incorporates foundational reasoning structures into interactive platforms, such that they can converse and reason in real time. They use very large amounts of data and sophisticated algorithms to generate human-like text, making the interaction feel natural while still following the formal structures of argumentation. The foundation models, like ChatGPT, play an important role as the medium through which these dialogues are facilitated.

It’s worth noting that these two concepts can be interpreted differently based on one’s perspective, a topic open for further exploration. It should also be emphasized that our future primary focus remains on argumentation theory, which encompasses both logical and dialogical reasoning. While we recognize the value of specific technologies like distributed ledger or foundational models, our main interest lies in their potential to enhance core theoretical development.

8.2.2 Five Examples

This section presents five examples in the literature of computational argumentation. They show different aspects of dialogues and they can be seen as future case studies driving the development of dialogue technologies. For brevity, we will focus on aspects specifically related to dialogue. For a more comprehensive understanding of these examples and underlying theories, readers are referred to the original papers from which they are sourced.

Example 8.1 (Jiminy Moral Advisor [195])**.** *Jiminy uses techniques from normative systems and formal argumentation to resolve moral dilemmas among multiple stakeholders. Each stakeholder owns their own normative system, which is in the sense of being distributed. Jiminy combines norms into arguments, identifies their conflicts as moral dilemmas, and evaluates the arguments to resolve each dilemma (whenever possible). In particular cases, Jiminy decides which of the stakeholders take preference over the others.*

The current Jiminy mainly adopts only argumentation as inference. One characteristic that should be noted is that Jiminy concerns ethics and morality. It also concerns the preference orders over the agents, which touches upon the meta-level reasoning. This involves a higher-level reflection on the dialogue process, i.e. participants are not just evaluating the content of the arguments but also the agents presenting those arguments. One can ask

How can dialogue technologies incorporate meta-level reasoning to better navigate moral dilemmas?

In the authors' perspective, one future direction could be how to generalize and transform the current Jiminy moral advisor into an interactive Dialogue Jiminy by replacing argumentation as inference with argumentation as dialogue. Regarding this aspect, one solution is to create a communication language and a protocol for persuasion dialogues on moral dilemmas [260]. Another is to study strategic aspects of these dialogues for the participants [306]. As envisioned Pardo et al. [243], one can expect to advance the theory of dialogue technology by improving the state of the art in text mining and explainability in AI (XAI) for norms and decisions through a combination of symbolic AI (Dialogue Jiminy) and sub-symbolic or data-driven methods (LLMs). Regarding this aspect, one can create a natural language interface between the Dialogue Jiminy and the stakeholders, and use machine learning to construct two language modules. The first is to use NLP to transform the stakeholders' informal norms into the avatars' formal rules; the second is to use natural language generation to synthesize formal dialogues into explanations (in plain language) of why a particular decision was passed as a moral recommendation to the agent.

Example 8.2 (Twelve angry men [67])**.** *The second example is from the NoDE benchmark, among which there are the annotated datasets extracted from of "Twelve Angry Men" play. Twelve Angry Men presents a scenario where a jury engages in intense dialogue to decide the fate of a teenager charged with murder. Amidst conflicting viewpoints and personal biases, the jurors grapple with their responsibility to deliver a just verdict based on the evidence presented. The play is divided into three acts: the end of each act corresponds to a fixed point in time (i.e. the halfway votes of the jury, before the official one).*

The authors translate natural language dialogues into abstract bipolar argumentation graphs. They first extract the dialogues from the script of the play and automatically generate arguments with unique labels and subsequently pair arguments. The relation between these pairs is then categorized as attacks, support, or null, leading to the construction of bipolar argumentation graphs, see Figure 4.3. The annotation of natural language dialogue to abstract argumentation provides a structured and nuanced representation of conversations. The process of translating natural language dialogues into abstract argumentation graphs presents a promising avenue for future research. As dialogues become increasingly complex in various domains, there's a growing need for tools that can automatically and visually represent and simplify these interactions. Incorporating such annotated data into the training of foundational models like ChatGPT can enhance their understanding of the structure of dialogues.

Example 8.3 (Child's custody)**.** *The third example is adapted from the divorce example used in Chapter 4. The following dialogue could happen among a couple arguing the child's best interest is whether she lives with her mother (M) or she lives with her father (F).*

The dialogue could be:

1 *M: The child's best interest is that she lives with her mother. [assert]*
  F (Challenge [1])

2 *F: The child's best interest is that she lives with her father. [assert]*
  M (Question [2])

3 *F: Mother is less wealthy than the father because he inherited. Mother cannot provide good living conditions for the child. [Justify]*
M (Question [1])

The dialogue could be annotated with speech acts from [209], i.e., assertion, question, challenge, justify, and retract.

```
<dialogue>

    <statement id="1" type="assert">
        <speaker>M</speaker>
        <content>'Childs best interest is that she lives with her mother.
            </content>
        <response id="challenge1" respondent="F" type="challenge" target=
            "1"/>
    </statement>

    <statement id="2" type="assert">
        <speaker>F</speaker>
        <content>'Childs best interest is that she lives with her father.
            </content>
        <response id="question2" respondent="M" type="question" target="2
            "/>
    </statement>

    <statement id="3" type="justify">
        <speaker>F</speaker>
        <content>Mother is less wealthy than the father, because he
            inherited. Mother cannot provide good living conditions for
            the child.</content>
    </statement>

</dialogue>
```

Different from the annotations used in the second example, we are trying to annotate the dialogue with speech acts [209]. This approach could provide a structured framework to understand the intent and function of each utterance within a conversation. By categorizing statements as assertions, questions, challenges, justification, or other speech acts, we can gain insights into the dynamics of human interaction and the underlying patterns that drive meaningful dialogue. Speech act annotations may serve as the foundation for constructing formal models and contribute to the foundation of formal dialogue. These models can then be used to study and simulate dialogues, leading to a deeper understanding of conversational dynamics. This structured approach may serve as a bridge, facilitating the transition from raw natural language dialogue text to automated dialogues. For foundational models, such annotations offer a roadmap to better comprehend human conversations. By training on data enriched with speech act annotations, the foundational models can more accurately predict and generate responses that align with the intended function of the user's input, leading to more coherent, context-aware, and human-like interactions in automated dialogues.

Example 8.4 (Triple-A: Anything You Say May Be Used against You)**.** *The fourth example is from the work of Arisaka et al. [18]. The authors propose an abstract agent argumentation (Triple-A), a model that distinguishes the global argumentation of judges from the* *local argumentation of accused, prosecutors, witnesses, lawyers, and experts. The agents have partial knowledge of the arguments and attacks of other agents, and they decide autonomously whether to accept or reject their own arguments, and whether to bring their arguments forward in court. The arguments accepted by the judge are based on a game-theoretic equilibrium among the argumentation of the other agents. The theory can be used to distinguish various direct and indirect ways in which the arguments of an agent can be used against his or her other arguments. Given a global abstract agent argumentation framework, various agents view it differently.*

While this paper primarily focuses on static argumentation, its methodology suggests a potential bridge between argumentation as inference and dialogues. The authors apply argumentation semantics to derive justified conclusions. However, they also explore varied scenarios where agents employ diverse strategies of showing their local arguments, which in turn influence the semantics. It offers an interesting perspective on multi-agent interactions within formal dialogues, especially in settings like courtrooms and policy-making, etc. A promising direction for future research lies in exploring the dynamics of these interactions when agents possess partial knowledge. How might agents strategically present arguments when they're uncertain of others' knowledge bases? This model could also pave the way for developing advanced formal dialogue systems that consider the interactions between agents with varying degrees of information and goals.

The fifth example is dialogue games proposed by Martin Carminada [68]. He interprets a number of mainstream argumentation semantics by means of structured discussion. For example, the following is the discussion principle based on grounded semantics.

Definition 8.1. *Let (Ar, att) be an argumentation framework. A* grounded discussion *is a sequence of discussion moves constructed by applying the following principles.*

- *BASIS (HTB): If $A \in Ar$ then $[\texttt{HTB}(A)]$ is a grounded discussion*
- *STEP (HTB): If $[M_1, ..., M_n]$ $(n \geq 1)$ is a grounded discussion without HTB-CB repeats[1], and no CONCEDE or RETRACT move is applicable, and $M_n = \texttt{CB}(A)$ and B is an attacker of A then $[M_1, ..., M_n, \texttt{HTB}(B)]$ is also a grounded discussion*
- *STEP (CB): If $[M_1, ..., M_n]$ $(n \geq 1)$ is a grounded discussion without HTB-CB repeats, and no CONCEDE or RETRACT move is applicable, and M_n is not a CB move, and there is a move $M_i = \texttt{HTB}(A)$ $(i \in \{1, ..., n\})$ such that the discussion does not contain $\texttt{CONCEDE}(A)$, and for each move $M_j = \texttt{HTB}(A')$ $(j > i)$ the discussion contains a move $\texttt{CONCEDE}(A')$, and B is an attacker of A such that the discussion does not contain a move $\texttt{RETRACT}(B)$, then $[M_1, ..., M_n, \texttt{CB}(B)]$ is a grounded discussion*
- *STEP (CONCEDE): If $[M_1, .., M_n]$ $(n \geq 1)$ is a grounded discussion without HTB-CB repeats, and $\texttt{CONCEDE}(B)$ is applicable then $[M_1, ..., M_n, \texttt{CONCEDE}(B)]$ is a grounded discussion*
- *STEP (RETRACT): If $[M_1, ..., M_n]$ $(n \geq 1)$ is a grounded discussion without HTB-CB repeats, and $\texttt{RETRACT}(B)$ is applicable then $[M_1, ..., M_n, \texttt{RETRACT}(B)]$ is a grounded discussion.*

[1] We say that there is a HTB-CB repeat iff $\exists i, j \in \{1, ..., n\} \exists A \in Ar : (M_i = \texttt{HTB}(A) \vee M_i = \texttt{CB}(A)) \wedge (M_j = \texttt{HTB}(A) \vee M_j = \texttt{CB}(A)) \wedge i \neq j$.

Definition 8.2. *A grounded discussion* $[M_1, ..., M_n]$ *is called terminated iff there exists no move* M_{n+1} *such that* $[M_1, ..., M_n, M_{n+1}]$ *is a grounded discussion. A terminated grounded discussion (with* A *being the main argument) is won by the proponent iff the discussion contains* (A)*, otherwise it is won by the opponent.*

The discussion protocols (which we will often refer to as "discussion games") can be seen as proof procedures of their associated argumentation semantics□[68]. These strict discussion protocols are close to the notion of speech acts and communication protocols discussed by McBurney and Parsons [209, 210, 207], and more recent work by Singh and Chopra that uses these protocols as the guiding principle for complex systems engineering [90]. As expected by Martin Caminada, one of the possible applications of the discussion games is for the purpose of human-computer interaction.

The five examples reflect different aspects of future work on how to bridge distributed argumentation technology to dialogue technology. All these examples stem from the realm of computational argumentation. However, it's noteworthy that argumentation as dialogue, despite being a crucial component of argumentation, hasn't been as extensively explored as argumentation as inference. As foundational models emerge, formal dialogue presently plays a significant role in AI advancements. The real challenge is to advance from existing models to more sophisticated dialogue frameworks, some possible directions are shown across the five examples. These examples call for application-driven or example-driven research and a conceptual approach to guide the investigation.

Chapter 9

Summary

This thesis proposes distributed argumentation technology, a computational approach that incorporates argumentation reasoning mechanisms within multi-agent systems. For the formal foundations of distributed argumentation technology, in this thesis, we conduct a principle-based analysis of structured argumentation as well as abstract multi-agent and abstract bipolar argumentation.

- **Structured argumentation:** In Chapter 2, we compare weakest link principle and last link principle following the approach of Dung's two seminal papers [112, 115]. Additionally, we introduce a new definition of weakest link attack relation assignment based on lookahead, and compare this new lookahead definition with two existing ones in the literature using a principle-based analysis. We show that our lookahead definition does not satisfy context independence, we introduce a new principle called weak context independence, and we show that lookahead weakest link satisfies weak context independence. We also show that lookahead weakest link is the closest approximation to Brewka's prioritised default logic (PDL), known as the greedy approach. For PDL, we prove an impossibility result under Dung's axioms. Our results generalize earlier findings restricted to total orders to the more general case of modular orders.

- **Abstract Argumentation:**

  - **Abstract agent argumentation:** In Chapter 3, we study four types of semantics for them. First, agent defense semantics replaces Dung's notion of defense by some kind of agent defense. Second, social agent semantics prefers arguments that belong to more agents. Third, agent reduction semantics considers the perspective of individual agents. Fourth, agent filtering semantics are inspired by a lack of knowledge. We study five existing principles and we introduce twelve new ones. In total, we provide a full analysis of fifty-two agent semantics and the seventeen principles.

  - **Abstract Bipolar argumentation:** In Chapter 4, we introduce and study seven types of semantics for bipolar argumentation frameworks, each extending Dung's interpretation of attack with a distinct interpretation of support. First, we introduce three types of defence-based semantics by adapting the notions of defence. Second, we examine two types of selection-based semantics that select exten-

sions by counting the number of supports. Third, we analyze two types of traditional reduction-based semantics under deductive and necessary interpretations of support. We provide a full analysis of twenty-eight bipolar argumentation semantics and ten principles in total. In addition, we consider a legal divorce action in which the interpretation of support is in close relation to the interpretation of the law itself, where an argument may support another one, but we do not know yet which kind of support it is. Different agents have different interpretations of legal rules, which lead to various verdicts. However, this consideration does not invalidate our work on the principle-based approach. On the contrary, because of the ambiguity at the pragmatic and semantic level, a principle-based approach can be very useful to better understand the choices of a particular formalization. This serves the legal context by providing a more abstract conceptualization and analysis of the fundamental concept of formal argumentation.

Moreover, in Chapter 5, we propose distributed argumentation technology using distributed ledgers. We envision the Intelligent Human-input-based Blockchain Oracle (IHiBO), an artificial intelligence tool for storing argumentation reasoning. We present a decentralized and secure architecture for conducting decision-making, addressing key concerns of trust, transparency, and immutability. We model fund management with agent argumentation in IHiBO and analyze its compliance with European fund management legal frameworks. The approach explores integrating formal argumentation with contemporary technologies, thereby advancing the gap between theoretical constructs and real-world applications, i.e. The employment of blockchain technology for formal argumentation offers a modest attempt to bridge theory and practice. As a follow-up, in Chapter 6, we discuss how distributed argumentation technology can be used to advance risk management, regulatory compliance of distributed ledgers for financial securities,

In Chapter 7, we give a big picture of formal argumentation. First is formal argumentation vs. computational argumentation, then the branches of argumentation as inference and argumentation as dialogue, and we also provide a comprehensive overview concerning the three aspects we explored in this thesis: argument strength, multi-agent argumentation, and support in argumentation. Lastly, we briefly discuss argumentation used in legal AI. This allows us to position our contributions within the broader intellectual landscape of this field, thus underscoring the significance of our work.

Further, in Chapter 8, we explore the relationship between distributed argumentation technology and dialogue technology. We first discuss how to bridge argumentation as dialogue and argumentation as inference via dialogue games. Then we show five examples in the literature of formal argumentation. The five examples can be seen as future use-case studies that reflect different aspects of bridging distributed argumentation technology and dialogue technology.

The journey embarked upon in this doctoral thesis may conclude here, but the path it has blazed toward the intersection of law, science, and technology has just begun. As the understanding and capabilities in these realms grow, so too will the potential applications of the concepts and frameworks established within this thesis.

Bibliography

[1] Adiyanto Adiyanto and Rizki Febrianto. Authentication of transaction process in e-marketplace based on blockchain technology. *Aptisi Transactions On Technopreneurship (ATT)*, 2(1):68–74, 2020. 102

[2] John Adler, Ryan Berryhill, Andreas Veneris, Zissis Poulos, Neil Veira, and Anastasia Kastania. Astraea: A decentralized blockchain oracle. In *2018 IEEE iThings and IEEE GreenCom and IEEE CPSCom and IEEE SmartData*, pages 1145–1152. IEEE, 2018. 93

[3] Deirdre Ahern. Regulatory lag, regulatory friction and regulatory transition as fintech disenablers: calibrating an eu response to the regulatory sandbox phenomenon. *European Business Organization Law Review*, 22(3):395–432, 2021. 99

[4] Kamalani Aiyar, Malka N Halgamuge, and Azeem Mohammad. Probability distribution model to analyze the trade-off between scalability and security of sharding-based blockchain networks. In *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, pages 1–6. IEEE, 2021. 102

[5] Carlos E Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The journal of symbolic logic*, 50(2):510–530, 1985. 11

[6] Leila Amgoud and Jonathan Ben-Naim. Evaluation of arguments in weighted bipolar graphs. *International Journal of Approximate Reasoning*, 99:39–55, 2018. 62, 71

[7] Leila Amgoud, Jean-François Bonnefon, and Henri Prade. An argumentation-based approach to multiple criteria decision. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty: 8th European Conference, ECSQARU 2005, Barcelona, Spain, July 6-8, 2005. Proceedings 8*, pages 269–280. Springer, 2005. 117

[8] Leila Amgoud and Claudette Cayrol. Inferring from inconsistency in preference-based argumentation frameworks. *International Journal of Approximate Reasoning*, 29(2):125–169, 2002. 34, 49

[9] Leila Amgoud and Souhila Kaci. An argumentation framework for merging conflicting knowledge bases. *International Journal of Approximate Reasoning*, 45(2):321–340, 2007. 115

[10] Leila Amgoud and Srdjan Vesic. Rich preference-based argumentation frameworks. *International Journal of Approximate Reasoning*, 55(2):585–606, 2014. 39, 49, 71

[11] Hamed Amini, Damir Filipović, and Andreea Minca. Systemic risk and central clearing counterparty design. *Swiss Finance Institute Research Paper*, (13-34), 2015. 96

[12] Bo An, Nicola Gatti, and Victor Lesser. Extending alternating-offers bargaining in one-to-many and many-to-many settings. In *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, volume 2, pages 423–426. IEEE, 2009. 79

[13] Lennart Ante. Smart contracts on the blockchain–a bibliometric analysis and review. *Telematics and Informatics*, 57:101519, 2021. 105

[14] Grigoris Antoniou, David Billington, Guido Governatori, and Michael J Maher. Representation results for defeasible logic. *ACM Transactions on Computational Logic (TOCL)*, 2(2):255–287, 2001. 120

[15] Shahina Anwarul and Susheela Dahiya. A comprehensive review on face recognition methods and factors affecting facial recognition accuracy. *Proceedings of ICRIC 2019: Recent Innovations in Computing*, pages 495–514, 2020. 2

[16] Ofer Arieli and Christian Straßer. Sequent-based logical argumentation. *Argument & Computation*, 6(1):73–99, 2015. 111

[17] Ryuta Arisaka, Jérémie Dauphin, Ken Satoh, and Leendert van der Torre. Multi-agent argumentation and dialogue. *FLAP*, 9(4):921–954, 2022. 6, 13, 116, 122, 124

[18] Ryuta Arisaka, Ken Satoh, and Leendert van der Torre. Anything you say may be used against you in a court of law: Abstract agent argumentation (triple-a). In *AI Approaches to the Complexity of Legal Systems: AICOL International Workshops 2015-2017: AICOL-VI@ JURIX 2015, AICOL-VII@ EKAW 2016, AICOL-VIII@ JURIX 2016, AICOL-IX@ ICAIL 2017, and AICOL-X@ JURIX 2017, Revised Selected Papers*, pages 427–442. Springer, 2018. 128

[19] Ryuta Arisaka, Ken Satoh, and Leendert W. N. van der Torre. Anything you say may be used against you in a court of law - abstract agent argumentation (triple-a). In Ugo Pagallo, Monica Palmirani, Pompeu Casanovas, Giovanni Sartor, and Serena Villata, editors, *AI Approaches to the Complexity of Legal Systems*, volume 10791, pages 427–442. Springer, 2017. 34, 49

[20] Dev Arora, Siddharth Gautham, Harshit Gupta, and Bharat Bhushan. Blockchain-based security solutions to preserve data privacy and integrity. In *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, pages 468–472. IEEE, 2019. 102

[21] Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 54(4):328–346, 1950. 30

[22] Kenneth J Arrow. *Social choice and individual values*, volume 12. Yale university press, 2012. 117

[23] Kenneth J Arrow, Amartya Sen, and Kotaro Suzumura. *Handbook of social choice and welfare*, volume 2. Elsevier, 2010. 11

[24] Katie Atkinson, Pietro Baroni, Massimiliano Giacomin, Anthony Hunter, Henry Prakken, Chris Reed, Guillermo Simari, Matthias Thimm, and Serena Villata. Towards artificial argumentation. *AI magazine*, 38(3):25–36, 2017. 109

[25] Emilios Avgouleas and Alexandros Seretakis. Governing the digital finance value-chain in the eu: Mifid ii, the digital package, and the large gaps between! *European Company Law and Financial Law Review*, pages 1–35, 2021. 103

[26] Edmond Awad, Jean-François Bonnefon, Martin Caminada, Thomas W Malone, and Iyad Rahwan. Experimental assessment of aggregation principles in argumentation-enabled collective intelligence. *ACM Transactions on Internet Technology (TOIT)*, 17(3):1–21, 2017. 112, 116, 117

[27] Edmond Awad, Richard Booth, Fernando Tohmé, and Iyad Rahwan. Judgement aggregation in multi-agent argumentation. *J. Log. Comput.*, 27(1):227–259, 2017. 33

[28] Tim Baarslag, Koen Hindriks, and Catholijn Jonker. Acceptance conditions in automated negotiation. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, pages 95–111. Springer, 2013. 96

[29] Andrew B Baker and Matthew L Ginsberg. A theorem prover for prioritized circumscription. In *IJCAI*, pages 463–467. Citeseer, 1989. 4

[30] Tom Barbereau, Reilly Smethurst, Orestis Papageorgiou, Johannes Sedlmeir, and Gilbert Fridgen. Decentralised finance's unregulated governance: Minority rule in the digital wild west. *Available at SSRN*, 2022. 93

[31] P. Baroni, D. Gabbay, and M. Giacomin. *Handbook of Formal Argumentation*. College Publications, 2018. 3, 4, 6, 7

[32] Pietro Baroni, Guido Boella, Federico Cerutti, Massimiliano Giacomin, Leendert Van Der Torre, and Serena Villata. On the input/output behavior of argumentation frameworks. *Artificial Intelligence*, 217:144–197, 2014. 49, 116

[33] Pietro Baroni, Martin Caminada, and Massimiliano Giacomin. An introduction to argumentation semantics. *The knowledge engineering review*, 26(4):365–410, 2011. 11, 86

[34] Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre, editors. *Handbook of Formal Argumentation*, volume 1. College Publications, 2018. 33

[35] Pietro Baroni and Massimiliano Giacomin. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence*, 171(10-15):675–700, 2007. 10, 34, 35, 42, 43, 47, 65, 71, 161, 167, 177, 178, 190, 191, 195

[36] Pietro Baroni, Massimiliano Giacomin, and Giovanni Guida. Scc-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence*, 168(1-2):162–210, 2005. 42, 43, 44, 49, 66, 162, 163, 171

[37] Pietro Baroni, Antonio Rago, and Francesca Toni. How many properties do we need for gradual argumentation? In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018. 71

[38] Pietro Baroni, Marco Romano, Francesca Toni, Marco Aurisicchio, and Giorgio Bertanza. Automatic evaluation of design alternatives with quantitative argumentation. *Argument & Computation*, 6(1):24–49, 2015. 71

[39] Ringo Baumann, Gerhard Brewka, and Markus Ulbricht. Comparing weak admissibility semantics to their Dung-style counterparts–reduct, modularization, and strong equivalence in abstract argumentation. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning*, volume 17, pages 79–88, 2020. 44, 49, 71, 179

[40] Moritz Becker and Balázs Bodó. Trust in blockchain-based systems. *Internet Policy Review*, 10(2), 2021. 76

[41] Mathieu Beirlaen, Jesse Heyninck, Pere Pardo, and Christian Straßer. Argument strength in formal argumentation. *Journal of Logics and their Applications*, 5(3):629–676, 2018. 31, 113, 114, 115

[42] Trevor Bench-Capon, Henry Prakken, and Giovanni Sartor. *Argumentation in legal reasoning*. Springer, 2009. 122

[43] Trevor Bench-Capon and Giovanni Sartor. A model of legal reasoning with cases incorporating theories and values. *Artificial Intelligence*, 150(1-2):97–143, 2003. 122

[44] Salem Benferhat, Didier Dubois, Souhila Kaci, and Henri Prade. Bipolar representation and fusion of preferences on the possibilistic logic framework. *KR*, 2:421–432, 2002. 117

[45] Salem Benferhat, Didier Dubois, and Henri Prade. Argumentative inference in uncertain and inconsistent knowledge bases. In *Uncertainty in Artificial Intelligence*, pages 411–419. Elsevier, 1993. 114

[46] Abdeljalil Beniiche. A study of blockchain oracles. *arXiv preprint arXiv:2004.07140*, 2020. 78, 93

[47] Evangelos Benos, Rod Garratt, and Pedro Gurrola-Perez. The economics of distributed ledger technology for securities settlement. *Available at SSRN 3023779*, 2017. 102

[48] Ph Besnard et al. Semantics for evidence-based argumentation. In *Computational models of argument: proceedings of COMMA*, volume 172, page 276, 2008. 119

[49] Philippe Besnard and Anthony Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128(1-2):203–235, 2001. 111

[50] Philippe Besnard and Anthony Hunter. *Elements of Argumentation*. MIT Press, 2008. 16

[51] Philippe Besnard and Anthony Hunter. *Elements of Argumentation*. MIT Press, 2008. 113

[52] Philippe Besnard and Anthony Hunter. A review of argumentation based on deductive arguments. *Handbook of Formal Argumentation*, 1:437–484, 2018. 111

[53] Vinith V Bhandarkar, Akshay A Bhandarkar, and Aditya Shiva. Digital stocks using blockchain technology the possible future of stocks? *International Journal of Management (IJM)*, 10(3), 2019. 84

[54] Kariappa Bheemaiah. *The blockchain alternative: rethinking macroeconomic policy and economic theory*. Apress, 2017. 101

[55] Stéphane Blemus. Law and blockchain: A legal perspective on current regulatory trends worldwide. *Revue Trimestrielle de Droit Financier (Corporate Finance and Capital Markets Law Review) RTDF*, (4-2017), 2017. 102

[56] Alexander Bochman. Propositional argumentation and causal reasoning. In *International Joint Conference on Artificial Intelligence*, volume 19, page 388. LAWRENCE ERLBAUM ASSOCIATES LTD, 2005. 111

[57] Guido Boella, Dov M Gabbay, Le en de r t van der Torre, and Serena Villata. Support in abstract argumentation. In *Proceedings of the Third International Conference on Computational Models of Argument (COMMA'10)*, pages 40–51. Frontiers in Artificial Intelligence and Applications, IOS Press, 2010. 53, 65, 118, 119

[58] Andrei Bondarenko, Phan Minh Dung, Robert A Kowalski, and Francesca Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial intelligence*, 93(1-2):63–101, 1997. 4, 16, 111

[59] Richard Booth, Souhila Kaci, Tjitze Rienstra, and Leendert van der Torre. A logical theory about dynamics in abstract argumentation. In *International Conference on Scalable Uncertainty Management*, pages 148–161. Springer, 2013. 72

[60] Craig Boutilier. Toward a logic for qualitative decision theory. In *Principles of knowledge representation and reasoning*, pages 75–86. Elsevier, 1994. 117

[61] Felix Brandt, Vincent Conitzer, and Ulle Endriss. Computational social choice. *Multiagent systems*, 2:213–284, 2012. 117

[62] G. Brewka and T. Eiter. Preferred answer sets for extented logic programs. *Artif. Intell.*, 109:297–356, 1999. 16, 113

[63] Gerhard Brewka. Reasoning about priorities in default logic. In Barbara Hayes-Roth and Richard E. Korf, editors, *Proc. of the 12th National Conference on AI*, volume 2, pages 940–945. AAAI Press / The MIT Press, 1994. 18, 113

[64] Gerhard Brewka and Thomas Eiter. Preferred answer sets for extended logic programs. *Artificial intelligence*, 109(1-2):297–356, 1999. 114

[65] Gerhard Brewka and Thomas Eiter. Prioritizing default logic. In Steffen Hölldobler, editor, *Intellectics and Computational Logic*, volume 19 of *Applied Logic Series*, pages 27–45. Kluwer, 2000. 18

[66] Vitalik Buterin et al. Ethereum white paper, 2013. 80, 83, 84, 88

[67] Elena Cabrio and Serena Villata. Node: A benchmark of natural language arguments. In *Computational Models of Argument*, pages 449–450. IOS Press, 2014. 60, 127

[68] Martin Caminada. Argumentation semantics as formal discussion. *Journal of Applied Logics*, 4(8):2457–2492, 2017. 51, 129, 130

[69] Martin Caminada. Rationality postulates: Applying argumentation theory for non-monotonic reasoning. *FLAP*, 4(8), 2017. 16, 31, 115

[70] Martin Caminada. Argumentation semantics as formal discussion. *Handbook of Formal Argumentation*, 1:487–518, 2018. 112, 117

[71] Martin Caminada and Leila Amgoud. An axiomatic account of formal argumentation. In *AAAI*, volume 6, pages 608–613, 2005. 5, 112

[72] Martin Caminada and Leila Amgoud. On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171(5-6):286–310, 2007. 5, 112, 113, 114

[73] Martin Caminada and Gabriella Pigozzi. On judgment aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems*, 22(1):64–102, 2011. 33, 49, 112, 116

[74] Martin Caminada, Samy Sá, Joao Alcântara, and Wolfgang Dvořák. On the equivalence between logic programming semantics and argumentation semantics. *International Journal of Approximate Reasoning*, 58:87–111, 2015. 111

[75] Martin WA Caminada, Walter A Carnielli, and Paul E Dunne. Semi-stable semantics. *Journal of Logic and Computation*, 22(5):1207–1254, 2012. 5, 112

[76] Martinus Wigbertus Antonius Caminada, Sanjay Modgil, and Nir Oren. Preferences and unrestricted rebut. *Computational Models of Argument*, 2014. 111

[77] Claudette Cayrol, Andrea Cohen, and Marie-Christine Lagasquie-Schiex. Higher-order interactions (bipolar or not) in abstract argumentation: A state of the art. In Dov Gabbay, Massimiliano Giacomin, Guillermo R. Simari, and Matthias Thimm, editors, *Handbook of Formal Argumentation, Volume 2*, pages 15–130. College Publications, 2021. 71

[78] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Gradual valuation for bipolar argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 366–377. Springer, 2005. 71

[79] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 378–389. Springer, 2005. 10, 53, 118, 120

[80] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 378–389. Springer, 2005. 53, 63

[81] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Bipolar abstract argumentation systems. In *Argumentation in Artificial Intelligence*, pages 65–84. Springer, 2009. 10, 53, 63, 118, 120

[82] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Coalitions of arguments: A tool for handling bipolar argumentation frameworks. *International Journal of Intelligent Systems*, 25(1):83–109, 2010. 10, 53, 118

[83] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Bipolarity in argumentation graphs: Towards a better understanding. *International Journal of Approximate Reasoning*, 54(7):876–899, 2013. 10, 42, 53, 54, 63, 65, 118

[84] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. An axiomatic approach to support in argumentation. In *International Workshop on Theory and Applications of Formal Argumentation*, pages 74–91. Springer, 2015. 53, 71, 118

[85] Fernando Cerezetti, Max Chan, and Rafael Plata. Decentralized clearing? an assessment of the impact of dlts on ccps. *An Assessment of the Impact of DLTs on CCPs (February 14, 2023)*, 2023. 100

[86] Chainsafe. Chainbridge, 2022. 83

[87] James Chen. Fund manager, 2021. 75

[88] Weiwei Chen. Aggregation of support-relations of bipolar argumentation frameworks. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1804–1806, 2020. 71

[89] Ahmad Chokor and Elise Alfieri. Long and short-term impacts of regulation in the cryptocurrency market. *The quarterly review of economics and finance*, 81:157–173, 2021. 107

[90] Amit K. Chopra. Interaction-oriented software engineering: Programming abstractions for autonomy and decentralization. *AI Communications*, 35(4):381–391, 2022. 130

[91] Gordon L Clark. Financial intermediation, infrastructure investment and regional growth. *Area Development and Policy*, 2(3):217–236, 2017. 98

[92] Andrea Cohen, Sebastian Gottifredi, Alejandro Javier García, and Guillermo Ricardo Simari. A survey of different approaches to support in argumentation systems. *Knowl. Eng. Rev.*, 29(5):513–550, 2014. 53, 118, 120

[93] Robert G Cooper, Scott J Edgett, and Elko J Kleinschmidt. Portfolio management in new product development: Lessons from the leaders—i. *Research-Technology Management*, 40(5):16–28, 1997. 73, 75

[94] Robert Gravlin Cooper. *Winning at new products*. Addison-Wesley Reading, MA, 1986. 75

[95] 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. 33

[96] Sylvie Coste-Marquis, Caroline Devred, Sébastien Konieczny, Marie-Christine Lagasquie-Schiex, and Pierre Marquis. On the merging of dung's argumentation systems. *Artificial Intelligence*, 171(10-15):730–753, 2007. 112, 116

[97] Sylvie Coste-Marquis, Sébastien Konieczny, Pierre Marquis, and Mohand Akli Ouali. Selecting extensions in weighted argumentation frameworks. *COMMA*, 12:342–349, 2012. 71

[98] Raffaele Della Croce, J Paula, and A Laboul. Infrastructure financing instruments and incentives. *Organ Econ Co-operation Dev (OECD)*, 2015. 98

[99] Kristijonas Čyras, Ken Satoh, and Francesca Toni. Explanation for case-based reasoning via abstract argumentation. In *Computational Models of Argument*, pages 243–254. IOS Press, 2016. 74

[100] Matija Damjan. The interface between blockchain and the real world. *Ragion pratica*, pages 379–406, 2018. 93

[101] Jérémie Dauphin, Tjitze Rienstra, and Leendert van der Torre. New weak admissibility semantics for abstract argumentation. In *Logic and Argumentation: 4th International Conference, CLAR 2021, Hangzhou, China, October 20–22, 2021, Proceedings 4*, pages 112–126. Springer, 2021. 71

[102] Stefano De Angelis. Assessing security and performances of consensus algorithms for permissioned blockchains. *arXiv preprint arXiv:1805.03490*, 2018. 81, 88

[103] James P. Delgrande and Torsten Schaub. Expressing preferences in default logic. *Artificial Intelligence*, 123(1-2):41–87, 2000. 31, 113

[104] James P. Delgrande and Torsten Schaub. Expressing preferences in default logic. *Artif. Intell.*, 123(1-2):41–87, 2000. 114

[105] James P. Delgrande, Torsten Schaub, Hans Tompits, and Kewen Wang. A classification and survey of preference handling approaches in nonmonotonic reasoning. *Computational Intelligence*, 20(2):308–334, 2004. 17

[106] Jérôme Delobelle, Sébastien Konieczny, and Srdjan Vesic. On the aggregation of argumentation frameworks: operators and postulates. *J. Log. Comput.*, 28(7):1671–1699, 2018. 33

[107] Huimin Dong, Beishui Liao, Réka Markovich, and Leendert van der Torre. Defeasible deontic logic: Arguing about permission and obligation. *IfCoLog Journal of Logics and Their Applications*, 9(4):957–1018, 2022. 122

[108] Didier Dubois and Hélène Fargier. On the qualitative comparison of sets of positive and negative affects. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty: 8th European Conference, ECSQARU 2005, Barcelona, Spain, July 6-8, 2005. Proceedings 8*, pages 305–316. Springer, 2005. 117

[109] Darrell Duffie, Nicolae Gârleanu, and Lasse Heje Pedersen. Over-the-counter markets. *Econometrica*, 73(6):1815–1847, 2005. 96

[110] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357, 1995. 5, 6, 15, 33, 35, 54, 55, 112, 116, 117

[111] Phan Minh Dung. An axiomatic analysis of structured argumentation for prioritized default reasoning. volume 263 of *Frontiers in Artificial Intelligence and Applications*, pages 267–272. IOS Press, 2014. 16

[112] Phan Minh Dung. An axiomatic analysis of structured argumentation with priorities. *Artif. Intell.*, 231:107–150, 2016. 13, 16, 18, 19, 20, 21, 22, 23, 29, 32, 131

[113] Phan Minh Dung. A canonical semantics for structured argumentation with priorities. In Pietro Baroni, Thomas F. Gordon, Tatjana Scheffler, and Manfred Stede, editors, *Computational Models of Argument - Proceedings of COMMA*, volume 287 of *Frontiers in Artificial Intelligence and Applications*, pages 263–274. IOS Press, 2016. 9, 16, 22, 28, 29

[114] Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. Assumption-based argumentation. In Guillermo Ricardo Simari and Iyad Rahwan, editors, *Argumentation in Artificial Intelligence*, pages 199–218. Springer, 2009. 17

[115] Phan Minh Dung and Phan Minh Thang. Fundamental properties of attack relations in structured argumentation with priorities. *Artificial Intelligence*, 255:1–42, 2018. 9, 16, 18, 32, 131

[116] Phan Minh Dung, Phan Minh Thang, and Tran Cao Son. On structured argumentation with conditional preferences. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI*, pages 2792–2800. AAAI Press, 2019. 16

[117] Thomas Duttle, Katharina Holzinger, Thomas Malang, Thomas Schäubli, Frank Schimmelfennig, and Thomas Winzen. Opting out from european union legislation: the differentiation of secondary law. *Journal of European Public Policy*, 24(3):406–428, 2017. 96

[118] Nabil El Ioini and Claus Pahl. A review of distributed ledger technologies. In *On the Move to Meaningful Internet Systems. OTM 2018 Conferences: Confederated International Conferences: CoopIS, C&TC, and ODBASE 2018, Valletta, Malta, October 22-26, 2018, Proceedings, Part II*, pages 277–288. Springer, 2018. 4

[119] Edith Elkind, Talal Rahwan, and Nicholas R Jennings. Computational coalition formation. *Multiagent systems*, pages 329–380, 2013. 112, 116

[120] Edith Elkind, Talal Rahwan, and Nicholas R Jennings. Game theoretic foundations of multiagent systems. *Multiagent systems*, pages 811–848, 2013. 112, 116

[121] Steve Ellis, Ari Juels, and Sergey Nazarov. Chainlink a decentralized oracle network, 2017. 93

[122] Antonio Estache and Atsushi Iimi. Bidder asymmetry in infrastructure procurement: Are there any fringe bidders? *Review of Industrial Organization*, 36:163–187, 2010. 98

[123] Valentinos Evripidou and Francesca Toni. Quaestio-it. com: a social intelligent debating platform. *Journal of Decision Systems*, 23(3):333–349, 2014. 71

[124] Rino Falcone and Cristiano Castelfranchi. Social trust: A cognitive approach. In *Trust and deception in virtual societies*, pages 55–90. Springer, 2001. 91

[125] Peyman Faratin, Carles Sierra, and Nick R Jennings. Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems*, 24(3-4):159–182, 1998. 79, 85, 89

[126] Sara Feenan, Daniel Heller, Alexander Lipton, Massimo Morini, Rhomaios Ram, Robert Sams, Tim Swanson, Stanley Yong, and Diana Barrero Zalles. Decentralized financial market infrastructures. *The Journal of FinTech, Forthcoming*, 2020. 74, 76, 92, 93

[127] Eilís Ferran. *Building an EU securities market*. Cambridge University Press, 2004. 95

[128] Agata Ferreira. Regulating smart contracts: Legal revolution or simply evolution? *Telecommunications Policy*, 45(2):102081, 2021. 107

[129] ET Feteris et al. Legal argumentation in pragma-dialectical perspective. *Dutch Journal for Philosophy of Law and Legal Theory*, 25, 1996. 119

[130] Christian Fisch. Initial coin offerings (icos) to finance new ventures. *Journal of Business Venturing*, 34(1):1–22, 2019. 98

[131] John Fox and Subrata Das. Safe and sound. *Artificial intelligence in hazardous applications*, 307, 2000. 109

[132] Merritt B Fox, Lawrence R Glosten, Edward F Greene, and Sue S Guan. Distributed ledger technology and the securities markets of the future: a stakeholder survey. *Colum. Bus. L. Rev.*, page 651, 2021. 100

[133] Jozo Frankovic, Bin Liu, and Sandy Suardi. On spillover effects between cryptocurrency-linked stocks and the cryptocurrency market: Evidence from australia. *Global Finance Journal*, 54:100642, 2022. 100

[134] D. Gabbay, M. Giacomin, and G Simari. *Handbook of Formal Argumentation*. College Publications, 2021. 3, 4, 33

[135] Dov M Gabbay, Massimiliano Giacomin, Beishui Liao, and Leon van der Torre. Present and future of formal argumentation (dagstuhl perspectives workshop 15362). *Dagstuhl Manifestos*, 7(1):69–95, 2018. 11, 12, 109

[136] Dov M Gabbay and Gadi Rozenberg. Introducing abstract argumentation with many lives. *IfCoLog Journal of Logics and their Applications*, 7(3):295–336, 2020. 62

[137] Yang Gao, Francesca Toni, Hao Wang, and Fanjiang Xu. Argumentation-based multi-agent decision making with privacy preserved. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 1153–1161, 2016. 33

[138] Alejandro J Garcıa and Guillermo R Simari. Defeasible logic programming. Technical report, Technical report, Computer Science Department, Universidad Nacional del Sur, 1998. 4, 110

[139] Alejandro J García and Guillermo R Simari. Defeasible logic programming: An argumentative approach. *Theory and practice of logic programming*, 4(1-2):95–138, 2004. 4, 110

[140] Anis Gargouri, Sébastien Konieczny, Pierre Marquis, and Srdjan Vesic. On a notion of monotonic support for bipolar argumentation frameworks. In *20th International Conference on Autonomous Agents and MultiAgent Systems*, 2020. 54, 62, 65, 71, 118

[141] James W Garson. *Modal logic for philosophers*. Cambridge University Press, 2013. 3

[142] Hector Geffner and Judea Pearl. Conditional entailment: Bridging two approaches to default reasoning. *Artificial Intelligence*, 53(2-3):209–244, 1992. 4

[143] Massimiliano Giacomin. Handling heterogeneous disagreements through abstract argumentation (extended abstract). In Bo An, Ana L. C. Bazzan, João Leite, Serena Villata, and Leendert W. N. van der Torre, editors, *PRIMA 2017: Principles and Practice of Multi-Agent Systems*, volume 10621 of *Lecture Notes in Computer Science*, pages 3–11. Springer, 2017. 34, 49

[144] GnosisDAO. Gnosis whitepaper, 2017. 93

[145] Nelson Goodman. *Fact, fiction, and forecast*. Harvard University Press, 1983. 28

[146] Thomas Gordon. An overview of the carneades argumentation support system. 2010. 120

[147] Thomas F Gordon. Towards requirements analysis for formal argumentation. In Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre, editors, *Handbook of formal argumentation, Volume 1*, pages 145–156. College Publictions, 2018. 71, 118

[148] Thomas F Gordon, Henry Prakken, and Douglas Walton. The carneades model of argument and burden of proof. *Artificial Intelligence*, 171(10-15):875–896, 2007. 6, 120

[149] Thomas F Gordon and Douglas Walton. Proof burdens and standards. In *Argumentation in artificial intelligence*, pages 239–258. Springer, 2009. 120

[150] Nikos Gorogiannis and Anthony Hunter. Instantiating abstract argumentation with classical logic arguments: Postulates and properties. *Artificial Intelligence*, 175(9-10):1479–1497, 2011. 16

[151] Guido Governatori, Trevor Bench-Capon, Bart Verheij, Michał Araszkiewicz, Enrico Francesconi, and Matthias Grabmair. Thirty years of artificial intelligence and law: the first decade. *Artificial Intelligence and Law*, 30(4):481–519, 2022. 3, 122

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

[153] Guido Governatori, Michael J. Maher, Grigoris Antoniou, and David Billington. Argumentation semantics for defeasible logic. *J. Log. Comput.*, 14(5):675–702, 2004. 15

[154] Floriana Grasso, Alison Cawsey, and Ray Jones. Dialectical argumentation to solve conflicts in advice giving: a case study in the promotion of healthy nutrition. *International Journal of Human-Computer Studies*, 53(6):1077–1115, 2000. 109

[155] Lewis Gudgeon, Pedro Moreno-Sanchez, Stefanie Roos, Patrick McCorry, and Arthur Gervais. Sok: Layer-two blockchain protocols. In *International Conference on Financial Cryptography and Data Security*, pages 201–226. Springer, 2020. 80, 81, 83, 92

[156] Sven Ove Hansson and David Makinson. Applying normative rules with restraint. In *Logic and Scientific Methods: Volume One of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence, August 1995*, pages 313–332. Springer, 1997. 113

[157] Loha Hashimy and Philipp Sandner. The impact of financial regulation on the development of distributed ledger technology (dlt) firms. *Frontiers in Blockchain*, 3:21, 2020. 104

[158] Snoeck Henkemans and Arnolda Francisca. Analysing complex argumentation: The reconstruction of multiple and coordinatively compound argumentation in a critical discussion. 1992. 119

[159] Jesse Heyninck. *Investigations into the logical foundations of defeasible reasoning: an argumentative perspective.* PhD thesis, Ruhr University Bochum, Germany, 2019. 111

[160] Sebastiaan Niels Hooghiemstra. Distributed ledger technology ('dlt') and its impact (on the regulation of) european investment funds. *Available at SSRN 3735886*, 2020. 91

[161] John F. Horty. Argument construction and reinstatement in logics for defeasible reasoning. *Artif. Intell. Law*, 9(1):1–28, 2001. 37

[162] Matthew B Hoy. Alexa, siri, cortana, and more: an introduction to voice assistants. *Medical reference services quarterly*, 37(1):81–88, 2018. 2

[163] Anthony Hunter, Sylwia Polberg, and Matthias Thimm. Epistemic graphs for representing and reasoning with positive and negative influences of arguments. *Artif. Intell.*, 281:103236, 2020. 49

[164] Anthony Hunter and Matthew Williams. Aggregating evidence about the positive and negative effects of treatments. *Artificial intelligence in medicine*, 56(3):173–190, 2012. 109

[165] Michael Huth and Mark Ryan. *Logic in Computer Science: Modelling and reasoning about systems*. Cambridge university press, 2004. 2

[166] Trung Dong Huynh, Nicholas R Jennings, and Nigel R Shadbolt. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 13:119–154, 2006. 112, 116

[167] Suk Hyun, Donghyun Park, and Shu Tian. The price of going green: the role of greenness in green bond markets. *Accounting & Finance*, 60(1):73–95, 2020. 106

[168] Roger G Ibbotson and Jay R Ritter. Initial public offerings. *Handbooks in operations research and management science*, 9:993–1016, 1995. 96

[169] Henry Alexander Ignatious, Manzoor Khan, et al. An overview of sensors in autonomous vehicles. *Procedia Computer Science*, 198:736–741, 2022. 2

[170] International Token Standardization Association. List of synthetix derivative tokens, 2021. 83, 84

[171] Philip C Jackson. *Introduction to artificial intelligence*. Courier Dover Publications, 2019. 1

[172] Mathilde Janier, Mark Snaith, Katarzyna Budzynska, John Lawrence, and Chris Reed. A system for dispute mediation: The mediation dialogue game. In *The 6th International Conference on Computational Models of Argument*, pages 351–358. IOS Press, 2016. 109

[173] Anna Jobin, Marcello Ienca, and Effy Vayena. The global landscape of ai ethics guidelines. *Nature Machine Intelligence*, 1(9):389–399, 2019. 2

[174] Souhila Kaci, Leendert van der Torre, Srdjan Vesic, and Serena Villata. Preference in abstract argumentation. In Dov Gabbay, Massimiliano Giacomin, Guillermo R. Simari, and Matthias Thimm, editors, *Handbook of Formal Argumentation, Volume 2*, pages 211–248. College Publications, 2021. 49, 65, 71, 72

[175] Muhammad Kamran, Pakeezah Butt, Assim Abdel-Razzaq, and Hadrian Geri Djajadikerta. Is bitcoin a safe haven? application of fintech to safeguard australian stock markets. *Studies in Economics and Finance*, 39(3):386–402, 2022. 100

[176] Nikos Karanikolas, Pierre Bisquert, and Christos Kaklamanis. A voting argumentation framework: Considering the reasoning behind preferences. In *ICAART 2019-11th International Conference on Agents and Artificial Intelligence*, volume 1, pages 42–53. SCITEPRESS-Science and Technology Publications, 2019. 33

[177] Beomjoong Kim, Hyoung Joong Kim, and Junghee Lee. First smart contract allowing cryptoasset recovery. *KSII Transactions on Internet & Information Systems*, 16(3), 2022. 106

[178] Michinobu Kishi. Project stella and the impacts of fintech on financial infrastructures in japan. 2019. 100

[179] Sébastien Konieczny, Pierre Marquis, and Srdjan Vesic. On supported inference and extension selection in abstract argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 49–59. Springer, 2015. 71

[180] Kurt Konolige. On the relation between default and autoepistemic logic. *Artificial intelligence*, 35(3):343–382, 1988. 113

[181] Dionysios Kontarinis and Francesca Toni. Identifying malicious behavior in multi-party bipolar argumentation debates. In Michael Rovatsos, George A. Vouros, and Vicente Julián, editors, *Multi-Agent Systems and Agreement Technologies - 13th European Conference, EUMAS 2015, and Third International Conference, AT 2015, Athens, Greece, December 17-18, 2015, Revised Selected Papers*, volume 9571 of *Lecture Notes in Computer Science*, pages 267–278. Springer, 2015. 49

[182] Harold Koster. Towards better implementation of the european union's anti-money laundering and countering the financing of terrorism framework. *Journal of Money Laundering Control*, 2020. 105

[183] Leonard Kostovetsky. Whom do you trust?: Investor-advisor relationships and mutual fund flows. *The Review of Financial Studies*, 29(4):898–936, 2016. 73

[184] Yeşem Kurt Peker, Xavier Rodriguez, James Ericsson, Suk Jin Lee, and Alfredo J Perez. A cost analysis of internet of things sensor data storage on blockchain via smart contracts. *Electronics*, 9(2):244, 2020. 78, 80, 88

[185] Anne Lafarre and Christoph Van der Elst. Blockchain technology for corporate governance and shareholder activism. *European Corporate Governance Institute (ECGI)-Law Working Paper*, (390), 2018. 99

[186] Marie-Christine Lagasquie-Schiex. *Inconsistency handling in non-monotonic reasoning: from consistency restoration to argumentation*. PhD thesis, Université Toulouse 3 Paul Sabatier, 2009. 10, 117

[187] Jérôme Lang, Leendert Van der Torre, and Emil Weydert. Utilitarian desires. *Autonomous agents and Multi-agent systems*, 5:329–363, 2002. 117

[188] Stefan Lauren, Francesco Belardinelli, and Francesca Toni. Aggregating bipolar opinions. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, pages 746–754, 2021. 71

[189] Tuomo Lehtonen, Johannes P Wallner, and Matti Järvisalo. Computing stable conclusions under the weakest-link principle in the aspic+ argumentation formalism. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning*, volume 19, pages 215–225, 2022. 31, 115

[190] João Leite and João G. Martins. Social abstract argumentation. In Toby Walsh, editor, *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pages 2287–2292. IJCAI/AAAI, 2011. 34, 49, 112, 116

[191] Robert Leshner and Geoffrey Hayes. Compound: The money market protocol. *White Paper*, 2019. 93

[192] Olga Lewandowska. *Post-Trade Processing of OTC Derivatives*. BoD–Books on Demand, 2020. 96

[193] Bo Li, Peng Qi, Bo Liu, Shuai Di, Jingen Liu, Jiquan Pei, Jinfeng Yi, and Bowen Zhou. Trustworthy ai: From principles to practices. *ACM Computing Surveys*, 55(9):1–46, 2023. 2

[194] Beishui Liao, Nir Oren, Leender van der Torre, and Serena Villata. Prioritized norms in formal argumentation. *Journal of Logic and Computation*, 29(2):215–240, 2019. 19, 29, 31, 32, 115

[195] Beishui Liao, Pere Pardo, Marija Slavkovik, Michal Feldman, and Leendert van der Torre. The jiminy advisor: moral agreements among stakeholders based on norms and argumentation. to be appeared. 13, 113, 115, 124, 126

[196] Vladimir Lifschitz. Computing circumscription. In *IJCAI*, volume 85, pages 121–127, 1985. 113

[197] Bowen Liu, Pawel Szalachowski, and Jianying Zhou. A first look into defi oracles. *arXiv preprint arXiv:2005.04377*, 2020. 93

[198] Fenrong Liu. *Reasoning about preference dynamics*, volume 354. Springer Science & Business Media, 2011. 113

[199] Paul Lorenzen and Kuno Lorenz. *Dialogische logik*. Wissenschaftliche Buchgesellschaft, 1978. 3, 112, 116

[200] Ronald Loui. Theory and computation of uncertain inference and decision. 1987. 4

[201] Tom Lyons, L Courcelas, and K Timsit. Blockchain and the gdpr. In *The European Union Blockchain Observatory and Forum*, 2018. 92

[202] Neil MacCormick and Robert S Summers. Interpretation and justification. *Interpreting Statutes: a comparative study*, pages 512–515, 1991. 119

[203] Batta Mahesh. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, 9:381–386, 2020. 2

[204] Khalid Mansour and Ryszard Kowalczyk. A meta-strategy for coordinating of one-to-many negotiation over multiple issues. In *Foundations of Intelligent Systems*, pages 343–353. Springer, 2011. 79

[205] Faraz Masood and Arman Rasool Faridi. An overview of distributed ledger technology and its applications. *Int. J. Comput. Sci. Eng*, 6(10):422–427, 2018. 4

[206] Paul-Amaury Matt, Maxime Morge, and Francesca Toni. Combining statistics and arguments to compute trust. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 209–216. Citeseer, 2010. 92

[207] Peter McBurney, David Hitchcock, and Simon Parsons. The eightfold way of deliberation dialogue. *22*, 1:95–132, 2007. 130

[208] Peter McBurney and Simon Parsons. Games that agents play: A formal framework for dialogues between autonomous agents. *Journal of logic, language and information*, 11:315–334, 2002. 124

[209] Peter McBurney and Simon Parsons. Games that agents play: A formal framework for dialogues between autonomous agents. *Journal of Logic, Language and Information*, 11(3):315–334, 2002. 128, 130

[210] Peter McBurney, Rogier M. Van Eijk, Simon Parsons, and Leila Amgoud. A dialogue game protocol for agent purchase negotiations. *Autonomous Agents and Multi-Agent Systems*, 7(3):235–273, 2003. 130

[211] John McCarthy. Applications of circumscription to formalizing common-sense knowledge. *Artificial intelligence*, 28(1):89–116, 1986. 113

[212] Robert N McCauley and Catherine R Schenk. Central bank swaps then and now: swaps and dollar liquidity in the 1960s. 2020. 107

[213] H McDowell. Asx to replace equity post-trade systems with blockchain. *Global Custodian https://www. globalcustodian. com/asx-to-replace-equity-post-trade-systems-with-blockchain/accessed*, 25, 2018. 100

[214] H McDowell. Canadian securities exchange releases blockchain clearing platform. *Global Custodian https://www. globalcustodian. com/canadian-securities-exchange-releases-blockchain-clearing-platform/accessed*, 25, 2018. 100

[215] Pascal Mehrwald, Theresa Treffers, Maximilian Titze, and Isabell Welpe. Blockchain technology application in the sharing economy: a proposed model of effects on trust and intermediation. 2019. 103

[216] Sanjay Modgil and Henry Prakken. A general account of argumentation with preferences. *Artif. Intell.*, 195:361–397, 2013. 15, 16, 18, 19, 31, 115

[217] Sanjay Modgil and Henry Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397, 2013. 111

[218] Sanjay Modgil and Henry Prakken. The aspic+ framework for structured argumentation: a tutorial. *Argument & Computation*, 5(1):31–62, 2014. 5, 6, 7, 9, 28, 29, 111, 112, 114, 118

[219] Sanjay Modgil and Henry Prakken. The *ASPIC*$^+$ framework for structured argumentation: a tutorial. *Argument Comput.*, 5(1):31–62, 2014. 15, 16, 17, 18, 19, 26, 31, 115

[220] RA Moore and C Derry. Efficacy of otc analgesics. *International Journal of Clinical Practice*, 67:21–25, 2013. 96

[221] Massimo Morini. Managing derivatives on a blockchain. a financial market professional implementation. *A Financial Market Professional Implementation (May 5, 2017)*, 2017. 74, 93

[222] Roger B Myerson. *Game theory: analysis of conflict*. Harvard university press, 1991. 11

[223] Amro Najjar, Xavier Serpaggi, Christophe Gravier, and Olivier Boissier. Multi-agent negotiation for user-centric elasticity management in the cloud. In *2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing*, pages 357–362. IEEE, 2013. 79

[224] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2008. 4, 76, 80, 82, 84

[225] Keerthi Nelaturu, John Adler, Marco Merlini, Ryan Berryhill, Neil Veira, Zissis Poulos, and Andreas Veneris. On public crowdsource-based mechanisms for a decentralized blockchain oracle. *IEEE Transactions on Engineering Management*, 67(4):1444–1458, 2020. 93

[226] Samer Nofal, Katie Atkinson, and Paul E Dunne. Algorithms for argumentation semantics: labeling attacks as a generalization of labeling arguments. *Journal of Artificial Intelligence Research*, 49:635–668, 2014. 85

[227] Farid Nouioua. Afs with necessities: further semantics and labelling characterization. In *International Conference on Scalable Uncertainty Management*, pages 120–133. Springer, 2013. 119

[228] Farid Nouioua and Vincent Risch. Bipolar argumentation frameworks with specialized supports. In *2010 22nd IEEE International Conference on Tools with Artificial Intelligence*, volume 1, pages 215–218. IEEE, 2010. 53, 118

[229] Farid Nouioua and Vincent Risch. Argumentation frameworks with necessities. In *International Conference on Scalable Uncertainty Management*, pages 163–176. Springer, 2011. 119

[230] Donald Nute. Defeasible reasoning: a philosophical analysis in prolog. *Aspects of artificial intelligence*, pages 251–288, 1988. 4

[231] Donald Nute. Defeasible logic. *Handbook of logic in artificial intelligence and logic programming*, 3:353–395, 1994. 4, 120

[232] Santiago Ontañón and Enric Plaza. Argumentation-based information exchange in prediction markets. In Iyad Rahwan and Pavlos Moraitis, editors, *Argumentation in Multi-Agent Systems, Fifth International Workshop, ArgMAS 2008, Estoril, Portugal, May 12, 2008. Revised Selected and Invited Papers*, volume 5384 of *Lecture Notes in Computer Science*, pages 181–196. Springer, 2008. 117

[233] Simona-Vasilica Oprea, Adela Bâra, and Anca Ioana Andreescu. Two novel blockchain-based market settlement mechanisms embedded into smart contracts for securely trading renewable energy. *IEEE Access*, 8:212548–212556, 2020. 91, 93

[234] Nir Oren, Michael Luck, and Chris Reed. Moving between argumentation frameworks. In *Proceedings of the 2010 International Conference on Computational Models of Argument*. IOS Press, 2010. 71, 72, 119

[235] Nir Oren and Timothy J Norman. Semantics for evidence-based argumentation. In *Proceedings of the 2008 conference on Computational Models of Argument: Proceedings of COMMA 2008*, pages 276–284, 2008. 54, 118

[236] Nir Oren, Timothy J Norman, and Alun Preece. Evidential reasoning in bipolar argumentation frameworks. In *Proceedings of the Fourth International Workshop on Argumentation in Multiagent Systems, Honolulu, Hawaii*. Citeseer, 2007. 120

[237] Mohamad Osmani, Ramzi El-Haddadeh, Nitham Hindi, Marijn Janssen, and Vishanth Weerakkody. Blockchain for next generation services in banking and finance: cost, benefit, risk and opportunity analysis. *Journal of Enterprise Information Management*, 34(3):884–899, 2021. 102

[238] Evin Özkan, Neda Azizi, and Omid Haass. Leveraging smart contract in project procurement through dlt to gain sustainable competitive advantages. *Sustainability*, 13(23):13380, 2021. 104

[239] Rudy Palmieri. Regaining trust through argumentation in the context of the current financial-economic crisis. *Studies in Communication Sciences*, 9(2):59–78, 2009. 92

[240] Wenxuan Pan and Meikang Qiu. Application of blockchain in asset-backed securitization. In *2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing,(HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, pages 71–76. IEEE, 2020. 93

[241] Alison R Panisson, Simon Parsons, Peter McBurney, Rafael H Bordini, et al. Choosing appropriate arguments from trustworthy sources. In *COMMA*, pages 345–352, 2018. 49

[242] Pere Pardo and Christian Straßer. Modular orders on defaults in formal argumentation. *Journal of Logic and Computation*, 2022. 19, 31, 115

[243] Pere Pardo, Leendert van der Torre, and Liuwen Yu. Advanced intelligent systems and reasoning: Standardization, experimentation, explanation. *Logics for New-Generation AI*, to be appear, 2023. 127

[244] Simon Parsons, Peter McBurney, and Elizabeth Sklar. Reasoning about trust using argumentation: A position paper. In *International Workshop on Argumentation in Multi-Agent Systems*, pages 159–170. Springer, 2010. 92

[245] Gareth W Peters and Efstathios Panayi. *Understanding modern banking ledgers through blockchain technologies: Future of transaction processing and smart contracts on the internet of money*. Springer, 2016. 99

[246] Gabriella Pigozzi. Belief merging and judgment aggregation. 2015. 117

[247] Andrea Pinna and Wiebe Ruttenberg. Distributed ledger technologies in securities post-trading revolution or evolution? *ECB Occasional Paper*, (172), 2016. 97

[248] José Plug. Complex argumentation in judicial decisions. analysing conflicting arguments. In *FAPR*, pages 464–479, 1996. 59

[249] José Plug. The analysis and evaluation of counter-arguments in judicial decisions. 1999. 10, 119, 121

[250] Sylwia Polberg. Intertranslatability of abstract argumentation frameworks. Technical report, 2017. 54, 65, 118

[251] Sylwia Polberg and Nir Oren. Revisiting support in abstract argumentation systems. In *Computational Models of Argument - Proceedings of COMMA*, volume 266 of *Frontiers in Artificial Intelligence and Applications*, pages 369–376. IOS Press, 2014. 72

[252] Sylwia Polberg and Nir Oren. Revisiting support in abstract argumentation systems. In *COMMA*, pages 369–376, 2014. 119

[253] John L Pollock. Defeasible reasoning. *Cognitive science*, 11(4):481–518, 1987. 4, 5, 26

[254] John L Pollock. How to reason defeasibly. *Artificial Intelligence*, 57(1):1–42, 1992. 15

[255] John L Pollock. Justification and defeat. *Artificial Intelligence*, 67(2):377–407, 1994. 15

[256] John L Pollock. *Cognitive carpentry: A blueprint for how to build a person*. Mit Press, 1995. 9, 15, 16, 112, 114

[257] John L. Pollock. Defeasible reasoning with variable degrees of justification. *Artif. Intell.*, 133(1-2):233–282, 2001. 15, 19, 21

[258] John L. Pollock. Defeasible reasoning and degrees of justification. *Argument Comput.*, 1(1):7–22, 2010. 15

[259] Claudia Pop, Cristian Pop, Antal Marcel, Andreea Vesa, Teodor Petrican, Tudor Cioara, Ionut Anghel, and Ioan Salomie. Decentralizing the stock exchange using blockchain an ethereum-based implementation of the bucharest stock exchange. In *2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP)*, pages 459–466. IEEE, 2018. 83

[260] Henry Prakken. Formal systems for persuasion dialogue. *The knowledge engineering review*, 21(2):163–188, 2006. 127

[261] Henry Prakken. An abstract framework for argumentation with structured arguments. *Argument & Computation*, 1(2):93–124, 2010. 54, 118

[262] Henry Prakken. *Logical tools for modelling legal argument: a study of defeasible reasoning in law*, volume 32. Springer Science & Business Media, 2013. 13

[263] Henry Prakken. On support relations in abstract argumentation as abstractions of inferential relations. In *ECAI 2014*, pages 735–740. IOS Press, 2014. 71, 118

[264] Henry Prakken. *Historical overview of formal argumentation*, volume 1. College Publications, 2018. 50, 71

[265] Henry Prakken and Giovanni Sartor. Law and logic: A review from an argumentation perspective. *Artificial intelligence*, 227:214–245, 2015. 109

[266] Henry Prakken and Gerard Vreeswijk. Logics for defeasible argumentation. *Handbook of philosophical logic*, pages 219–318, 2002. 28

[267] Randy Priem. Distributed ledger technology for securities clearing and settlement: benefits, risks, and regulatory implications. *Financial Innovation*, 6(1):1–25, 2020. 74, 91, 92, 93, 97, 99

[268] Randy Priem. A european distributed ledger technology pilot regime for market infrastructures: finding a balance between innovation, investor protection and financial stability. *Journal of Financial Regulation and Compliance*, 2022. 100

[269] Provable Things Limited. Provable, 2019. 93

[270] Dean G Pruitt. *Negotiation behavior*. Academic Press, 2013. 79

[271] Lisha Qiao, Yiqi Shen4 Liuwen Yu, Beishui Liao, and Leendert van der Torre. Arguing coalitions in abstract argumentation. *Logics for New-Generation AI*, page 93, 2021. 36

[272] Iyad Rahwan and Guillermo R Simari. *Argumentation in artificial intelligence*, volume 47. Springer, 2009. 92

[273] Sarvapali D Ramchurn, Dong Huynh, and Nicholas R Jennings. Trust in multi-agent systems. *The knowledge engineering review*, 19(1):1–25, 2004. 112, 116

[274] Raymond Reiter. A logic for default reasoning. *Artificial intelligence*, 13(1-2):81–132, 1980. 110

[275] Nicholas Rescher. Plausible reasoning: An introduction to the theory and practice of plausibilistic inference. *Philosophy and Rhetoric*, 13(3), 1980. 111

[276] Tjitze Rienstra, Alan Perotti, Serena Villata, Dov M Gabbay, and Leendert van der Torre. Multi-sorted argumentation. In *International Workshop on Theorie and Applications of Formal Argumentation*, pages 215–231. Springer, 2011. 49, 112, 116

[277] Tjitze Rienstra, Chiaki Sakama, and Leendert van der Torre. Persistence and monotony properties of argumentation semantics. In *International Workshop on Theory and Applications of Formal Argumentation*, pages 211–225. Springer, 2015. 64

[278] Tjitze Rienstra, Chiaki Sakama, Leendert van der Torre, and Beishui Liao. A principle-based robustness analysis of admissibility-based argumentation semantics. *Argument Comput.*, 11(3):305–339, 2020. 44, 50, 54

[279] Jussi Rintanen. Prioritized autoepistemic logic. In *Logics in Artificial Intelligence: European Workshop JELIA'94 York, UK, September 5–8, 1994 Proceedings 4*, pages 232–246. Springer, 1994. 113

[280] Jay R Ritter and Ivo Welch. A review of ipo activity, pricing, and allocations. *The journal of Finance*, 57(4):1795–1828, 2002. 96

[281] Peter Robinson and Raghavendra Ramesh. General purpose atomic crosschain transactions. In *2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pages 1–3. IEEE, 2021. 85

[282] Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, pages 97–109, 1982. 88

[283] Stuart J Russell. *Artificial intelligence a modern approach*. Pearson Education, Inc., 2010. 1

[284] Reade Ryan and Mayme Donohue. Securities on blockchain. *The Business Lawyer*, 73(1):85–108, 2017. 99

[285] Goldman Sachs. Blockchain: Putting theory in practice. *Equity Research https://github. com/bellaj/Blockchain/blob/master/Goldman-Sachs-Blockchain-Putting-Theory-into-Practice. pdf Accessed*, 30, 2017. 97, 101

[286] Iqbal H Sarker. Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, 2(3):160, 2021. 2

[287] Giovanni Sartor, Michał Araszkiewicz, Katie Atkinson, Floris Bex, Tom van Engers, Enrico Francesconi, Henry Prakken, Giovanni Sileno, Frank Schilder, Adam Wyner, et al. Thirty years of artificial intelligence and law: the second decade. *Artificial Intelligence and Law*, pages 1–37, 2022. 3

[288] Johannes Sedlmeir, Philipp Ross, André Luckow, Jannik Lockl, Daniel Miehle, and Gilbert Fridgen. The dlps: a new framework for benchmarking blockchains. 2021. 80, 81, 88

[289] OECD Blockchain Policy Series. The tokenisation of assets and potential implications for financial markets. *The Secretary General of the OECD*, 2020. 107

[290] Mr Ghiath Shabsigh, Mr Tanai Khiaonarong, and Mr Harry Leinonen. *Distributed ledger technology experiments in payments and settlements*. International Monetary Fund, 2020. 101

[291] Sugam Sharma, Samia Gamoura, Deva Prasad, and Arti Aneja. Emerging legal informatics towards legal innovation: Current status and future challenges and opportunities. *Legal Information Management*, 21(3-4):218–235, 2021. 3

[292] A Shobanadevi, Sumegh Tharewal, Mukesh Soni, D Dinesh Kumar, Ihtiram Raza Khan, and Pankaj Kumar. Novel identity management system using smart blockchain technology. *International Journal of System Assurance Engineering and Management*, pages 1–10, 2021. 102

[293] Mohammed Shuaib, Salwani Mohd Daud, Shadab Alam, and Wazir Zada Khan. Blockchain-based framework for secure and reliable land registry system. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 18(5):2560–2571, 2020. 103

[294] Avi Silberschatz, Michael Stonebraker, and Jeff Ullman. Database systems: Achievements and opportunities. *Communications of the ACM*, 34(10):110–120, 1991. 115

[295] Guillermo R Simari and Ronald P Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artificial intelligence*, 53(2-3):125–157, 1992. 4

[296] Amritraj Singh, Kelly Click, Reza M Parizi, Qi Zhang, Ali Dehghantanha, and Kim-Kwang Raymond Choo. Sidechain technologies in blockchain networks: An examination and state-of-the-art review. *Journal of Network and Computer Applications*, 149:102471, 2020. 83, 92

[297] Harjit Singh, Geetika Jain, Alka Munjal, and Sapna Rakesh. Blockchain technology in corporate governance: disrupting chain reaction or not? *Corporate Governance: The International Journal of Business in Society*, 20(1):67–86, 2019. 103

[298] Noam Slonim, Yonatan Bilu, Carlos Alzate, Roy Bar-Haim, Ben Bogin, Francesca Bonin, Leshem Choshen, Edo Cohen-Karlik, Lena Dankin, Lilach Edelstein, et al. An autonomous debating system. *Nature*, 591(7850):379–384, 2021. 109

[299] Reza Soltani, Marzia Zaman, Rohit Joshi, and Srinivas Sampalli. Distributed ledger technologies and their applications: A review. *Applied Sciences*, 12(15):7898, 2022. 4

[300] Ruben Stranders, Mathijs de Weerdt, and Cees Witteveen. Fuzzy argumentation for trust. In *International Workshop on Computational Logic in Multi-Agent Systems*, pages 214–230. Springer, 2007. 92

[301] Katia P Sycara. Multiagent systems. *AI magazine*, 19(2):79–79, 1998. 79

[302] Nick Szabo. Smart contracts: building blocks for digital markets. *EXTROPY: The Journal of Transhumanist Thought,(16)*, 18(2):28, 1996. 78

[303] Sek-Wah Tan and Judea Pearl. Specification and evaluation of preferences under uncertainty. In *Principles of Knowledge Representation and Reasoning*, pages 530–539. Elsevier, 1994. 117

[304] Yuqing Tang, Kai Cai, Elizabeth Sklar, Peter McBurney, and Simon Parsons. A system of argumentation for reasoning about trust. In *Proceedings of the 8th European Workshop on Multi-Agent Systems, Paris, France*, 2010. 92

[305] WT Luke Teacy, Jigar Patel, Nicholas R Jennings, and Michael Luck. Travos: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, 2006. 92

[306] Matthias Thimm. Strategic argumentation in multi-agent systems. *KI-Künstliche Intelligenz*, 28:159–168, 2014. 127

[307] Muthukkumarasamy Thuvarakan. Regulatory changes for redesigned securities markets with distributed ledger technology. *The Knowledge Engineering Review*, 35, 2020. 91

[308] Priscilla Toffano and Kathy Yuan. E-shekels across borders: a distributed ledger system to settle payments between israel and the west bank. 2019. 102

[309] Rijo Jackson Tom, Suresh Sankaranarayanan, and Joel JPC Rodrigues. Agent negotiation in an iot-fog based power distribution system for demand reduction. *Sustainable Energy Technologies and Assessments*, 38:100653, 2020. 79

[310] Francesca Toni. A tutorial on assumption-based argumentation. *Argument & Computation*, 5(1):89–117, 2014. 6, 111

[311] Stephen E Toulmin. *The uses of argument*. Cambridge university press, 1958. 3, 4

[312] Kentaroh Toyoda, Koji Machi, Yutaka Ohtake, and Allan N. Zhang. Function-level bottleneck analysis of private proof-of-authority ethereum blockchain. *IEEE Access*, 8:141611–141621, 2020. 84, 88, 89

[313] David Uzsoki. Tokenization of infrastructure. *A blockchain-based solution to financing sustainable infrastructure. IISD, MAVA*, 2019. 98

[314] Leendert van der Torre and Srdjan Vesic. The principle-based approach to abstract argumentation semantics. *FLAP*, 4(8), 2017. 10, 34, 39

[315] Leendert van der Torre and Srdjan Vesic. The principle-based approach to abstract argumentation semantics. In Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre, editors, *Handbook of formal argumentation, Volume 1*, pages 797–838. College Publictions, 2018. 54, 71

[316] Frans H. van Eemeren, Bart Garssen, Erik C. W. Krabbe, A. Francisca Snoeck Henkemans, Bart Verheij, and Jean H. M. Wagemans, editors. *Handbook of Argumentation Theory*. Springer, 2014. 4, 5, 110

[317] Frans H Van Eemeren, Rob Grootendorst, and Tjark Kruiger. Handbook of argumentation theory. In *Handbook of Argumentation Theory*. De Gruyter Mouton, 2019. 11, 110, 118

[318] Koen Vanderheyden and Tim Reucroft. Central securities depositories regulation: The next systemic crisis waiting to happen? *Journal of Securities Operations & Custody*, 7(3):242–252, 2015. 96

[319] Leendert ven der Torre, Livio Robaldo, , Giovanni Casini, and Serena Villata, editors. *Handbook of Legal AI*, volume 1. College Publications, 2022. 3, 121

[320] Bart Verheij. Deflog: on the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation*, 13(3):319–346, 2003. 120

[321] Serena Villata, Michal Araszkiewicz, Kevin Ashley, Trevor Bench-Capon, L Karl Branting, Jack G Conrad, and Adam Wyner. Thirty years of artificial intelligence and law: the third decade. *Artificial Intelligence and Law*, pages 1–31, 2022. 3

[322] Serena Villata, Guido Boella, Dov M Gabbay, and Leendert Van Der Torre. Arguing about the trustworthiness of the information sources. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 74–85. Springer, 2011. 92

[323] Rosa Vinciguerra, Francesca Cappellieri, and Michele Pizzo. A disruptive innovation of blockchain in finance and banking: A jump into the future. *Intellectual Capital, Smart Technologies and Digitalization: Emerging Issues and Opportunities*, pages 159–170, 2021. 100

[324] Eric Wall and Gustaf Malm. Using blockchain technology and smart contracts to create a distributed securities depository. 2016. 93

[325] Douglas Walton. On a razor's edge: Evaluating arguments from expert opinion. *Argument & computation*, 5(2-3):139–159, 2014. 90

[326] Ruth Wandhofer. *Technology innovation in financial markets: implications for money, payments and settlement finality*. PhD thesis, City, University of London, 2019. 101

[327] Michael P Wellman, Amy Greenwald, and Peter Stone. *Autonomous bidding agents: Strategies and lessons from the trading agent competition*. Mit Press, 2007. 79

[328] Sam M Werner, Daniel Perez, Lewis Gudgeon, Ariah Klages-Mundt, Dominik Harz, and William J Knottenbelt. Sok: Decentralized finance (defi). *arXiv preprint arXiv:2101.08778*, 2021. 74, 83, 93

[329] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1), 2016. 88

[330] Pascal Witzig and Victoriya Salomon. Cutting out the middleman: a case study of blockchain-induced reconfigurations in the swiss financial services industry. Technical report, Université de Neuchâtel, 2018. 99

[331] Michael Wooldridge. *An introduction to multiagent systems*. John wiley & sons, 2009. 79

[332] Christoph Wronka. Financial crime in the decentralized finance ecosystem: new challenges for compliance. *Journal of Financial Crime*, 30(1):97–113, 2023. 103

[333] Yining Wu and Mikołaj Podlaszewski. Implementing crash-resistance and non-interference in logic-based argumentation. *Journal of Logic and Computation*, 25(2):303–333, 2015. 5, 112

[334] Anthony P. Young, Sanjay Modgil, and Odinaldo Rodrigues. Prioritised default logic as rational argumentation. In Catholijn M. Jonker, Stacy Marsella, John Thangarajah, and Karl Tuyls, editors, *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 626–634. ACM, 2016. 15, 16, 18, 19, 21, 29, 31, 115

[335] Anthony P Young, Sanjay Modgil, and Odinaldo Rodrigues. Prioritised default logic as rational argumentation. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, pages 626–634, 2016. 111

[336] Anthony P. Young, Sanjay Modgil, and Odinaldo Rodrigues. On the interaction between logic and preference in structured argumentation. In Elizabeth Black, Sanjay Modgil, and Nir Oren, editors, *Theory and Applications of Formal Argumentation - 4th International Workshop*, volume 10757, pages 35–50. Springer, 2017. 19, 31, 115

[337] Liuwen Yu, Dongheng Chen, Lisha Qiao, Yiqi Shen, and Leendert van der Torre. A principle-based analysis of abstract agent argumentation semantics. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning*, volume 18, pages 629–639, 2021. 71

[338] Liuwen Yu and Leendert Van der Torre. A principle-based approach to bipolar argumentation. In *NMR 2020 Workshop Notes*, page 227, 2020. 50, 53, 68, 71

[339] Liuwen Yu, Mirko Zichichi, Réka Markovich, and Amro Najjar. Enhancing trust in trust services: Towards an intelligent human-input-based blockchain oracle (ihibo). In *Hawaii International Conference on System Sciences 2022 (HICSS-55)*, 2022. To appear. 74, 88

[340] Dirk A Zetzsche, Filippo Annunziata, Douglas W Arner, and Ross P Buckley. The markets in crypto-assets regulation (mica) and the eu digital finance strategy. *Capital Markets Law Journal*, 16(2):203–225, 2021. 107

[341] Dirk A Zetzsche, Ross P Buckley, Douglas W Arner, and Linus Fohr. The ico gold rush: It's a scam, it's a bubble, it's a super challenge for regulators. *Harv. Int'l LJ*, 60:267, 2019. 98

[342] Bo Zhang, Jun Zhu, and Hang Su. Toward the third generation artificial intelligence. *Science China Information Sciences*, 66(2):1–19, 2023. 1

[343] Yan Zhang and Norman Y Foo. Answer sets for prioritized logic programs. In *ILPS*, pages 69–83, 1997. 113

[344] Zibin Zheng, Shaoan Xie, Hong-Ning Dai, Weili Chen, Xiangping Chen, Jian Weng, and Muhammad Imran. An overview on smart contracts: Challenges, advances and platforms. *Future Generation Computer Systems*, 105:475–491, 2020. 78

[345] Huasheng Zhu and Zach Zhizhong Zhou. Analysis and outlook of applications of blockchain technology to equity crowdfunding in china. *Financial innovation*, 2(1):1–11, 2016. 99

[346] Mirko Zichichi. miker83z/IHiBO: Intelligent Human-input-based Blockchain Oracle. https://github.com/miker83z/IHiBO, August 2021. 88

[347] Mirko Zichichi, Stefano Ferretti, and Gabriele D'Angelo. Are Distributed Ledger Technologies Ready for Intelligent Transportation Systems? In *Proc. of the 3rd Workshop on Cryptocurrencies and Blockchains for Distributed Systems (CryBlock 2020), co-located with the 26th Annual International Conference on Mobile Computing and Networking (MobiCom 2020), ACM*, pages 1–6. ACM, 2020. 88

[348] Mirko Zichichi, Stefano Ferretti, and Gabriele D'Angelo. A framework based on distributed ledger technologies for data management and services in intelligent transportation systems. *IEEE Access*, 8:100384–100402, 2020. 78, 80

Chapter 10

Appendix

10.1 Proofs of The Principle-based Analysis of Bipolar Argumentation

Proposition 10.1. *TR satisfies P3.10-P3.15 for all the semantics.*

Proof. Semantics under TR does not concern the concept of agent, thus, they all satisfy P3.10-P3.15. □

Proposition 10.2. *TR does not satisfy P3.16 and P3.17*

Proof. We use a counterexample to prove TR does not satisfy P3.16 and 3.17, as shown in Figure 10.1. The complete, grounded, complete, stable semantics under TR is $\{a\}$.

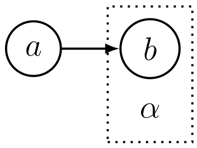


Figure 10.1: R does not satisfy P3.16 and 3.17

□

Proposition 10.3 (Sem_1, Sem_2 × P3.1 and P3.2)**.** *All the four kinds of Dung's semantics under admissibility*$_1$ *and/or admissibility*$_2$ *satisfy P3.1 and 3.2.*

Proof. Straightforward by definition. □

Proposition 10.4 (Sem_1 × P3.6)**.** *The grounded, complete and preferred semantics under admissibility*$_1$ *satisfy P3.6.*

Proof. Straightforward by definition. □

Proposition 10.5 (Sem_2 × P3.7)**.** *The grounded, complete and preferred semantics under admissibility*$_2$ *satisfy P3.7.*

Proof. Straightforward by definition. □

Proposition 10.6 (Sem_1 × P3.7)**.** *The grounded, complete and preferred semantics under admissibility*$_1$ *satisfy P3.7.*

Proof. It suffices to show that, given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, for every $E \subseteq \mathscr{A}$ and $c \in \mathscr{A}$, if E defend$_1$ c then E defend$_2$ c. This is straightforward by definition. □

Proposition 10.7 (Sem$_2$ × P3.6)**.** *None of the four kinds of Dung's semantics under admissibility$_2$ satisfy P3.6.*

Proof. Consider the AAF below:

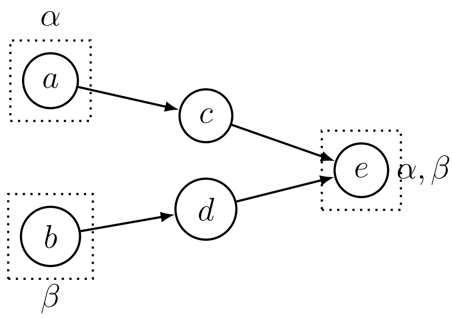


Figure 10.2: Agent defense

It is easy to see that, for every $E \subseteq \mathscr{A}$, if E is admissible$_1$ then $e \notin E$ (since otherwise E cannot defend$_1$ itself). On the other hand, it is easy to see that $\{a, b, e\}$ is the grounded extension of the above AAF under admissibility$_2$, thus the grounded semantics under admissibility$_2$ does not satisfy P3.6. Furthermore, since the grounded extension is the least complete extension, the complete semantics and preferred semantics under admissibility$_2$ does not satisfy P3.6 as well. The stable semantics under admissibility$_2$ is considered in Proposition 10.8. □

Proposition 10.8 (Sem$_1$, Sem$_2$ × P3.6, P3.7)**.** *The stable semantics under admissibility$_1$ and/or admissibility$_2$ satisfies neither P3.6 nor P3.7.*

Proof. Consider the AAF below: It is easy to see that $\{a, b, e\}$ is a stable extension of the

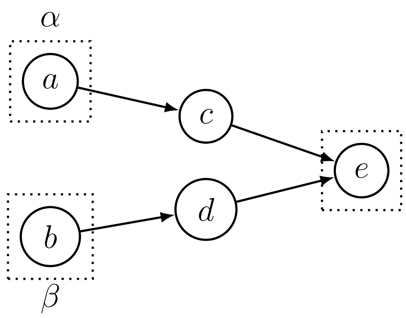


Figure 10.3: Agent defense

AAF. However $\{a, b, e\}$ is neither admissible$_1$ nor admissible$_2$. □

Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, let $\mathscr{U} \subseteq \mathscr{A}$. For every $E \subseteq \mathscr{U}$ and $c \in \mathscr{U}$, we use "E defend$_i^{\mathscr{A}}$ c" ("E defend$_i^{\mathscr{U}}$ c", respectively) to denote that E defend$_i$ c in AAF (in $AAF \downarrow_{\mathscr{U}}$, respectively), where $i \in \{1, 2\}$.

Lemma 10.1. *Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, let $\mathscr{U} \subseteq \mathscr{A}$ be an unattacked set. For every $i \in \{1, 2\}$, $E \subseteq \mathscr{A}$ and $c \in \mathscr{U}$, E* defend$_i^{\mathscr{A}}$ *c iff $E \cap \mathscr{U}$* defend$_i^{\mathscr{U}}$ *c.*

Proof. We only consider the case $i = 1$. From left to right: Assume E defend$_1^{\mathscr{A}} c$, then there must be $\alpha \in S_c$ such that for all arguments $b \in c^-$, there exists $a \in E \cap A_\alpha$ such that $a \in b^-$. Now let $b \in c^- \cap \mathscr{U}$ be arbitrary, it follows that there exists $a \in E \cap A_\alpha$ such that $a \in b^-$. Since $\mathscr{U}$ is unattacked and $b \in \mathscr{U}$, $a \in \mathscr{U}$ as well, i.e., $a \in (E \cap \mathscr{U}) \cap (A_\alpha \cap \mathscr{U})$. Since b is arbitrary, E defend$_1^{\mathscr{U}} c$.

From right to left: Assume $E \cap \mathscr{U}$ defend$_1^{\mathscr{U}} c$, then there must be $\alpha \in S_c$ such that for all arguments $b \in c^- \cap \mathscr{U}$, there exists $a \in E \cap (A_\alpha \cap \mathscr{U})$ such that $a \in b^-$. Since $\mathscr{U}$ is unattacked and $c \in \mathscr{U}$, $c^- \cap \mathscr{U} = c^-$. It follows that for all arguments $b \in c^-$, there exists $a \in E \cap A_\alpha$ such that $a \in b^-$. That is, E defend$_1^{\mathscr{A}} c$.

□

Proposition 10.9 (Sem$_1$, Sem$_2$× P3.3)**.** *The grounded, complete and preferred semantics under admissibility$_1$ and/or admissibility$_2$ satisfy P3.3, whereas the stable semantics under admissibility$_1$ and/or admissibility$_2$ does not satisfy P3.3.*

Proof. Since the stable semantics under admissibility$_1$ and/or admissibility$_2$ is the same as that in abstract argumentation frameworks, the second half follows from [35] directly.

For the first half, we first consider the complete semantics. Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$, let $\mathscr{U} \subseteq \mathscr{A}$ be an unattacked set. If $E \subseteq \mathscr{A}$ is a complete extension of AAF, we show that $E \cap \mathscr{U}$ is a complete extension of $AAF\downarrow_{\mathscr{U}}$. It is easy to see that $E \cap \mathscr{U}$ is conflict-free. For every $c \in E \cap \mathscr{U}$, since E defend$_i^{\mathscr{A}} c$, $E \cap \mathscr{U}$ defend$_i^{\mathscr{U}} c$ by Lemma 10.1. For every $c \in \mathscr{U} \backslash E$, since E does not defend$_i^{\mathscr{A}} c$, $E \cap \mathscr{U}$ does not defend$_i^{\mathscr{U}} c$ by Lemma 10.1.

On the other hand, if $E \subseteq \mathscr{U}$ is a complete extension of $AAF\downarrow_{\mathscr{U}}$, by Lemma 10.1 it is easy to see that E is still admissible$_i$ in AAF. For every $j \in \mathbb{N}$, we inductively define a set $E_j \subseteq \mathscr{A}$ as follows: (1) $E_0 = E$; (2) $E_{n+1} = E_n \cup \text{defend}_i^{\mathscr{A}}(E_n)$. Let $E^* = \bigcup_{j \in \mathbb{N}} E_j$. We show that $E^* \cup \mathscr{U} = E$ and E^* is a complete extension of AAF. For the former, using Lemma 10.1, we can show that $E_j \cap \mathscr{U} = E$ for every $i \in \mathbb{N}$: The case $j = 0$ is trivial. Suppose $E_n \cap \mathscr{U} = E$, then $E_{n+1} \cap \mathscr{U} = (E_n \cap \mathscr{U}) \cup (\text{defend}_i^{\mathscr{A}}(E_n) \cap \mathscr{U})$. By Lemma 10.1, $\text{defend}_i^{\mathscr{A}}(E_n) \cap \mathscr{U} = \text{defend}_i^{\mathscr{U}}(E_n \cap \mathscr{U}) = \text{defend}_i^{\mathscr{U}}(E) \subseteq E$ (note that E is a complete extension of $AAF\downarrow_{\mathscr{U}}$). Thus $E_{n+1} \cap \mathscr{U} = E$. Therefore $E^* \cup \mathscr{U} = E$. For the latter, since E is admissible$_i$ in AAF, it is easy to prove that E^* is a complete extension of AAF. This complete the proof for the case of the complete semantics.

Next we consider the grounded semantics. Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$, let $\mathscr{U} \subseteq \mathscr{A}$ be an unattacked set. Suppose that $E \subseteq \mathscr{A}$ is the grounded extension of AAF. $E \cap \mathscr{U}$ must be the grounded extension of $AAF\downarrow_{\mathscr{U}}$, since other there is $E' \subset E \cap U$ which is the grounded extension of $AAF\downarrow_{\mathscr{U}}$. Consider the set $(E')^*$ defined as in the last paragraph, it holds that $(E')^* \cup \mathscr{U} = E'$ and $(E')^*$ is a complete extension of AAF. It follows that $E \nsubseteq (E')^*$, contradicting that E is the grounded extension of AAF. On the other hand, suppose that $E \subseteq \mathscr{U}$ is the grounded extension of $AAF\downarrow_{\mathscr{U}}$. Consider the set E^* defined as in the last paragraph, it is a complete extension of AAF. Let $E' \subseteq \mathscr{A}$ be an arbitrary complete extension of AAF. We know that $E' \cap \mathscr{U}$ is a complete extension of $AAF\downarrow_{\mathscr{U}}$, thus $E \subseteq E' \cap \mathscr{U} \subseteq E'$. By induction on j, we can prove that $E_j \subseteq E'$ for every $j \in \mathbb{N}$. Thus $E^* \subseteq E'$. Since E' is arbitrary, E^* is the grounded extension of AAF.

Finally we consider the preferred semantics. Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$, let $\mathscr{U} \subseteq \mathscr{A}$ be an unattacked set. Suppose that $E \subseteq \mathscr{A}$ is a preferred extension of AAF. Since preferred extensions are maximal complete extensions, E is also a complete extension of AAF. Thus $E \cap \mathscr{U}$ is a complete extension of $AAF\downarrow_{\mathscr{U}}$. We show that $E \cap \mathscr{U}$ is a preferred extension of $AAF\downarrow_{\mathscr{U}}$. Suppose not, then there is $F \subseteq \mathscr{U}$ such that $E \cap \mathscr{U} \subset F$ and F is a complete extension of $AAF\downarrow_{\mathscr{U}}$. Consider the set $F \cup E$, we show that $F \cup E$ is admissible which will contradict that E is maximal admissible. We first show $F \cup E$ is conflict-free. Suppose not, since E and F are conflict-free and $\mathscr{U}$ is unattacked, the only possibility is that there are $a \in F\backslash E$ and $b \in E\backslash F$ such that a attacks b. Since E defend$_i^{\mathscr{A}}$ b, there must be $c \in E$ such that c attacks a. Note that $a \in F \subseteq \mathscr{U}$ and $\mathscr{U}$ is unattacked, then $c \in \mathscr{U} \cap E \subset F$. It implies that F is not conflict-free, contradiction! It remains to show $F \cup E$ can defend itself. This is trivial in view of Lemma 10.1.

On the other hand, suppose $E \subseteq \mathscr{U}$ is a preferred extension of $AAF\downarrow_{\mathscr{U}}$. Consider the set E^* defined as before. E^* is a complete extension of AAF and $E^* \cup \mathscr{U} = E$. Since $\mathscr{A}$ is finite, there must be a maximal complete extension (preferred extension) F of AAF such that $F \supseteq E^*$. Consider the set $F \cap \mathscr{U}$, we know that $F \cap \mathscr{U}$ is a preferred extension of $AAF\downarrow_{\mathscr{U}}$ by the last paragraph. Note that $E \subseteq F \cap \mathscr{U}$, thus $E = F \cap \mathscr{U}$. □

Proposition 10.10 (Sem$_1$, Sem$_2 \times$ P3.4)**.** *The grounded, complete and preferred semantics under admissibility$_1$ and/or admissibility$_2$ do not satisfy P3.4, whereas the stable semantics under admissibility$_1$ and/or admissibility$_2$ does satisfy P3.4.*

Proof. Since the stable semantics under admissibility$_1$ and/or admissibility$_2$ is the same as that in abstract argumentation frameworks, the second half follows from [36] directly. Specifically we can define the founction $\mathscr{G}$ in such a way that $\mathscr{G}(AAF, \mathscr{C}) = SE(TR(AAF), \mathscr{C})$ where the function SE is defined in page 184 of ([36]).

For the first half of the lemma, consider the $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ below: It is clear that,

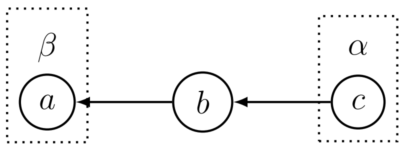


Figure 10.4: Agent Reinstatement

under both admissibility$_1$ and admissibility$_2$, $\{a\}$ is the grounded extension of the above AAF , as well as the complete extension and the preferred extension.

We first consider the grounded semantics. Suppose, towards to a contradiction, that there is a function $\mathscr{G}$ as described in P3.4. Thus $\mathscr{G}(AAF, \mathscr{A}) = \{\{a\}\}$ (1). On the other hand, let $E \in \mathscr{G}(AAF, \mathscr{A})$. There are three SCCs in AAF, i.e., $S_1 = \{a\}$, $S_2 = \{b\}$, and $S_3 = \{c\}$. By (1) we know that $E \cap S_1 = \{a\}$. Now consider the SSC S_2. Since $UP_{AAF}(S_2, E) = U_{AAF}(S_2, E) = \emptyset$, $E \cap S_2 = \emptyset$. S_3 remains to be considered. Note that $UP_{AAF}(S_3, E) = U_{AAF}(S_3, E) = \{c\}$. So $E \cap S_3 \in \mathscr{G}(AAF\downarrow_{\{c\}}, \{c\})$. Note that $\mathscr{G}(AAF\downarrow_{\{c\}}, \{c\}$ must be the grounded extension of $AAF\downarrow_{\{c\}}$ which is $\{c\}$. So $E = \{a, c\}$, contradicting (1).

The cases for the complete and preferred can be shown in a similar way. □

Proposition 10.11 (Sem$_1$, Sem$_2 \times$ P3.5)**.** *The grounded, complete and preferred semantics under admissibility$_1$ and/or admissibility$_2$ do not satisfy P3.5, whereas the stable semantics under admissibility$_1$ and/or admissibility$_2$ does satisfy P3.5.*

Proof. Since the stable semantics under admissibility$_1$ and/or admissibility$_2$ is the same as that in abstract argumentation frameworks, the second half follows from [36] directly.

For the first half of the lemma, consider the $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ in Figure 10.4.

It is clear that, under both admissibility$_1$ and admissibility$_2$, $\{a\}$ is the grounded extension of the above AAF, as well as the complete extension and the preferred extension.

Let us consider, for example, the complete semantics. $\{a\}$ is a complete extension of the AAF and $\{c\}$ is also a complete extension of $AAF^{\{a\}} = AAF \downarrow_{\{c\}}$. However, $\{a, c\}$ is not a complete extension of the AAF. □

Lemma 10.2. *Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, a set $E \subseteq A$ and a SCC $S \in SCCS_{AAF}$, for all $a \in S$, we have:*

1. *for each $i \in \{1, 2\}$: $a \in AUP^i_{AAF}(S, E)$ iff $\forall b \in E \backslash S$: $b \nrightarrow a$.*
2. *$a \in AU^1_{AAF}(S, E)$ iff $\forall b \in E \backslash S$: $b \nrightarrow a$ and $\exists \alpha \in \mathscr{S}_a$, $\forall b \in \mathscr{A} \backslash S$ with $b \rightarrow a$: $\exists c \in E \cap \mathscr{A}_\alpha$ such that $c \rightarrow b$.*
3. *$a \in AU^2_{AAF}(S, E)$ iff $\forall b \in E \backslash S$: $b \nrightarrow a$ and $\forall b \in \mathscr{A} \backslash S$ with $b \rightarrow a$: $\exists \alpha \in \mathscr{S}_a$ & $c \in E \cap \mathscr{A}_\alpha$ such that $c \rightarrow b$.*

Proposition 10.12 (Sem$_1$ × P3.9)**.** *The complete, preferred and grounded semantics under admissibility$_1$ do not satisfy P3.9.*

Proof. Consider the $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ in Figure 10.2: It is clear that, under both admissibility$_1$, $\{a, b\}$ is the grounded extension of the above AAF, as well as the complete extension and the preferred extension.

We first consider the grounded semantics. Suppose, towards to a contradiction, that there is a function $\mathscr{G}$ as described in P3.9. Thus $\mathscr{G}(AAF, \mathscr{A}) = \{\{a, b\}\}$ (1). On the other hand, let $E \in \mathscr{G}(AAF, \mathscr{A})$. There are five SCCs in AAF, i.e., $S_1 = \{a\}$, $S_2 = \{b\}$, $S_3 = \{c\}$, $S_4 = \{d\}$ and $S_5 = \{e\}$. By (1) we know that $E \cap S_1 = \{a\}$ and $E \cap S_2 = \{b\}$. Now consider the SSC S_3. Since $\text{AUP}^2_{AAF}(S_3, E) = \text{AU}^2_{AAF}(S_3, E) = \varnothing$, $E \cap S_3 = \varnothing$. By the same reason, $E \cap S_4 = \varnothing$. S_5 remains to be considered. Note that $\text{AUP}^2_{AAF}(S_5, E) = \text{AU}^2_{AAF}(S_5, E) = \{c\}$. So $E \cap S_5 \in \mathscr{G}(AAF \downarrow_{\{c\}}, \{c\})$. Note that $\mathscr{G}(AAF \downarrow_{\{c\}}, \{c\}$ must be the grounded extension of $AAF \downarrow_{\{c\}}$ (under admissibility$_1$) which is $\{c\}$. So $E = \{a, b, c\}$, contradicting (1).

The cases for the complete and preferred can be shown in a similar way. □

Proposition 10.13. *The stable semantics under admissibility$_1$ and/or admissibility$_2$ satisfy P3.8 and P3.9.*

Proof. We define the function $\mathscr{G}$ as follows: for any $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ and $C \subseteq \mathscr{A}$, $E \in \mathscr{G}(AAF, C)$ if and only E is a stable extension in $\langle \mathscr{A}, \rightarrow \rangle$. By Proposition 32 in [36], it is straightforward to show that $\mathscr{G}$ satisfies the propertied in P3.8 and P3.9. □

Proposition 10.14 (Sem$_1$ × P3.8)**.** *The grounded, complete and preferred semantics under admissibility$_1$ does not satisfy P3.8.*

Proof. For the complete and preferred semantics, consider the following $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, We first consider the complete semantics. Suppose, toward to a contradiction, that there is function $\mathscr{G}$ as described in P3.8. Let $E \in \mathscr{G}(AAF, \mathscr{A})$. There are three SCCs: $S_1 = \{a\}$, $S_2 = \{b\}$ and $S_3 = \{c, d, e, f\}$. For S_1, since $\text{AUP}^1_{AAF}(S_1, E) = \{a\}$, $AAF \downarrow_{\text{AUP}^1_{AAF}(S_1, E)}$ consists of the single point a. The complete extension of $AAF \downarrow_{\text{AUP}^1_{AAF}(S_1, E)}$ under admissibility$_1$ is $\{a\}$. Since $\text{AUP}^1_{AAF}(S_1, E) = \text{AU}^1_{AAF}(S_1, E) = \{a\}$, by P3.8, $(E \cap S_1) = \mathscr{G}(AAF \downarrow_{\{a\}}, \{a\}) = \{a\}$. For S_2, since $\text{AUP}^1_{AAF}(S_2, E) = \text{AU}^1_{AAF}(S_2, E) = \varnothing$, $(E \cap S_2) = \varnothing$. For S_3,

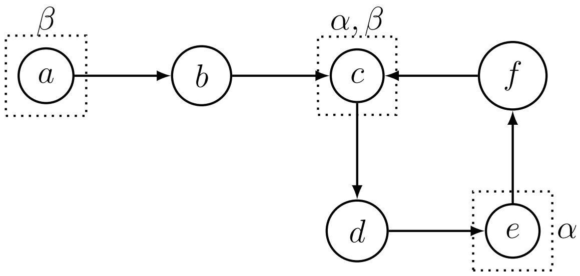


Figure 10.5: Agent Reinstatement

we first note that $\mathrm{AUP}^1_{AAF}(S_3, E) = \mathrm{AU}^1_{AAF}(S_3, E) = \{c, d, e, f\}$. By P3.8, $\mathscr{G}(AAF\downarrow_{\{c,d,e,f\}}, \{c, d, e, f\})$ must be the set of complete extensions of $AAF\downarrow_{\{c,d,e,f\}}$. Since $\{c, e\}$ is the only complete extension of $AAF\downarrow_{\{c,d,e,f\}}$, we have $E \cap S_3 = \mathscr{G}(AAF\downarrow_{\{c,d,e,f\}}, \{c, d, e, f\}) = \{c, e\}$. In sum, $E = \{a, c, e\}$. However, E cannot defend$_1$ itself in AAF (consider the argument c), contradiction!

The case for preferred semantics can be shown similarly.

For the grounded semantics, consider the following $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, Suppose, to-

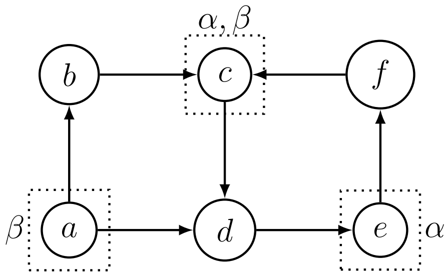


Figure 10.6: Agent Reinstatement

wards to a contradiction, that there is function $\mathscr{G}$ as described in P3.8. Let $E \in \mathscr{G}(AAF, \mathscr{A})$. There are three SCCs: $S_1 = \{a\}$, $S_2 = \{b\}$ and $S_3 = \{c, d, e, f\}$. For S_1, since $\mathrm{AUP}^1_{AAF}(S_1, E) = \{a\}$, $AAF\downarrow_{\mathrm{AUP}^1_{AAF}(S_1,E)}$ consists of the single point a. The grounded extension of $AAF\downarrow_{\{a\}}$ under admissibility$_1$ is $\{a\}$. Since $\mathrm{AUP}^1_{AAF}(S_1, E) = \mathrm{AU}^1_{AAF}(S_1, E) = \{a\}$, by P3.8, $(E \cap S_1) = \mathscr{G}(AAF\downarrow_{\{a\}}, \{a\}) = \{a\}$. For S_2, since $\mathrm{AUP}^1_{AAF}(S_2, E) = \mathrm{AU}^1_{AAF}(S_2, E) = \varnothing$, $(E \cap S_2) = \varnothing$. For S_3, we first note that $\mathrm{AUP}^1_{AAF}(S_3, E) = \mathrm{AU}^1_{AAF}(S_3, E) = \{c, e, f\}$. By principle 8, $\mathscr{G}(AAF\downarrow_{\{c,e,f\}}, \{c, e, f\})$ must be the grounded extension of $AAF\downarrow_{\{c,e,f\}}$. Since the grounded extension of $AAF\downarrow_{\{c,e,f\}}$ is $\{c, e\}$, we have $E \cap S_3 = \mathscr{G}(AAF\downarrow_{\{c,e,f\}}, \{c, e, f\}) = \{c, e\}$. In sum, $E = \{a, c, e\}$. However, E cannot defend$_1$ itself in AAF (consider the argument c), contradiction! □

Proposition 10.15 (Sem$_2$ × P3.8)**.** *The complete, grounded and preferred semantics for admissibility$_2$ do not satisfy P3.8.*

Proof. Consider the following $AAF_1 = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset_1 \rangle$ (on the left) and $AAF_2 = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset_2 \rangle$ (on the right):

Let us consider, for example, the complete semantics under admissibility$_2$. We denote the five SCCs in AAF_1 (or AAF_2) as: $S_1 = \{a\}$, $S_2 = \{b\}$, $S_3 = \{c\}$, $S_4 = \{d\}$ and $S_5 = \{e\}$. It is clear that, under admissibility$_2$, $E = \{a, b, e\}$ is a complete extension of AAF_1. Thus $\forall i \in \{1, \ldots, 5\}$:

$$E \cap S_i \in \mathscr{G}(AAF_1\downarrow_{\mathrm{AUP}^1_{AAF_1}(S_i,E)}, \mathrm{AU}^1_{AAF_1}(S_i, E)).$$

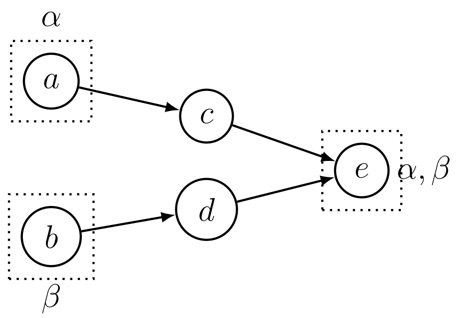


Figure 10.7: Sem_1 does not satisfy P3.15

Henceforth, we can also verify that $\forall i \in \{1, \ldots, 5\}$:

$$E \cap S_i \in \mathscr{G}(AAF_2 \downarrow_{\text{AUP}^1_{AAF_2}(S_i,E)}, \text{AU}^1_{AAF_2}(S_i, E)).$$

By P3.8, it must be that $E \in \mathscr{G}(AAF_2, \mathscr{A})$. But E is not a complete extension of AAF_2 under admissibility$_2$, contradiction! □

Proposition 10.16 ($\text{Sem}_2 \times$ P3.9)**.** *The complete, grounded and preferred semantics for admissibility$_2$ do not satisfy P3.9.*

Proof. For the grounded semantics, consider the $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ in Figure 10.6. Suppose, towards to a contradiction, that there is function $\mathscr{G}$ as described in P3.9. Let $E \in \mathscr{G}(AAF, \mathscr{A})$. There are three SCCs: $S_1 = \{a\}$, $S_2 = \{b\}$ and $S_3 = \{c, d, e, f\}$. For S_1, since $\text{AUP}^2_{AAF}(S_1, E) = \{a\}$, $AAF \downarrow_{\text{AUP}^2_{AAF}(S_1,E)}$ consists of the single point a. The grounded extension of $AAF \downarrow_{\{a\}}$ under admissibility$_2$ is $\{a\}$. Since $\text{AUP}^2_{AAF}(S_1, E) = \text{AU}^2_{AAF}(S_1, E) = \{a\}$, by P3.8, $(E \cap S_1) = \mathscr{G}(AAF \downarrow_{\{a\}}, \{a\}) = \{a\}$. For S_2, since $\text{AUP}^2_{AAF}(S_2, E) = \text{AU}^2_{AAF}(S_2, E) = \varnothing$, $(E \cap S_2) = \varnothing$. For S_3, we first note that $\text{AUP}^2_{AAF}(S_3, E) = \text{AU}^2_{AAF}(S_3, E) = \{c, e, f\}$. By principle 8, $\mathscr{G}(AAF \downarrow_{\{c,e,f\}}, \{c, e, f\})$ must be the grounded extension of $AAF \downarrow_{\{c,e,f\}}$. Since the grounded extension of $AAF \downarrow_{\{c,e,f\}}$ (under admissibility$_2$) is $\{c, e\}$, we have $E \cap S_3 = \mathscr{G}(AAF \downarrow_{\{c,e,f\}}, \{c, e, f\}) = \{c, e\}$. In sum, $E = \{a, c, e\}$. However, E cannot defend$_2$ itself in AAF (consider the argument d), contradiction!

The similar arguments hold for the complete and preferred semantics. □

Proposition 10.17. *Sem_1 does not satisfy P3.13 for all the semantics.*

Proof. Figure 3.2 is a counterexample to prove Sem_1 does not satisfy Principle3.13. □

Proposition 10.18. *Sem_1 does not satisfy P3.15 for all the semantics.*

Proof. We use a counterexample to prove, as shown in Figure 10.1.

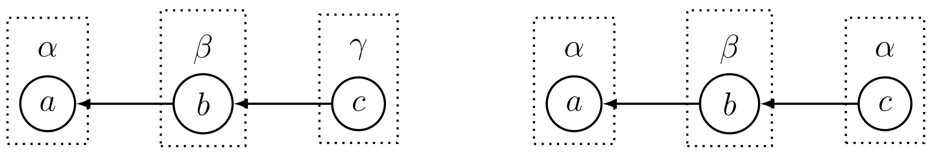


Figure 10.8: Sem_1 does not satisfy P3.15

□

Proposition 10.19. *Sem_1 does not satisfy P3.16 for all the semantics.*

Proof. A counterexample is Figure 3.2. □

Proposition 10.20. *Sem_1 does not satisfy P3.16 for all the semantics.*

Proof. A counterexample is 10.1 □

Proposition 10.21.

Proposition 10.22 (SR_1 × P3.1, P3.2)**.** *All the four kinds of Dung's semantics for SR_1 do not satisfy P3.1. Thus, they do not satisfy P3.2, as well.*

Proof. A counter example is Figure 10.1. □

Proposition 10.24 (SR_3 × P3.1)**.** *All the four kinds of Dung's semantics for SR_3 satisfy P3.1.*

Proof. Let an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given. It suffices to show that for any $a, b \in \mathscr{A}$, if there is no attack between them in $SR_3(AAF)$, then there is not attack between them in AAF. Suppose not, without loss of generality, we assume that a attacks b in AAF. If b does not attack a in AAF, then a attacks b in $SR_3(AAF)$, contradiction! If b attacks a in AAF, since it must be either $|\mathscr{S}(b)| \not> |\mathscr{S}(a)|$ or $|\mathscr{S}(a)| \not> |\mathscr{S}(b)|$, there must be attack between them in $SR_3(AAF)$, contradiction! □

Proposition 10.26 (SR_2 × P3.2)**.** *All the four kinds of Dung's semantics for SR_2 do not satisfy P3.2.*

Proof. A counter example is Figure 10.1. □

Proposition 10.27 (SR_3 × P3.2)**.** *All the four kinds of Dung's semantics for SR_3 satisfy P3.2.*

Proof. Let an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given. It suffices to show that any admissible set $E \subseteq \mathscr{A}$ in $SR_3(AAF)$ is admissible in $\langle \mathscr{A}, \rightarrow \rangle$. Suppose, toward to a contradiction, that E is not admissible in $\langle \mathscr{A}, \rightarrow \rangle$. By Proposition 10.24, it must be that there is $c \in E$ such that E cannot defend c in $\langle \mathscr{A}, \rightarrow \rangle$. Thus, since E is conflict-free in $\langle \mathscr{A}, \rightarrow \rangle$, there must be $b \in \mathscr{A} \backslash \mathscr{E}$ such that b attacks c and a does not attack b for all $a \in E$ (in $\langle \mathscr{A}, \rightarrow \rangle$). Thus, specially, c does not attack b (in $\langle \mathscr{A}, \rightarrow \rangle$). Therefore, by definition, b also attacks c in $SR_3(AAF)$. But for all $a \in E$, a does not attack b in $SR_3(AAF)$. So E cannot defend c in $SR_3(AAF)$, contradiction! □

Proposition 10.28 (SR_4 × P3.2)**.** *The complete semantics, preferred semantics and stable semantics for SR_4 do not satisfy P3.2, whereas the grounded semantics for SR_4 do satisfy P3.2.*

Proof. For the first half, a counter example is as follows: Note that $\{b\}$ is one of complete extensions, preferred extensions and stable extensions of $SR_4(AAF)$, but it is not admissible in AAF.

For the second half of the lemma, let an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and $E \subseteq \mathscr{A}$ be the grounded extension of $SR_4(AAF)$. Since E is conflict-free, we only need to show E can defend all of its members in AAF. Let D be the characteristic function of $SR_4(AAF)$,

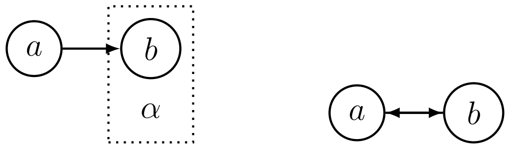


Figure 10.9: SR_4 does not satisfy P3.2

thus $E = \bigcup_{i=1,\ldots,\infty} D^i(\varnothing)$. We prove, by induction on the value of i, that E can defend $D^i(\varnothing)$ in AAF for $i = 1, \ldots, \infty$. If $i = 1$, let $c \in D^1(\varnothing) = D(\varnothing)$ be arbitrary. For any argument b such that b attacks c and c does not attack a in AAF, it must be that b attacks c in $SR_4(AAF)$ by definition. Thus such b does not exist, so E can defend $D^1(\varnothing)$ in AAF. Now suppose that E can defend $D^n(\varnothing)$ in AAF. Let $c \in D^{n+1}(\varnothing)$ be arbitrary. For any argument b such that b attacks c and c does not attack a in AAF, it must be that b attacks c in $SR_4(AAF)$ by definition. Since $c \in D^{n+1}(\varnothing) = D(D^n(\varnothing))$, there must be $a \in D^n(\varnothing)$ such that a attacks b in $SR_4(AAF)$. If a does not attacks b in AAF, it must hold that b attacks a in AAF by definition. By IH, there is $a' \in E$ such that a' attacks b in AAF, as desired. □

Proposition 10.29 (SR_1, SR_3 × P3.3)**.** *The grounded, complete, preferred semantics for SR_1 and SR_3 satisfy P3.3, whereas the stable semantics for SR_1 and SR_3 do not satisfy P3.3.*

Proof. For any $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, if a set of arguments $\mathscr{U} \subseteq \mathscr{A}$ is unattacked in AAF, then it is also unattacked in $SR_1(AAF)$ and $SR_3(AAF)$. Thus the first half follows from [35] directly.

For the second half of the lemma, consider the following $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$: It is

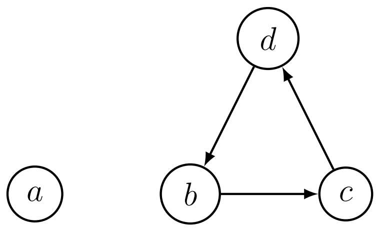


Figure 10.10: Stable semantics for SR_1 and SR_3 do not satisfy P3.3

easy to see that $SR_1(AAF) = SR_3(AAF) = \langle \mathscr{A}, \rightarrow \rangle$. $\{a\}$ is unattacked and is a stable extension of $AAF\downarrow_{\{a\}}$ and $AAF\downarrow_{\{a\}}$. But there is no stable extension E of AAF such that $a \in E$. □

Proposition 10.30 (SR_2 × P3.3)**.** *All the four kinds of Dung's semantics for SR_2 do not satisfy P3.3.*

Proof. A counter example is Figure 10.11. □

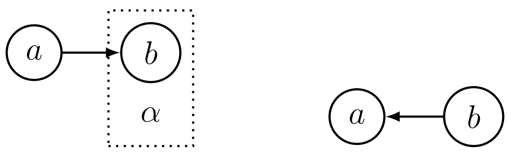


Figure 10.11: SR_4 does not satisfy P3.3

Proposition 10.31 (SR$_4$ × P3.3)**.** *All the four kinds of Dung's semantics for SR$_4$ do not satisfy P3.3.*

Proof. A counter example AAF is Figure 10.9. It is easy to see that $\{a\}$ is unattacked and is the grounded extension, complete extension, preferred extension and stable extension of $SR_4(AAF\downarrow_{\{a\}})$. But the grounded extension of $SR_4(AAF)$ is $\varnothing$ and $\{b\}$ is a complete, preferred and stable extension of $SR_4(AAF)$. □

Proposition 10.32 (SR$_1$, SR$_2$, SR$_4$ × P3.4)**.** *All the four kinds of Dung's semantics for SR$_1$, SR$_2$, and SR$_4$ do not satisfy P3.4.*

Proof. We first consider the case for SR$_1$, consider the following $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$:

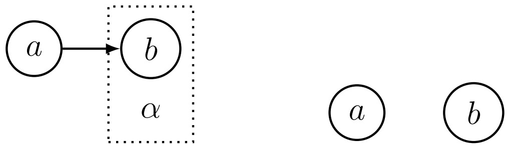


Figure 10.12: SR$_1$, SR$_2$, SR$_4$ does not satisfy P3.4

It is easy to see that $\{a,b\}$ is the grounded extension of $SR_1(AAF)$, as well as a complete extension, a preferred extension and a stable extension. Let consider, for example, the grounded semantics for SR$_1$: Suppose, toward to a contradiction, that there is a function $\mathscr{G}$ as described in P3.4. Then $\{a,b\} \in \mathscr{G}(AAF, \mathscr{A})$ (1). On the other hand, let $E \in \mathscr{G}(AAF, \mathscr{A})$. There are two SCCs in AAF, i.e., $S_1 = \{a\}$ and $S_2 = \{b\}$. Since S_1 is the initial SCC, $UP_{AAF}(S_1, E) = U_{AAF}(S_1, E) = \{a\}$ for any E. Thus $E \cap S_1 \in \mathscr{G}(AAF\downarrow_{\{a\}}, \{a\})$. Since the grounded (complete/ preferred/ stable) extension of $SR_1(AAF\downarrow_{\{a\}})$ is $\{a\}$, $E \cap S_1 = \{a\}$. For S_2, $UP_{AAF}(S_2, E) = U_{AAF}(S_2, E) = \varnothing$, thus $E \cap S_2 = \varnothing$, contradicting (1). The cases for the complete, preferred, and stable semantics for SR$_1$ can be shown similarly.

We then consider the case for SR$_2$. It is easy to see that the grounded extension of $SR_2(AAF)$ is $\{b\}$, as well as a complete extension, a preferred extension and a stable extension. But, by the same reasoning as above, $E \cap S_1 = \{a\}$ for any function $\mathscr{G}$ and $E \in \mathscr{G}(AAF, \mathscr{A})$.

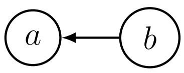


Figure 10.13: SR$_2$ does not satisfy P3.4

The case for SR$_3$ remains to be considered. For the grounded semantics, note that $\varnothing$ is the grounded extension of $SR_4(AAF)$. But, by the same reasoning as above, $E \cap S_1 = \{a\}$ for any function $\mathscr{G}$ and $E \in \mathscr{G}(AAF, \mathscr{A})$. For the complete, preferred, and stable semantics, note that $\{b\}$ is a complete, preferred and stable extension of $SR_4(AAF)$. But, by the same reasoning as above, $E \cap S_2 = \varnothing$ for any function $\mathscr{G}$ and $E \in \mathscr{G}(AAF, \mathscr{A})$.

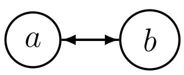


Figure 10.14: SR$_4$ does not satisfy P3.4

□

Proposition 10.33 (SR$_1$ × P3.5). *The grounded, complete, and preferred semantics for SR$_1$ do not satisfy the P3.5, whereas the stable semantics for SR$_1$ do satisfy the P3.5.*

Proof. Consider the following $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$:

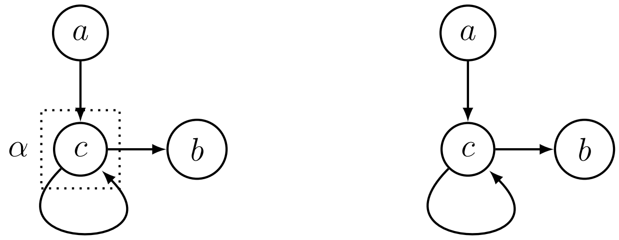


Figure 10.15: SR$_1$ does not satisfy P3.5

It is easy to see that $\{a\}$ is the grounded extension of $SR_1(AAF)$, as well a complete extension and a preferred extension. Let $E = \{a\}$, then then AAF^E consists of the single point b. It is also easy to see that $\{b\}$ is the grounded, complete and preferred extension of AAF^E. However, $\{a, b\}$ is not admissible in $SR_2(AAF)$.

For the second half of the proposition, let $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and E be a stable extension of $SR_1(AAF)$. It is easy to know that E^* is empty, thus E' is also empty. □

Proposition 10.34 (SR$_2$ × P3.5). *All the four kinds of Dung's semantics for SR$_2$ does not satisfy P3.5.*

Proof. Consider the counter example in Figure 10.11. It is easy to see that $\{b\}$ is the grounded, complete, preferred and stable extension of $SR_2(AAF)$. Let $E = \{b\}$. Because AAF^E consists of the single point a, thus a is the grounded, complete, preferred and stable extension of $SR_2(AAF^E)$. However, $\{a, b\}$ is not consistent in $SR_2(AAF)$. □

Proposition 10.35 (SR$_3$ × P3.5). *All the four kinds of Dung's semantics for SR$_3$ satisfy P3.5.*

Proof. We first consider *the complete semantics*: let $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and E be a complete extension of $SR_3(AAF)$ and E' be a complete extension of $SR_3(AAF^E)$. We show that $E \cup E'$ is a complete extension of $SR_3(AAF)$.

We first show that $E \cup E'$ is conflict-free in $SR_3(AAF)$. Suppose not, then the only possibility is that there are $a \in E$ and $b \in E'$ such that a attacks b or b attacks a in $SR_3(AAF)$. The former case is impossible, because it follows that a attacks b in AAF and, thus, $b \in E^+$. In the latter case, since E defends a in $SR_3(AAF)$, there must be $c \in E$ such that c attacks b in $SR_3(AAF)$. It follows that c attacks b in AAF. Thus $b \in E^+$, contradiction!

We then show that $E \cup E'$ can defend itself in $SR_3(AAF)$. It is obvious that, for every $c \in E$, $E \cup E'$ defends c in $SR_3(AAF)$. Let $c \in E'$ be arbitrary and b attacks c in $SR_3(AAF)$. Since $E \cup E'$ is consistent in $E \cup E'$, $b \in AAF^E$ or $b \in E^+$. In the former case, there is $a \in E'$ such that a attacks b in $E \cup E'$ (since E' is a complete extension of $SR_3(AAF^E)$). In the latter case, there is $a \in E$ such that a attacks b in AAF. Thus it must be either a attacks b or b attacks a in $SR_3(AAF)$. If it is the case that b attacks a in $SR_3(AAF)$, since E defend a in $SR_3(AAF)$, there must be $a' \in E$ such that a' attacks b in $SR_3(AAF)$.

Finally we show that $E \cup E'$ does not defend c for any $c \in \mathscr{A} \backslash (E \cup E')$. Suppose not, it follows that $c \notin E^+$ (since, otherwise, there is $b \in E$ such that b attacks c in AAF. Therefore it must be either b attacks c or c attacks b in $SR_3(AAF)$. In either case, it follows that there is $b' \in E$ such that b' attacks c in $SR_3(AAF)$. Since we assume that $E \cup E'$ defends c, there is $a \in E \cup E'$ such that a attacks b' in $SR_3(AAF)$, contradicting that $E \cup E'$ is conflict-free). For any argument $b \in E^*$ such that b attacks c in $SR_3(AAF^E)$ (thus, in $SR_3(AAF)$), there is no $a \in E$ such that a attacks b in $SR_3(AAF)$ since, otherwise, it implies that $b \in E^+$. Thus, since we assume that $E \cup E'$ defends c in $SR_3(AAF)$, there must be $a \in E'$ such that a attacks b in $SR_3(AAF)$ (thus, in $SR_3(AAF^E)$). It follows that E' defends c in $SR_3(AAF^E)$. Since E' is a complete extension of $SR_3(AAF^E)$, $c \in E'$, contradiction!

We then consider *the grounded semantics*: Let $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and E be the grounded extension of $SR_3(AAF)$ and E' be the grounded extension of $SR_3(AAF^E)$. It suffices to show that $E' = \varnothing$. We show the stronger claim that there is no argument c in $SR_3(AAF^E)$ that receives no attack (in $SR_3(AAF^E)$). Suppose not, then, for any $b \in \mathscr{A}$ with b attacks c in $SR_3(AAF)$, $b \in E \cup E^+$. If $b \in E$, it follows that $c \in E^+$, contradicting that c is in $SR_3(AAF^E)$. If $b \in E^+$, then there is $a \in E$ such that a attacks b in $SR_3(AAF)$. So E defends c in $SR_3(AAF)$. Thus $c \in E$ since E is the grounded extension of $SR_3(AAF)$, contradiction!

For *the preferred semantics*, if E is a preferred extension of $SR_3(AAF)$ and E' is a preferred extension of $SR_3(AAF^E)$, then we can show that $E' = \varnothing$. Since, otherwise, $E \cup E'$ is a complete extension of $SR_3(AAF)$ by our previous result and $E \cup E' \supset E$.

Finally, for the stable semantics, suppose E is a stable extension of $SR_3(AAF)$ and E' is a stable extension of $SR_3(AAF^E)$. It is easy to see that $E' = \varnothing$ because $E^+ = \mathscr{A} \backslash E$.

□

Proposition 10.36 (SR$_4$ × P3.5)**.** *The complete, preferred and stable semantics for SR$_4$ does not satisfy P3.5, whereas the grounded semantics for SR$_4$ do satisfy P3.5.*

Proof. For the first half, consider the counter example Figure 10.9.

It is easy to see that $\{b\}$ is a complete, preferred and stable extension of $SR_4(AAF)$. Let $E = \{b\}$, $SR_4(AAF^E)$ consists of the single point a. Thus $\{a\}$ is the complete, preferred and stable extension of $SR_4(AAF^E)$. However $\{b, a\}$ is not conflict-free in $SR_4(AAF)$.

For the second half, let $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and E be the grounded extension of $SR_4(AAF)$ and E' be the grounded extension of $SR_4(AAF^E)$. It suffices to show that $E' = \varnothing$. We show the stronger claim that there is no argument c in $SR_4(AAF^E)$ that receives no attack (in $SR_4(AAF^E)$). Suppose not, then, for any $b \in \mathscr{A}$ with b attacks c in $SR_4(AAF)$, $b \in E \cup E^+$. We show that $b \notin E$ by showing the following claim:

$$\text{For any } b \in E, \text{ if } b \text{ attacks } c \text{ in } SR_4(AAF), \text{ then } c \in E^+. \tag{10.1}$$

Proof of Claim. Let D be the characteristic function of $SR_4(AAF)$, thus $E = \bigcup_{i=1,\dots,\infty} D^i(\varnothing)$. The proof is carried out by induction on the value of i. If $b \in D^1(\varnothing)$, then c does not attack b in $SR_4(AAF)$. By Def. 3.10, it must be either

- $b \rightarrow c$ and $|\mathscr{S}_c| \not> |\mathscr{S}_b|$, or
- $b \rightarrow c$, $c \rightarrow b$ and $|\mathscr{S}_b| > |\mathscr{S}_c|$.

In either case, $c \in E^+$. Assume the claim holds for $i = n$. If $b \in D^{n+1}(\varnothing)$, we distinguish two cases: If c does not attack b in $SR_4(AAF)$, by the same reasoning as above, $c \in E^+$; otherwise, there must be $b' \in D^n(\varnothing)$ such that b' attacks c in $SR_4(AAF)$. Apply the IH, we have $c \in E^+$.

Therefore, $b \in E^+$. Namely, there is $a \in E$ such that $a \to b$. By Def. 3.10, it must be either a attacks b or b attacks a in $SR_4(AAF)$. In both cases, there is $a' \in E$ such that a' attacks b in $SR_4(AAF)$. Since b is arbitrary, we conclude that E defends c in $SR_4(AAF)$. Thus $c \in E$, contradiction! □

Lemma 10.3. *Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$, a set $E \subseteq A$ and a SCC $S \in SCCS_{AAF}$, for all $a \in S$, we have:*

1. *$a \in UP_{AAF}(S, E)$ iff for all $b \in E \backslash S$, $b \not\to a$.*
2. *$a \in U_{AAF}(S, E)$ iff for all $b \in E \backslash S$, $b \not\to a$ and for all $b \in \mathscr{A} \backslash S$ such that $b \to a$, there is $c \in E$ such that $c \to b$.*

Definition 10.1. *Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and a set $C \subseteq \mathscr{A}$, a set $E \subseteq \mathscr{A}$ is an* admissible set in *C iff $E \subseteq C$ and E is admissible in $SR_3(AAF)$. The set of admissible sets in C is denoted as $\mathscr{AS}(AAF, C)$.*

Definition 10.2. *Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and a set $C \subseteq \mathscr{A}$, a set $E \subseteq \mathscr{A}$ is an* complete extension in *C iff E is an admissible set in C and every $a \in C$ which is defended by E in $SR_3(AAF)$ belongs to E. The set of complete extensions in C is denoted as $\mathscr{CE}(AAF, C)$*

Definition 10.3. *Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and a set $C \subseteq \mathscr{A}$, a set $E \subseteq \mathscr{A}$ is a* preferred extension in *C if and only if E ia a maximal element (w.r.t. set inclusion) of $\mathscr{AS}(AAF, C)$. The set of preferred extensions in C is denoted as $\mathscr{PE}(AAF, C)$.*

Definition 10.4. *Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and a set $C \subseteq \mathscr{A}$, a set $E \subseteq \mathscr{A}$ is a* grounded extension in *C if and only if E ia the least element (w.r.t. set inclusion) of $\mathscr{CE}(AAF, C)$. The set of grounded extensions in C is denoted as $\mathscr{GE}(AAF, C)$.*

According to [36], the grounded extension in C exists and is unique for any $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and set $C \subseteq \mathscr{A}$.

Definition 10.5. *Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and a set $C \subseteq \mathscr{A}$, a set $E \subseteq \mathscr{A}$ is a* stable extension in *C if and only if $E \subseteq C$ and E is a stable extension in $SR_3(AAF)$. The set of stable extensions in C is denoted as $\mathscr{SE}(AAF, C)$.*

Lemma 10.4. *Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$, a set of arguments $E \subseteq \mathscr{A}$ that is admissible in $SR_3(AAF)$, and an argument a which is defended by E in $SR_3(AAF)$, denoting $SCC_{AAF}(a)$ as S, it holds that:*

- *$a \in U_{AAF}(S, E)$; and*
- *a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.*

Proof. To show $a \in U_{AAF}(S, E)$, by Lemma 10.3, we need to show that

(1) for all $b \in E \backslash S$, $b \not\to a$, and

(2) for all $b \in \mathscr{A} \backslash S$ such that $b \rightarrow a$, there is $c \in E$ such that $c \rightarrow b$.

For (1), suppose there is $b \in E \backslash S$ such that $b \rightarrow a$. Then $a \not\rightarrow b$ (otherwise $b \in S$, contradiction!). Thus b attacks a in $SR_3(AAF)$ by Def. 3.10. Since a is defended by E in $SR_3(AAF)$, there must be $c \in E$ such that c attacks b in $SR_3(AAF)$. This implies that E is not conflict-free in $SR_3(AAF)$, contradiction! For (2), let $b \in \mathscr{A} \backslash S$ be such that $b \rightarrow a$. Thus $a \not\rightarrow b$. By Def. 3.10, b attacks a in $SR_3(AAF)$. Since a is defended by E in $SR_3(AAF)$, there must be $c \in E$ such that c attacks b in $SR_3(AAF)$. Therefore, it must be that $c \rightarrow b$ by Def. 3.10.

The second item remains to be considered. We first note that it follows from the first item that $E \cap S \subseteq U_{AAF}(S, E)$ because every element of $E \cap S$ is defended by E in $SR_3(AAF)$. Let $b \in UP_{AAF}(S, E)$ be such that b attacks a in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$. Thus b attacks a in $SR_3(AAF)$. Since $b \in UP_{AAF}(S, E)$, by Def. 3.10, for all $c \in E \backslash S$, c does not attack b in $SR_3(AAF)$. But a is defended by E in $SR_3(AAF)$, thus there must be $c \in E \cap S$ such that c attacks b in $SR_3(AAF)$ (therefore, in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$). It follows that a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$ since b is arbitrary. □

Lemma 10.5. *Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, let $E \subseteq \mathscr{A}$ be a set of arguments such that, for all $S \in SCCS_{AAF}$*

$$(E \cap S) \in \mathscr{A}\mathscr{S}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E)).$$

Given $\hat{S} \in SCCS_{AAF}$ and an argument $a \in U_{AAF}(\hat{S}, E)$. Then for all $b \in \mathscr{A} \backslash \hat{S}$ such that b attacks a in $SR_3(AAF)$, there is $a' \in E$ such that a' attacks b in $SR_3(AAF)$.

Proof. By Def. 3.10, $b \rightarrow a$. Since $a \in U_{AAF}(\hat{S}, E)$, there must be $c \in E$ such that $c \rightarrow b$. If c attacks b in $SR_3(AAF)$, then we are done. Otherwise, by Definition 10, the only possibility is that $c \rightarrow b$, $b \rightarrow c$, and $|\mathscr{S}_c| < |\mathscr{S}_b|$. In this case, b attacks c in $SR_3(AAF)$. We show that there must be $c' \in E$ such that c' attacks b in $SR_3(AAF)$. Denote $SCC_{AAF}(b)$ as S'. If $b \in S' \backslash UP_{AAF}(S', E)$, then, by definition, there must be $c' \in E \backslash S'$ such that $c' \rightarrow b$. Note that $b \not\rightarrow c'$ (otherwise $c' \in S'$), thus c' attacks b in $SR_3(AAF)$ by Def. 3.10. If $b \in UP_{AAF}(S', E)$. Note that $E \cap S' \in \mathscr{A}\mathscr{S}(AAF \downarrow_{UP_{AAF}(S',E)}, U_{AAF}(S', E))$, $c \in E \cap S'$, and b attacks c in $SR_3(AAF \downarrow_{UP_{AAF}(S',E)})$. Thus there must be $c' \in E \cap S'$ such that c' attacks b in $SR_3(AAF \downarrow_{UP_{AAF}(S',E)})$ (thus in $SR_3(AAF)$). □

Lemma 10.6. *Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, let $E \subseteq \mathscr{A}$ be a set of arguments such that, for all $S \in SCCS_{AAF}$*

$$(E \cap S) \in \mathscr{A}\mathscr{S}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E)).$$

Given a $\hat{S} \in SCCS_{AAF}$ and an argument $a \in U_{AAF}(\hat{S}, E)$ which is defended by $E \cap \hat{S}$ in $SR_3(AAF \downarrow_{UP_{AAF}(\hat{S},E)})$. Then a is defended by E in $SR_3(AAF)$.

Proof. We first note that $E \cap \hat{S} \subseteq UP_{AAF}(\hat{S}, E)$. Let $b \in \mathscr{A}$ be such that b attacks a in $SR_3(AAF)$. We distinguish three cases: (1) If $b \in UP_{AAF}(\hat{S}, E)$, then b attacks a in $SR_3(AAF \downarrow_{UP_{AAF}(\hat{S},E)})$. Since a is defended by $E \cap \hat{S}$ in $SR_3(AAF \downarrow_{UP_{AAF}(\hat{S},E)})$, there must be $c \in E \cap \hat{S}$ such that c attacks b in $SR_3(AAF \downarrow_{UP_{AAF}(\hat{S},E)})$ (and, thus, in $SR_3(AAF)$). (2) If $b \in \hat{S} \backslash UP_{AAF}(\hat{S}, E)$, by definition, there must be $c \in E \backslash \hat{S}$ such that $c \rightarrow b$. Note that $b \not\rightarrow c$ (otherwise $c \in \hat{S}$), thus c attacks b in $SR_3(AAF)$ by Def. 3.10. (3) If $b \notin \hat{S}$, by Lemma 10.5, there is $a' \in E$ such that a' attacks b in $SR_3(AAF)$.

□

Lemma 10.7. *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ *and a set of arguments* $E \subseteq \mathscr{A}$, $\forall C \subseteq \mathscr{A}$, $E \in \mathscr{AS}(AAF, C)$ *if and only if* $\forall S \in SCCS_{AAF}$

$$(E \cap S) \in \mathscr{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C).$$

Proof. From left to right. We first show that (1) $E \cap S \subseteq U_{AAF}(S,E) \cap C$: It suffices to show $E \cap S \subseteq U_{AAF}(S,E)$ since $E \subseteq C$. Let $a \in E \cap S$. Since E is admissible in $SR_3(AAF)$ and a is defended by E in $SR_3(AAF)$, it follows from Lemma 10.4 that $a \in U_{AAF}(S,E)$. We then show that (2) $E \cap S$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$: It is easy to see that $E \cap S$ is conflict-free in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$. Let $a \in E \cap S$. Since E is admissible in $SR_3(AAF)$ and a is defended by E in $SR_3(AAF)$, it follows from Lemma 10.4 that a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.

From right to left. Since $\forall S \in SCCS_{AAF}$, $(E \cap S) \subseteq C$, it follow that (1) $E \subseteq C$. We then show that (2) E is conflict-free in $SR_3(AAF)$. Let $a, b \in E$. If $SCCS_{AAF}(a) = SCCS_{AAF}(b) = S$, since $E \cap S$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, there is no attack between them in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$ (thus in $SR_3(AAF)$). If $SCCS_{AAF}(a) \neq SCCS_{AAF}(b)$, since $E \cap SCCS_{AAF}(b) \subseteq U_{AAF}(SCCS_{AAF}(b), E)$ and $a \in E \backslash SCCS_{AAF}(b)$, $a \not\rightarrow b$. Thus, by Def. 3.10, a does not attack b in $SR_3(AAF)$. Similarly, b does not attack a in $SR_3(AAF)$. Finally, we show that (3) E can defend itself in $SR_3(AAF)$. Let $a \in E$ and denote $SCCS_{AAF}(a)$ as $\hat{S}$. Since $E \cap \hat{S}$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(\hat{S},E)})$ and $a \in E \cap \hat{S}$, a is defended by $E \cap \hat{S}$ in $SR_3(AAF \downarrow_{UP_{AAF}(\hat{S},E)})$. Thus, apply Lemma 10.6, we have a is defended by E in $SR_3(AAF)$. □

Complete Semantics

Proposition 10.37. *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ *and a set of arguments* $E \subseteq \mathscr{A}$, $\forall C \subseteq \mathscr{A}$, $E \in \mathscr{CE}(AAF, C)$ *if and only if* $\forall S \in SCCS_{AAF}$

$$(E \cap S) \in \mathscr{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C).$$

Proof. From left to right. It follows from Lemma 10.7 that $(E \cap S) \in \mathscr{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$. It remains to show that $\forall a \in U_{AAF}(S,E) \cap C$ such that a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, $a \in E \cap S$. Since a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, by Lemma 10.7, a is defended by E in $SR_3(AAF)$. Since $a \in C$ and E is a complete extension in C, $a \in E$.

From right to left. It follows from Lemma 10.7 that $E \in \mathscr{AS}(AFF, C)$. It remains to show that $\forall a \in C$, if a is defended by E in $SR_3(AAF)$, then $a \in E$. By Lemma 10.7, we know that E is admissible in $SR_3(AAF)$. Since a is defended by E in $SR_3(AAF)$, by Lemma 10.4, $a \in U_{AAF}(S,E)$ and a is defended by $E \cap S$ in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$. Thus $a \in E \cap S$ since $(E \cap S) \in \mathscr{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S,E) \cap C)$. □

Preferred Semantics

Lemma 10.8. *Given an* $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, *an admissible set* E *in* $SR_3(AAF)$, *and a* $S \in SCCS_{AAF}$, *let* $\hat{E}$ *be a set of arguments such that:*

- $E \cap S \subseteq \hat{E} \subseteq U_{AAF}(S,E)$;
- $\hat{E}$ *is admissible in* $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.

It holds that $E \cup \hat{E}$ is admissible in $SR_3(AAF)$.

Proof. We first show that $E \cup \hat{E}$ is conflict-free in $SR_3(AAF)$. Suppose not, then the only possibility is that there is $a \in E \backslash \hat{E}$ (thus $a \in E \backslash S$) and $b \in \hat{E} \backslash E$ such that a attacks b or b attacks a in $SR_3(AAF)$. In the former case, by Def. 3.10, it must be that $a \to b$. But it contradicts that $b \in \hat{E} \subseteq U_{AAF}(S, E)$. In the latter case, since E is admissible in $SR_3(AAF)$, there must be $a' \in E$ such that a' attacks b in $SR_3(AAF)$. $a' \notin \hat{E}$ (otherwise $\hat{E}$ is not conflict-free in $SR_3(AAF)$), thus $a' \in E \backslash \hat{E}$. This is also impossible by the same reason mentioned before.

We then show $E \cup \hat{E}$ can defend itself in $SR_3(AAF)$. It suffices to show that for all $a \in \hat{E} \backslash E$, $E \cup \hat{E}$ defends a in $SR_3(AAF)$. Let $b \in \mathscr{A}$ be such that b attacks a in $SR_3(AAF)$. We distinguish three cases: (a) If $b \in UP_{AAF}(S, E)$, then b attacks a in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$. Since $\hat{E}$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, there must be $c \in \hat{E}$ such that c attacks b in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$ (thus in $SR_3(AAF)$). (b) If $b \in S \backslash UP_{AAF}(S, E)$, by definition, there must be $c \in E \backslash S$ such that $c \to b$. Since $b \not\to c$, by Def. 3.10, c attacks b in $SR_3(AAF)$. (c) If $b \in \mathscr{A} \backslash S$. We first note that, by Lemma 10.7, we have (1) $\forall S \in SCCS_{AAF}$: $(E \cap S) \in \mathscr{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E))$. We also have (2) $a \in U_{AAF}(S, E)$. Based on (1) and (2), we can apply Lemma 10.5 and obtain that there is $c \in E$ such that c attacks b in $SR_3(AAF)$. □

Proposition 10.38. *Given an $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathscr{A}$, $\forall C \subseteq \mathscr{A}$, $E \in \mathscr{PE}(AAF, C)$ if and only if $\forall S \in SCCS_{AAF}$*

$$(E \cap S) \in \mathscr{PE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C).$$

Proof. From left to right. By Lemma 10.7, it follows that $(E \cap S) \in \mathscr{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C)$. Suppose, towards to a contradiction, that $E \cap S$ is not a maximal element (w.r.t set inclusion) in $\mathscr{AS}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C)$, i.e., there is $\hat{E} \in \mathscr{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C)$ such that $\hat{E} \supset (E \cap S)$. Note that:

- $E \cap S \subseteq \hat{E} \subseteq U_{AAF}(S, E)$;
- $\hat{E}$ is admissible in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.

Thus we can apply Lemma 10.8 and obtain that $E \cup \hat{E}$ is admissible in $SR_3(AAF)$. Since $E \cup \hat{E} \supset E$, it contradicts that $E \in \mathscr{PE}(AAF, C)$.

From right to left. By Lemma 10.7, it follows that $E \in \mathscr{AS}(AAF, C)$. Suppose, towards to a contradiction, that there is $\hat{E} \supset E$ and $\hat{E} \in \mathscr{AS}(AAF, C)$. Then there is $S \in SCCS_{AAF}$ such that $\hat{E} \cap S \supset E \cap S$. Let $\hat{S} \in SCCS_{AAF}$ be such that

(1) $\hat{E} \cap \hat{S} \supset E \cap \hat{S}$,

(2) for any $S \in SCCS_{AAF}$ such that S is an ancestor of $\hat{S}$ in the condensation of the graph $(\mathscr{A}, \to)$, $\hat{E} \cap S = E \cap S$.

By Lemma 10.7, we have that

$$(\hat{E} \cap \hat{S}) \in \mathscr{AS}(AAF \downarrow_{UP_{AAF}(\hat{S},\hat{E})}, U_{AAF}(\hat{S}, \hat{E}) \cap C).$$

But, by (2), it is easy to see that $UP_{AAF}(\hat{S}, \hat{E}) = UP_{AAF}(\hat{S}, E)$ and $U_{AAF}(\hat{S}, \hat{E}) = U_{AAF}(\hat{S}, E)$. Thus

$$(\hat{E} \cap \hat{S}) \in \mathscr{AS}(AAF \downarrow_{UP_{AAF}(\hat{S},E)}, U_{AAF}(\hat{S}, E) \cap C).$$

But $\hat{E} \cap \hat{S} \supset E \cap \hat{S}$, it contradicts that $(E \cap \hat{S}) \in \mathscr{PE}(AAF \downarrow_{UP_{AAF}(\hat{S},E)}, U_{AAF}(\hat{S}, E) \cap C)$. □

Grounded Semantics

Proposition 10.39. *Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathscr{A}$, $\forall C \subseteq \mathscr{A}$, $E \in \mathscr{GE}(AAF, C)$ if and only if $\forall S \in SCCS_{AAF}$*

$$(E \cap S) \in \mathscr{GE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C).$$

Proof. From left to right. By Proposition 10.37, we have that $\forall S \in SCCS_{AAF}$: $(E \cap S) \in \mathscr{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C)$. Suppose, towards to a contradiction, that there is $S \in SCCS_{AAF}$ such that $E \cap S$ is not the least element in $\mathscr{CE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C)$. Let $\hat{S} \in SCCS_{AAF}$ be such that

- for any $S \in SCCS_{AAF}$ such that S is an ancestor of $\hat{S}$ in the condensation of the graph $(\mathscr{A}, \rightarrow)$: $E \cap S \in \mathscr{GE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C)$.
- $\exists \hat{E} \subset (E \cap \hat{S}) : \hat{E} \in \mathscr{GE}(AAF \downarrow_{UP_{AAF}(\hat{S},E)}, U_{AAF}(\hat{S}, E) \cap C)$

We can thus construct a set $E' \subseteq \mathscr{A}$ such that

- for any $S \in SCCS_{AAF}$ such that S is an ancestor of $\hat{S}$ in the condensation of the graph $(\mathscr{A}, \rightarrow)$: $(E' \cap S) = (E \cap S)$;
- $E' \cap \hat{S} = \hat{E}$;
- $\forall S \in SCCS_{AAF}$: $(E' \cap S) \in \mathscr{GE}(AAF \downarrow_{UP_{AAF}(S,E')}, U_{AAF}(S, E') \cap C)$.

It is easy to see that $E' \in \mathscr{CE}(AAF, C)$ by Proposition 10.37. But $E \nsubseteq E'$, contradiction!

From right to left. By Proposition 10.37, we have $E \in \mathscr{CE}(AAF, C)$. Suppose that $\exists E' \in \mathscr{GE}(AAF, C)$: $E' \subset E$. Then there must be $S \in SCCS_{AAF}$ such that $E' \cap S \subset E \cap S$. Let $\hat{S} \in SCCS_{AAF}$ be such that

- for any $S \in SCCS_{AAF}$ such that S is an ancestor of $\hat{S}$ in the condensation of the graph $(\mathscr{A}, \rightarrow)$: $E' \cap S = E \cap S$;
- $E' \cap \hat{S} \subset E \cap \hat{S}$.

Since $UP_{AAF}(\hat{S}, E') = UP_{AAF}(\hat{S}, E)$, $U_{AAF}(\hat{S}, E') = U_{AAF}(\hat{S}, E)$, and $E' \cap \hat{S} \in \mathscr{CE}(AAF \downarrow_{UP_{AAF}(\hat{S},E')}, U_{AAF}(\hat{S}, E') \cap C)$ (by Proposition 10.37), we have $E' \cap \hat{S} \in \mathscr{CE}(AAF \downarrow_{UP_{AAF}(\hat{S},E)}, U_{AAF}(\hat{S}, E) \cap C)$. But $E' \cap \hat{S} \subset E \cap \hat{S}$, contradiction! □

Stable Semantics

Lemma 10.9. *Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathscr{A}$, E is a stable extension of $SR_3(AAF)$ if and only if $\forall S \in SCCS_{AAF}$: $E \cap S$ is a stable extension of $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.*

Proof. From left to right. We need to show that

(1) $E \cap S \subseteq UP_{AAF}(S, E)$.

(2) $E \cap S$ is conflict-free in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.

(3) $\forall a \in UP_{AAF}(S, E)$: $a \notin (E \cap S)$ implies that $\exists b \in E \cap S$ such that b attacks a in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$.

For (1), let $a \in E \cap S$ be arbitrary. Suppose, towards to a contradiction, that there is $b \in E \backslash S$ such that $b \to a$. We have $a \not\to b$ since otherwise $b \in S$. Thus b attacks a in $SR_3(AAF)$ by Def. 3.10, contradicting that E is conflict-free in $SR_3(AAF)$. For (2), it is easy to see that $E \cap S$ is conflict-free in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, since otherwise E is not conflict-free in $SR_3(AAF)$. For (3), let $a \in UP_{AAF}(S, E) \backslash (E \cap S)$ be arbitrary. Since $a \in UP_{AAF}(S, E)$, there is no $b \in E \backslash S$ such that $b \to a$. Thus, by Def. 3.10, there is no $b \in E \backslash S$ such that b attacks a in $SR_3(AAF)$. But E is a stable extension of $SR_3(AAF)$, thus it can only be that $\exists b \in E \cap S$ such that b attacks a in $SR_3(AAF)$ (thus in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$).

From right to left. By Lemma 10.7, we have that E is admissible in $SR_3(AAF)$, thus E is conflict-free in $SR_3(AAF)$. Let $a \in \mathscr{A} \backslash E$ be arbitrary. Denote $SCCS_{AAF}(a)$ as S. If $b \notin UP_{AAF}(S, E)$, then $\exists b \in E \backslash S : b \to a$. We have $a \not\to b$, since otherwise $b \in S$. Thus, by Def. 3.10, b attacks a in $SR_3(AAF)$. If $b \in UP_{AAF}(S, E)$, since $E \cap S$ is a stable extension of $SR_3(AAF \downarrow_{UP_{AAF}(S,E)})$, there must be $b \in E \cap S$ such that b attacks a in $SR_3(AAF \downarrow_{UP_{AAF}(S,E)}$) (thus in $SR_3(AAF)$). □

Proposition 10.40. *Given an* $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ *and a set of arguments* $E \subseteq \mathscr{A}$, $\forall C \subseteq \mathscr{A}$, $E \in \mathscr{SE}(AAF, C)$ *if and only if* $\forall S \in SCCS_{AAF}$

$$(E \cap S) \in \mathscr{SE}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C).$$

Proof. From left to right. By Lemma 10.9, $E \cap S$ is a stable extension of $AAF \downarrow_{UP_{AAF}(S,E)}$. Thus it suffices to show that $E \cap S \subseteq U_{AAF}(S, E)$. Let $a \in E \cap S$ be arbitrary and $b \in \mathscr{A} \backslash S$ be such that $b \to a$. We have $a \not\to a$, since otherwise $b \in S$. Thus, by Def. 3.10, b attacks a in $SR_3(AAF)$. Since E is conflict-free, $b \notin E$. Thus, E attacks b in $SR_3(AAF)$. Therefore, by Def. 3.10, there must be $c \in E$ such that $c \to b$. Since b is arbitrary, $a \in U_{AAF}(S, E)$.

From right to left. By Lemma 10.9, we have that E is a stable extension of $SR_3(AAF)$. Besides, it is easy to see that $E \subseteq C$. □

Proposition 10.41 (SR$_3 \times$ P3.4)**.** *All the four kinds of Dung's semantics for SR$_3$ satisfy P3.4.*

Proof. Let us consider, for example, the stable semantics for SR$_3$. By definition, it is easy to see that E is a stable extension of $SR_3(AAF)$ if and only if $E \in \mathscr{SE}(AAF, \mathscr{A})$ for any given $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and $E \subseteq \mathscr{A}$. For any AAF with $|SCCS_{AAF}| = 1$, define the function $\mathscr{B}$ as such that $\mathscr{B}(AAF, C) = \mathscr{SE}(AAF, C)$. Then, by Proposition 10.40, we have that the function $\mathscr{G} = \mathscr{SE}$ satisfies all the conditions in P3.4.

As far as complete, preferred and grounded semantics for SR$_3$ are concerned, proof are similar and are based on Proposition 10.37, 10.38 and 10.39, respectively. □

Proposition 10.42 (AR$_1$ × P3.1, P3.2)**.** *All the four kinds of Dung's semantics for AR$_1$ do not satisfy P3.2. Thus, they do not satisfy P3.2, as well.*

Proof. A counter example is Figure 10.1. □

Proposition 10.45. *(AR_4 × P3.1) All the four kinds of Dung's semantics for AR_4 satisfy P3.2.*

Proof. Obviously from Proposition 10.42. □

Proposition 10.46 **(AR_2 × P3.2).** *All the four kinds of Dung's semantics for AR_2 do not satisfy P3.2.*

Proof. A counter example is Figure 10.1. □

Proposition 10.47 **(AR_3 × P3.2).** *All the four kinds of Dung's semantics for AR_3 satisfy P3.2.*

Proof. Let an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given. It suffices to show that any admissible set $E \subseteq \mathscr{A}$ in $AR_3(AAF)$ is also admissible in $\langle \mathscr{A}, \rightarrow \rangle$. Suppose, towards to a contradiction, that E is not admissible in $\langle \mathscr{A}, \rightarrow \rangle$. By Proposition 10.42, it must be that there is $c \in E$ such that E cannot defend c in $\langle \mathscr{A}, \rightarrow \rangle$. That is, there is $b \in \mathscr{A}$ such that $b \rightarrow c$ and $a \not\rightarrow b$ for all $a \in E$. Thus, particularly, $c \not\rightarrow b$. By Def. 3.10, it follows that b attacks c in $AR_3(AAF)$. Since $a \not\rightarrow b$ for all $a \in E$, it holds that a does not attack b for all $a \in E$ in $AR_3(AAF)$. Thus, E cannot defend c in $AR_3(AAF)$, contradiction! □

Proposition 10.48 **(AR_4 × P3.2).** *The complete semantics, preferred semantics and stable semantics for AR_4 do not satisfy P3.2, whereas the grounded semantics for AR_4 do satisfy P3.2.*

Proof. For the first half, a counter example is Figure 10.9. Note that $\{b\}$ is one of complete extensions, preferred extensions and stable extensions of $AR_4(AAF)$, but it is not admissible in $\langle \mathscr{A}, \rightarrow \rangle$.

For the second half of the lemma, let an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and $E \subseteq \mathscr{A}$ be the grounded extension of $AR_4(AAF)$. Since E is conflict-free in $\langle \mathscr{A}, \rightarrow \rangle$ by Proposition 10.45, we only need to show E can defend all of its members in $\langle \mathscr{A}, \rightarrow \rangle$. Let D be the characteristic function of $AR_4(AAF)$, thus $E = \bigcup_{i=1,\dots,\infty} D^i(\varnothing)$. We prove, by induction on the value of i, that E can defend $D^i(\varnothing)$ in $\langle \mathscr{A}, \rightarrow \rangle$ for $i = 1, \dots, \infty$. If $i = 1$, let $c \in D^1(\varnothing) = D(\varnothing)$ be arbitrary. For any argument b such that $b \rightarrow c$ and $c \not\rightarrow b$, it must be that b attacks c in $AR_4(AAF)$ by Def. 3.10. Thus such b does not exist, so E can defend $D^1(\varnothing)$ in $\langle \mathscr{A}, \rightarrow \rangle$. Now suppose that E can defend $D^n(\varnothing)$ in $\langle \mathscr{A}, \rightarrow \rangle$. Let $c \in D^{n+1}(\varnothing)$ be arbitrary. For any argument b such that $b \rightarrow c$ and $c \not\rightarrow b$, it must be that b attacks c in $AR_4(AAF)$ by Def. 3.10. Since $c \in D^{n+1}(\varnothing) = D(D^n(\varnothing))$, there must be $a \in D^n(\varnothing)$ such that a attacks b in $AR_4(AAF)$. If $a \not\rightarrow b$, it must hold that $b \rightarrow a$ by Def. 3.10. By IH, there is $a' \in E$ such that $a' \rightarrow b$, as desired. □

Proposition 10.49 **(AR_1, AR_3 × P3.3).** *The grounded, complete, preferred semantics for AR_1 and AR_3 satisfy P3.3, whereas the stable semantics for AR_1 and AR_3 do not satisfy P3.3.*

Proof. For any $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$, if a set of arguments $\mathscr{U} \subseteq \mathscr{A}$ is unattacked in $\langle \mathscr{A}, \rightarrow \rangle$, then it is also unattacked in $AR_1(AAF)$ and $AR_3(AAF)$. Thus the first half follows from [35] directly.

For the second half of the lemma, consider the following $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S} = \{\alpha\}, \sqsubset \rangle$ in Figure 10.16. It is easy to see that $AR_1(AAF) = AR_3(AAF) = \langle \mathscr{A}, \rightarrow \rangle$. $\{a\}$ is unattacked and is a stable extension of $AAF \downarrow_{\{a\}}$. But there is no stable extension of $\langle \mathscr{A}, \rightarrow \rangle$. □

Proposition 10.50 (AR_2, AR_4 × P3.3). *All the four kinds of Dung's semantics for AR_2 and/ or AR_4 do not satisfy P3.3.*

Proof. The counter examples are the same as in Proposition 10.30 and Proposition 10.31. □

Proposition 10.51 (AR_1, AR_2, AR_4 × P3.4). *All the four kinds of Dung's semantics for AR_1, AR_2, and AR_4 do not satisfy P3.4.*

Proof. See Proposition 10.32. □

Proposition 10.52 (AR_3× P3.4). *All the four kinds of Dung's semantics for AR_3 satisfy P3.4.*

Proof. The proofs are similar to those for SR_3. We just substitute AR_3 for SR_3 in them. □

Proposition 10.53 (AR_1 × P3.5). *The grounded, complete, and preferred semantics for AR_1 do not satisfy the P3.5, whereas the stable semantics for AR_1 do satisfy the P3.5.*

Proof. See Proposition 10.33. □

Proposition 10.54 (AR_2 × P3.5). *All the four kinds of Dung's semantics for AR_2 does not satisfy P3.5.*

Proof. See Proposition 10.34. □

Proposition 10.55 (AR_3 × P3.5). *All the four kinds of Dung's semantics for AR_3 satisfy P3.5.*

Proof. The proof can be obtained by substituting AR_3 for SR_3 in the proof of Proposition 10.35. □

Proposition 10.56 (AR_4 × P3.5). *The complete, preferred and stable semantics for AR_4 does not satisfy P3.5, whereas the grounded semantics for AR_4 do satisfy P3.5.*

Proof. The proof is similar to that of Proposition 10.36. □

Proposition 10.58 (OR × P3.2). *All the four kinds of Dung's semantics for OR do not satisfy P3.2.*

Proof. A counter example is Figure 10.1. □

Proposition 10.59 (NBR × P3.1, P3.2). *All the four kinds of Dung's semantics for NBR do not satisfy P3.1 and P3.2.*

Proof. A counter example is Figure 10.1. □

Proposition 10.60 (OR × P3.3). *The complete, grounded and preferred semantics for OR satisfy P3.3, whereas the stable semantics for OR does not satisfy P3.3.*

Proof. Let an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and $\mathscr{U} \subseteq \mathscr{A}$ be an unattacked set. Since the set of arguments in $OR(AAF \downarrow_{\mathscr{U}})$ is unattacked in $OR(AAF)$, the first half of the lemma follows directly from [35]. For the second half of the lemma, a counter example is as follows: □

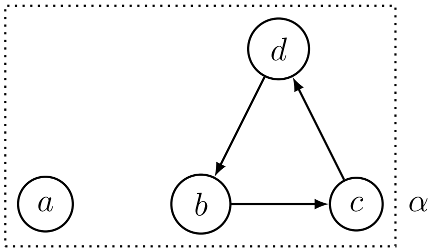


Figure 10.16: Stable semantics for OR do not satisfy P3.3

Proposition 10.61 (NBR × P3.3)**.** *The complete, grounded and preferred semantics for NBR satisfy P3.3, whereas the stable semantics for NBR does not satisfy P3.3.*

Proof. The proof is similar to that of Proposition 10.60. □

Proposition 10.62 (NBR× P3.4)**.** *All the four kinds of Dung's semantics for NBR do not satisfy P3.4.*

Proof. A counter example is Figure 10.12. □

Proposition 10.63 (OR × P3.5)**.** *All the four kinds of Dung's semantics for OR satisfy P3.5.*

Proof. Let an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and let $OR(AAF) = (\mathscr{A}', \rightarrow')$. We show that, for any $E \subseteq \mathscr{A}'$, $OR(AAF^E) = (OR(AAF))^E$, where $(OR(AAF))^E$ is the E-reduct of $OR(AAF)$ defined in [39]. Let $(OR(AAF))^E = (\mathscr{A}'', \rightarrow'')$, we first show $\mathscr{A}' = \mathscr{A}''$.

$a \in \mathscr{A}'$
iff $a \notin E$, a is not attacked by E in AAF, and there is $\alpha \in \mathscr{S}$ s.t. $a \sqsubset \alpha$ (by definition)
iff there is $\alpha \in \mathscr{S}$ s.t. $a \sqsubset \alpha$, $a \notin E$, and a is not attacked by E in $OR(AAF)$ (since $E \subseteq \mathscr{A}'$)
iff $a \in \mathscr{A}''$

We then show $\rightarrow' = \rightarrow''$. For any $a, b \in \mathscr{A}' = \mathscr{A}''$:

$$a \rightarrow' b \text{ iff } a \rightarrow b \text{ iff } a \rightarrow'' b.$$

Thus, $OR(AAF^E) = (OR(AAF))^E$. Therefore, the lemma follows directly from [39]. □

Proposition 10.64 (NBR × P3.5)**.** *The grounded, complete, and preferred semantics for NBR do not satisfy the P3.5, whereas the stable semantics for NBR do satisfy the P3.5.*

Proof. Consider the Figure 10.15. It is easy to see that $\{a\}$ is the grounded extension of $SR_1(AAF)$, as well a complete extension and a preferred extension. Let $E = \{a\}$, then then AAF^E consists of the single point b. It is also easy to see that $\{b\}$ is the grounded, complete and preferred extension of AAF^E. However, $\{a, b\}$ is not admissible in $NBR(AAF)$.

For the second half of the proposition, let $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ be given and E be a stable extension of $NBR_1(AAF)$. It is easy to know that E^* is empty, thus E' is also empty. □

Complete, Grounded and Preferred Semantics

In this subsection, let an $AAF_1 = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ *be given. Define* $AAF_2 = \langle \mathscr{A}, \rightarrow, \{\alpha\}, \sqsubset' \rangle$ *as such that* $a \sqsubset' \alpha$ *if and only if* $\exists \beta \in \mathscr{S}$ *such that* $a \sqsubset \beta$. *It is easy to see that* $OR(AAF_1) = OR(AAF_2)$. *Denote* $OR(AAF_1) = OR(AAF_2)$ *as* $\langle \mathscr{A}', \rightarrow' \rangle$. *We also note that* $SCCS_{AAF_1} = SCCS_{AAF_2}$, $UP_{AAF_1}(S, E) = UP_{AAF_2}(S, E)$ *and* $U_{AAF_1}(S, E) = U_{AAF_2}(S, E)$ *for any* $E \subseteq \mathscr{A}$ *and* $S \in SCCS_{AAF_1}$.

Lemma 10.10. *The following hold:*

(1) For any arguments a, b *in* $\mathscr{A}'$, $a \rightarrow' b$ *if and only* a *attacks* b *in* $SR_3(AAF_2)$.

(2) $\mathscr{A}'$ *is unattacked in* $SR_3(AAF_2)$.

Proof. Immediately from Def. 3.10. □

Lemma 10.11. *For any* $E \subseteq \mathscr{A}$, E *is a complete extension in* $OR(AAF_1) = OR(AAF_2)$ *if and only if* $E \in \mathscr{CE}(AAF_2, \mathscr{A}')$ *where the function* $\mathscr{CE}$ *is defined in Definition 10.2.*

Proof. From left to right. We need to show that

(1) $E \subseteq \mathscr{A}'$.

(2) E is conflict-free in $SR_3(AAF_2)$.

(3) for every $a \in E$, E defends a in $SR_3(AAF_2)$.

(4) for every $a \in \mathscr{A}'$, if E defends a in $SR_3(AAF_2)$ then $a \in E$.

The first item is trivial. The second item follows directly from Lemma 10.10 (1). For (3), let $b \in \mathscr{A}$ be such that b attacks a in $SR_3(AAF_2)$. Since $\mathscr{A}'$ is unattacked in $SR_3(AAF_2)$ and $a \in E \subseteq \mathscr{A}'$, $b \in \mathscr{A}'$. Thus $b \rightarrow' a$ by Lemma 10.10 (1). Since E is a complete extension in $OR(AAF_2)$, there must be $c \in E$ such that $c \rightarrow' b$. Thus, by Lemma 10.10 (1) again, c attacks b in $SR_3(AAF_2)$. Since b is arbitrary, E defends a in $SR_3(AAF_2)$. For (4), we show that E defends a in $OR(AAF_2)$, thus $a \in E$ since E is a complete extension in $OR(AAF_2)$. Let $b \in \mathscr{A}'$ be such that $b \rightarrow' a$, then b attacks a in $SR_3(AAF_2)$ by Lemma 10.10 (1). Since E defends a in $SR_3(AAF_2)$, there must be $c \in E$ such that c attacks b in $SR_3(AAF_2)$. By Lemma 10.10 (1) again, $c \rightarrow' b$. Since b is arbitrary, E defends a in $OR(AAF_2)$.

From right to left. We need to show that

(1) $E \subseteq \mathscr{A}'$.

(2) E is conflict-free in $OR(AAF_2)$.

(3) for every $a \in E$, E defends a in $OR(AAF_2)$.

(4) for every $a \in \mathscr{A}'$, if E defends a in $OR(AAF_2)$ then $a \in E$.

The first item is trivial. The second item follows directly from Lemma 10.10 (1). The remaining proof is similar to that for the direction from left to right. □

Lemma 10.12. *The following hold for any* $E \subseteq \mathscr{A}$:

(1) *E is a preferred extension in $OR(AAF_1) = OR(AAF_2)$ if and only if $E \in \mathscr{PE}(AAF_2, \mathscr{A}')$ where the function $\mathscr{PE}$ is defined in Definition 10.3.*

(2) *E is a grounded extension in $OR(AAF_1) = OR(AAF_2)$ if and only if $E \in \mathscr{GE}(AAF_2, \mathscr{A}')$ where the function $\mathscr{GE}$ is defined in Definition 10.4.*

Proof. Directly from Lemma 10.11. □

Proposition 10.65. *The complete, preferred and grounded semantics for OR satisfy P3.4.*

Proof. Let us consider, for example, the complete semantics for OR. Let $\mathscr{G}(AAF_1 \downarrow_X, Y) = \mathscr{CE}(AAF_2 \downarrow_X, Y \cap \mathscr{A}')$ for any $Y \subseteq X \subseteq \mathscr{A}$. Then $\mathscr{G}(AAF_1, \mathscr{A}) = \mathscr{G}(AAF_1 \downarrow_{\mathscr{A}}, \mathscr{A}) = \mathscr{CE}(AAF_2 \downarrow_{\mathscr{A}}, \mathscr{A} \cap \mathscr{A}') = \mathscr{CE}(AAF_2, \mathscr{A}')$. By Lemma 10.11, we know that, for any $E \subseteq \mathscr{A}$, E is a complete extension in $OR(AAF_1)$ if and only if $E \in \mathscr{G}(AAF_1, \mathscr{A})$. It remains to show that for any $E \subseteq \mathscr{A}$ and $C \subseteq \mathscr{A}$: $E \in \mathscr{G}(AAF_1, C)$ if and only if $\forall S \in SCCS_{AAF_1}$

$$(E \cap S) \in \mathscr{G}(AAF_1 \downarrow_{UP_{AAF_1}(S,E)}, U_{AAF_1}(S, E) \cap C).$$

We have the following equivalent conditions:
$E \in \mathscr{G}(AAF_1, C)$ iff $E \in \mathscr{CE}(AAF_2, C \cap \mathscr{A}')$ (by Definition of $\mathscr{G}$)
iff $\forall S \in SCCS_{AAF_2}$: $(E \cap S) \in \mathscr{CE}(AAF_2 \downarrow_{UP_{AAF_2}(S,E)}, U_{AAF_2}(S, E) \cap (C \cap \mathscr{A}'))$ (by Proposition 10.37)
iff $\forall S \in SCCS_{AAF_1}$: $(E \cap S) \in \mathscr{CE}(AAF_2 \downarrow_{UP_{AAF_1}(S,E)}, U_{AAF_1}(S, E) \cap (C \cap \mathscr{A}'))$ (by ($SCCS_{AAF_1} = SCCS_{AAF_2}$, $UP_{AAF_1}(S, E) = UP_{AAF_2}(S, E)$ and $U_{AAF_1}(S, E) = U_{AAF_2}(S, E)$))
iff $\forall S \in SCCS_{AAF_1}$: $(E \cap S) \in \mathscr{G}(AAF_1 \downarrow_{UP_{AAF_1}(S,E)}, U_{AAF_1}(S, E) \cap C)$ (by Definition of $\mathscr{G}$)

As far as preferred and grounded semantics for OR are concerned, proof are similar and are based on Proposition 10.38 and 10.39, respectively. □

Stable Semantics

Lemma 10.13. *Given an $AAF = \langle \mathscr{A}, \rightarrow, \mathscr{S}, \sqsubset \rangle$ and a set of arguments $E \subseteq \mathscr{A}$, E is a stable extension of $OR(AAF)$ if and only if $\forall S \in SCCS_{AAF}$: $E \cap S$ is a stable extension of $OR(AAF \downarrow_{UP_{AAF}(S,E)})$.*

Proof. Denote $OR(AAF)$ as $\langle \mathscr{A}', \rightarrow' \rangle$.

From left to right. We need to show that

(1) $E \cap S \subseteq \mathscr{A}' \cap UP_{AAF}(S, E)$.

(2) $E \cap S$ is conflict-free in $OR(AAF \downarrow_{UP_{AAF}(S,E)})$.

(3) $\forall a \in \mathscr{A}' \cap UP_{AAF}(S, E)$: $a \notin E \cap S$ implies that $\exists b \in E \cap S$ such that b attacks a in $OR(AAF \downarrow_{UP_{AAF}(S,E)})$.

For (1), since E is a stable extension of $OR(AAF)$, $E \cap S \subseteq E \subseteq \mathscr{A}'$. It remains to show $E \cap S \subseteq UP_{AAF}(S, E)$. Let $a \in E \cap S$. Since E is conflict-free in $\langle \mathscr{A}, \rightarrow \rangle$, there is no $b \in E \backslash S$ such that $b \rightarrow a$. Thus $a \in UP_{AAF}(S, E)$. For (2), it is obvious that $E \cap S$ is conflict-free in $OR(AAF \downarrow_{UP_{AAF}(S,E)})$ since otherwise E would not be conflict-free in $\langle \mathscr{A}, \rightarrow \rangle$. For (3), since $a \in \mathscr{A}' \backslash E$ and E is a stable extension of $OR(AAF)$, there must be $b \in E$ such that $b \rightarrow a$. Since $a \in UP_{AAF}(S, E)$, $b \notin E \backslash S$. Thus $b \in E \cap S$. Since $b \rightarrow a$ and $b \in E \cap S \subseteq \mathscr{A}' \cap UP_{AAF}(S, E)$, b attacks a in $OR(AAF \downarrow_{UP_{AAF}(S,E)})$.

From right to left. We need to show that

(1) $E \subseteq \mathscr{A}'$.

(2) E is conflict-free in $OR(AAF)$.

(3) $\forall a \in \mathscr{A}'$: $a \notin E$ implies that $\exists b \in E$ such that $b \to a$.

For (1), since $E \cap S \subseteq \mathscr{A}'$ for all $S \in SCCS_{AAF}$, $E \subseteq \mathscr{A}'$. For (2), suppose, toward to a contradiction, that there are $a, b \in E$ such that $a \to b$. We distinguish two cases: (a) If $SCCS_{AAF}(a) = SCCS_{AAF}(b) = S$, it implies that $E \cap S$ is not conflict-free in $OR(AAF \downarrow_{UP_{AAF}(S,E)})$, contradiction!

(b) If $SCCS_{AAF}(a) \neq SCCS_{AAF}(b)$. Denote $SCCS_{AAF}(b)$ as S. Thus $b \in E \cap S$ and $b \notin UP_{AAF}(S, E)$, contradicting that $E \cap S \subseteq \mathscr{A}' \cap UP_{AAF}(S, E)$. □

Proposition 10.66. *The stable semantics for OR satisfies P3.4.*

Proof. For any $AAF = \langle \mathscr{A}, \to, \mathscr{S}, \sqsubset \rangle$ and $C \subseteq \mathscr{A}$, let $\mathscr{G}(AAF, C)$ be the set of stable extensions in $OR(AAF)$. Thus, for any $E \subseteq \mathscr{A}$, E is a stable extension in $OR(AAF)$ iff $E \in \mathscr{A}(AAF, \mathscr{A})$. On the other hand, we have that

	$E \in \mathscr{G}(AAF, C)$	
iff	E is a stable extension in $OR(AAF)$	(Definition of $\mathscr{G}$)
iff	$\forall S \in SCCS_{AAF}$: $(E \cap S)$ is a stable extension of $OR(AAF \downarrow_{UP_{AAF}(S,E)})$	(Lemma 10.13)
iff	$\forall S \in SCCS_{AAF}$: $(E \cap S) \in \mathscr{G}(AAF \downarrow_{UP_{AAF}(S,E)}, U_{AAF}(S, E) \cap C)$	(Definition of $\mathscr{G}$)

□

10.2 Proofs of The Principle-based Analysis of Abstract Agent Argumentation

Proposition 10.67. *σ_0^x satisfy P4.1.*

Proof. Obviously. □

Proposition 10.68. *σ_1^x do not satisfy P4.1.*

Proof. We use a counterexample from Figure 10.17. From the left part of the Figure, we can see that $\sigma_1^{c,g,p} = \{\{c, d\}\}$ and $\sigma_1^s = \varnothing$ while when adding the support from c to a as in the right part of the Figure, $\sigma_1^{c,g,p,s} = \{\{a, c, d\}\}$. □

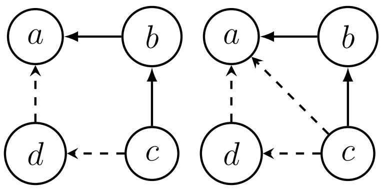


Figure 10.17: σ_1^x do not satisfy P4.1

Proposition 10.69. *σ_2^x do not satisfy P4.1.*

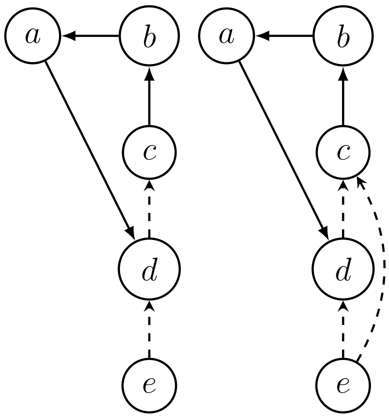


Figure 10.18: σ_2^x do not satisfy P4.1

Proof. We use a counterexample from Figure 10.18. From the left part of the Figure, we can see that $\sigma_2^{c,g,p} = \{\{c, e\}\}$ and $\sigma_1^s = \varnothing$ while when adding the support from e to c as in the right part of the Figure, $\sigma_2^{c,g,p,s} = \{\{a, c, e\}\}$. □

Proposition 10.70. *σ_3^x do not satisfy P4.1.*

Proof. We use a counterexample from Figure 10.19. From the left part of the Figure, we can see that $\sigma_3^{c,g,p,s} = \{\{a, c, e, f\}\}$ while when adding the support from f to b as in the right part of the Figure, $\sigma_3^{c,g,p} = \{\{c, e, f\}\}$ and $\sigma_3^s = \varnothing$ □

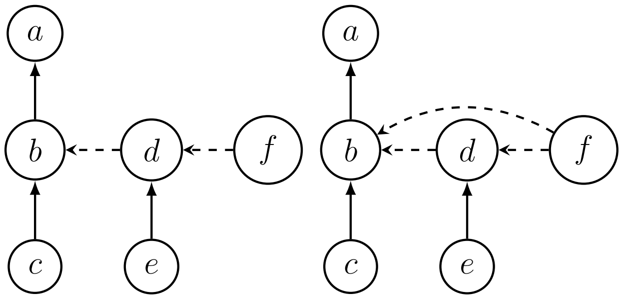


Figure 10.19: σ_3^x do not satisfy P4.1

Proposition 10.71. *$\sigma_4^{c,p,s}$ do not satisfy P4.1.*

Proof. We use a counter example from Figure 10.20. $\sigma_0^{c,p,s} = \{\{a, c, e\}, \{b, d, f\}\}$. In the left part of the Figure, we have $NS_I(\mathscr{F}, \{a, c, e\}) = NS_I(\mathscr{F}, \{b, d, f\}) = 0$ so $\sigma_4^{c,p,s} = \{\{a, c, e\}, \{b, d, f\}\}$. In the right part of the Figure, when adding the support from a to c, we have $NS_I(\mathscr{F}, \{a, c, e\}) = 1$ while $NS_I(\mathscr{F}, \{b, d, f\}) = 0$ so $\sigma_4^{c,p,s} = \{\{a, c, e\}\}$.

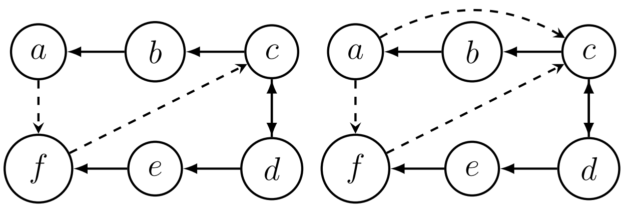


Figure 10.20: σ_4^x do not satisfy P4.1

□

Proposition 10.72. *$\sigma_5^{c,p,s}$ do not satisfy P4.1.*

Proof. We use a counterexample from Figure 10.21. $\sigma_0^c = \{\{a\},\{a,c,e\},\{a,d,f\}\}$, $\sigma_0^{p,s} = \{\{a,c,e\},\{a,d,f\}\}$. In the left part of the Figure we have $NS_E(\mathscr{F},\{a\}) = NS_E(\mathscr{F},\{a,c,e\}) = NS_E(\mathscr{F},\{a,d,f\}) = 1$ so $\sigma_5^c = \sigma_0^c$ and $\sigma_5^{p,s} = \sigma_0^{p,s}$. In the right part of the Figure, when adding the support from e to a, we have $NS_E(\mathscr{F},\{a,c,e\}) = 1$ while $NS_E(\mathscr{F},\{a\}) = NS_E(\mathscr{F},\{a,d,f\}) = 2$ so $\sigma_5^c = \{\{a\},\{a,d,f\}\}$ and $\sigma_5^{p,s} = \{\{a,d,f\}\}$.

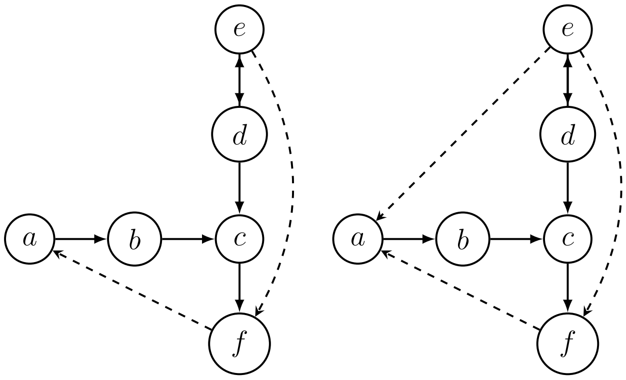


Figure 10.21: σ_5^x do not satisfy P4.1

□

Proposition 10.73. *σ_6^x and σ_7^x satisfy P4.1.*

Proof. It is obvious from Definition 8, that each indirect attack is the result of a **sequence of supports** and an attack. Hence, for all arguments a, b, c such that a supports b and b supports c, applying the deductive or necessary reduction to BAF will lead to the same result as having a supports b, b supports c and a supports c and applying the deductive or necessary reduction to BAF. □

Proposition 10.74. *σ_0^x satisfy Principles 4.2, 4.3, 4.4, 4.7, 4.8.*

Proof. Since the notions of conflict-freeness and defence related to σ_0^x do not take into account the supporting arguments, for each $x \in \{c, g, p, s\}$, we have

$$\sigma_0^x(Ar, att, sup) = \sigma_0^x(Ar, att, \varnothing).$$

Satisfaction of the aforementioned principles follows from this observation. □

Proposition 10.75. *σ_0^x, σ_1^x, σ_2^x, σ_3^x, σ_4^x, σ_5^x and σ_7^x do not satisfy P4.5.*

Proof. We use a counterexample from Figure 10.22. There is a unique complete/grounded /preferred/stable extension, $E = \{a, c\}$. So even if $(a, b) \in sup$, we have $b \notin E$. □

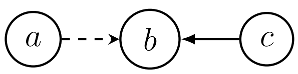


Figure 10.22: σ_0^x, σ_1^x, σ_2^x, σ_3^x, σ_4^x, σ_5^x and σ_7^x do not satisfy P4.5

Proposition 10.76. *σ_0^x, σ_1^x, σ_2^x, σ_3^x, σ_4^x, σ_5^x and σ_6^x do not satisfy P4.6.*

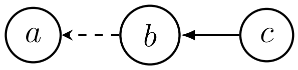


Figure 10.23: σ_0^x, σ_1^x, σ_2^x, σ_3^x, σ_4^x, σ_5^x and σ_6^x do not satisfy P4.6

Proof. We use a counterexample from Figure 10.23. There is a unique complete/grounded/preferred/stable extension, $E = \{a, c\}$. So even if $(a, b) \in sup$, we have $b \notin E$. □

Proposition 10.77. *σ_1^x do not satisfy P4.2.*

Proof. We use a counterexample from Figure 10.24. We have that $\sigma_1^x(Ar, att, sup) = \{\{a, c\}\}$. If we do not consider that c supports a, we have $\sigma_1^{c,p,g}(Ar, att, \varnothing) = \{\{c\}\}$, $\sigma_1^s(Ar, att, \varnothing) = \varnothing$. Then $\sigma_1^x(Ar, att, sup) \notin \sigma_1^x(Ar, att, \varnothing)$. □

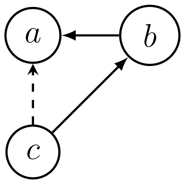


Figure 10.24: σ_1^x do not satisfy P4.2

Proposition 10.78. *σ_1^x do not satisfy P4.3.*

Proof. We use a counterexample from Figure 10.24, where we have the initial BAF $\mathscr{F}$, $\sigma_1^x(\mathscr{F}) = \{\{a, c\}\}$. If we remove the support from c to a, we have a new BAF $\mathscr{F}'$ in Figure 10.25, and we have $\sigma_1^x(\mathscr{F}') = \{\{c\}\}$. □

Proposition 10.79. *σ_1^x satisfy P4.4.*

Proof. For all BAFs $\mathscr{F} = \langle Ar, att, sup\rangle$, for all extensions $E \in \sigma_1^x$, if argument a is not in E, then the support from a has no influence on other arguments. Thus, we have $E \in \sigma_1^x(Ar, att, sup \setminus \{(a, b)\})$. □

Proposition 10.80. *σ_1^x do not satisfy P4.7.*

Proof. We use a counterexample from Figure 10.25. We have $\sigma_1^{g,c,p}(Ar, att, \varnothing) = \{\{c\}\}$, $\sigma_1^s(Ar, att, \varnothing) = \varnothing$. It is the case that $\sigma_0^x(Ar, att, \varnothing) = \{\{a, c\}\}$. It follows that $\sigma_1^x(Ar, att, \varnothing) \neq \sigma_0^x(Ar, att, \varnothing)$. □

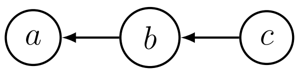


Figure 10.25: σ_1^x do not satisfy P4.7

Proposition 10.81. *σ_1^x satisfy P4.8.*

Proof. From Definition 3, for all BAFs $\mathscr{F} = \langle Ar, att, sup\rangle$, for all arguments $a \in Ar$, if we add the support relation from argument $b \in Ar$ to a, there will not be any new attacks towards a, thus, the status of a will not diminish. □

Proposition 10.82. *σ_2^x do not satisfy P4.2.*

Proof. We use a counterexample from Figure 10.26. We have that $\sigma_2^c(Ar, att, sup) = \{\{a, c, d\}\}$, but if we do not take support into account, $\sigma_2^c(Ar, att, \varnothing) = \{\{c, d\}\}$. Thus, $\sigma_2^c(Ar, att, sup) \notin \sigma_2^c(Ar, att, \varnothing)$. □

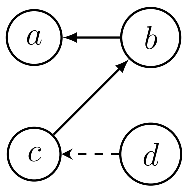


Figure 10.26: σ_2^x do not satisfy P4.2

Proposition 10.83. *σ_2^x do not satisfy P4.3.*

Proof. We use a counterexample from Figure 10.27. For the initial BAF $\mathscr{F}$ on the left, we have $\sigma_2^x(\mathscr{F}) = \{a, c, d\}$, while in the new BAF $\mathscr{F}'$ on the right, we remove the support from d to c and we have $\sigma_2^{c,p,g}(\mathscr{F}') = \{\{c, d\}\}$, $\sigma_2^s(\mathscr{F}') = \varnothing$. □

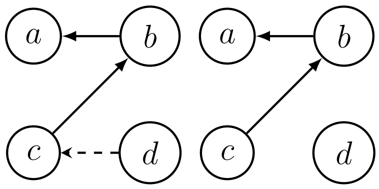


Figure 10.27: σ_2^x do not satisfy P4.3

Proposition 10.84. *σ_2^x satisfy P4.4.*

Proof. For all BAFs $\mathscr{F} = \langle Ar, att, sup \rangle$, for all extensions $E \in \sigma_2^x$, if argument a is not in E, then the support from a has no influence on other arguments. Thus, we have $E \in \sigma_2^x(Ar, att, sup \setminus \{(a, b)\})$. □

Proposition 10.85. *σ_2^x do not satisfy P4.7.*

Proof. We use a counterexample from Figure 10.28. We have that $\sigma_2^c(Ar, att, \varnothing) = \{\{c, d\}\}$, $\sigma_0^c(Ar, att, \varnothing) = \{\{a, c, d\}\}$. Thus, $\sigma_2^c(Ar, att, \varnothing) \neq \sigma_0^c(Ar, att, \varnothing)$. □

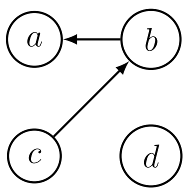


Figure 10.28: σ_2^x do not satisfy P4.7

Proposition 10.86. *σ_2^x satisfy P4.8.*

Proof. Obviously. □

Proposition 10.87. *σ_3^x do not satisfy P4.2.*

Proof. We use a counterexample from Figure 10.29. We have that $\sigma_3^c(Ar, att, sup) = \{\{c, d\}\}$. If we do not take support into account, we have that $\sigma_3^c(Ar, att, \varnothing) = \{\{a, c, d\}\}$. Then, $\sigma_3^c(Ar, att, sup) \notin \sigma_3^c(Ar, att, \varnothing)$. □

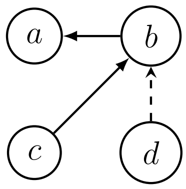


Figure 10.29: σ_3^x do not satisfy P4.2

Proposition 10.88. *σ_3^x satisfy P4.3.*

Proof. Obviously. □

Proposition 10.89. *σ_3^x satisfy P4.4.*

Proof. For all BAFs $\mathscr{F} = \langle Ar, att, sup \rangle$, for all extensions $E \in \sigma_3^x$, if argument a is not in E, then the support from a has no influence on other arguments. Thus, we have $E \in \sigma_3^x(Ar, att, sup \setminus \{(a, b)\})$. □

Proposition 10.90. *σ_3^x satisfy P4.7.*

Proof. For any BAF $\mathscr{F} = \langle Ar, att, \varnothing \rangle$, since there are no supporting arguments, the attacking defence definition becomes the same as Dung's defence definition, and so for every extension $E \in \sigma_3^x(\mathscr{F})$, $\sigma_3^x(Ar, att, \varnothing) = \sigma_0^x(Ar, att, \varnothing)$. □

Proposition 10.91. *σ_3^x satisfy P4.8.*

Proof. Obviously. □

Proposition 10.92. *σ_4^x satisfy Principles 4.2, 4.7, 4.8.*

Proof. Obviously from Definition 7 and Proposition 10.74. □

Proposition 10.93. *σ_4^x do not satisfy P4.3.*

Proof. We use a counterexample from Figure 10.30. We have the initial BAF $\mathscr{F}$ on the left, $\sigma_4^x(\mathscr{F}) = \{\{a.b.d\}\}$. However, if we remove the support from b to a and the support from b to d, we have a new BAF $\mathscr{F}'$ on the right, which is $\sigma_4^x(\mathscr{F}') = \{\{b, c\}\}$. □

Proposition 10.94. *σ_4^x satisfy P4.4.*

Proof. For all BAFs $\mathscr{F}$, for all extensions $E \in \sigma_4^x(\mathscr{F})$ that are selected from $\sigma_0^x(\mathscr{F})$ based on the number of support they receive from arguments not in E, this has nothing to do with the support inside E. □

Proposition 10.95. *σ_5^x satisfy Principles 4.2, 4.7, 4.8.*

Proof. Obviously from Definition 7 and Proposition 10.74. □

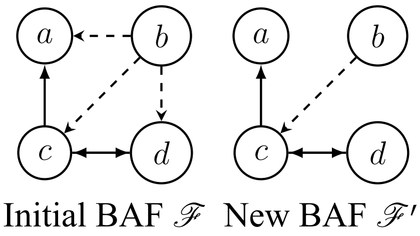


Figure 10.30: σ_4^x do not satisfy P4.3.

Proposition 10.96. *σ_5^x satisfy P4.3.*

Proof. For all BAFs $\mathscr{F}$, for all extensions $E \in \sigma_5^x(\mathscr{F})$ that are selected from $\sigma_0^x(\mathscr{F})$ based on the number of support E receives from arguments outside E, this has nothing to do with the support inside E. □

Proposition 10.97. *σ_5^x do not satisfy P4.4.*

Proof. We use a counterexample from Figure 10.31. We have the initial BAF $\mathscr{F}$ on the left, which is $\sigma_5^x(\mathscr{F}) = \{\{a, b, d\}\}$. However, if we remove the support from c to b and the support from c to a, we have a new BAF $\mathscr{F}'$, which is $\sigma_5^x(\mathscr{F}') = \{\{b, c\}\}$. □

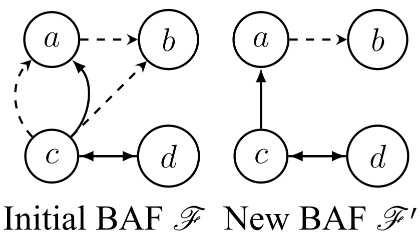


Figure 10.31: σ_5^x do not satisfy P4.4.

Proposition 10.98. *σ_6^x do not satisfy P4.2.*

Proof. We use a counterexample from Figure 10.32. We have the initial BAF $\mathscr{F}$ on the left, and after deductive reduction, we have the new framework $RD(\mathscr{F})$ on the right, which is $\sigma_0^x(RD(\mathscr{F})) = \{\varnothing\}$. However, if we do not take support into account, we have a unique complete/grounded/preferred/stable extension w.r.t. $\sigma_6^x(Ar, att, \varnothing) = \{\{a\}\}, \varnothing \notin \{\{a\}\}$. □

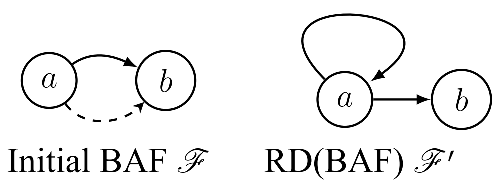


Figure 10.32: σ_6^x do not satisfy P4.2.

Proposition 10.99. *σ_6^x satisfy P4.3.*

Proof. For all BAFs, for all extensions $E \in \sigma_6^x(\mathscr{F})$, for all arguments a and b such that $a, b \in E$ and a supports b, if we remove the support from a to b, for all the arguments c attacked by b, we need to remove the supported attack from a to c, but c is still not accepted because of b. For all the arguments c attacking b, we need to remove mediated attacks from c to a, a is still accepted by E, nothing will change by removing the support from a to b, thus, E is still an extension. □

Proposition 10.100. *σ_6^x satisfy P4.4.*

Proof. For all BAFs $\mathscr{F}$, for all extensions $E \in \sigma_6^x(\mathscr{F})$, for all arguments a and b such that a supports b, if $a \notin E$ and $b \in E$, if we remove the support from a to b, for all arguments c attacked by b, we need to remove the supported attack from a to c, since a is not accepted, there is an argument d attacking a without being attacked, so there is no change to E. The other situation is if b is attacked by an argument c. Since b is accepted, there is an argument $d \in E$ attacking c, and a is also not accepted, so the removal of a mediated attack from c to a will not influence E. □

Proposition 10.101. *σ_6^x satisfy P4.5.*

Proof. We use proof by contradiction. Let $\langle Ar, att, sup\rangle$ be a BAF, let there be an extension $E \in \sigma_6^c$. Assume that σ_6^c does not satisfy P4.5, such that $\exists a \in E$, $b \in Ar\backslash E$, such that $(a, b) \in sup$. As $b \notin E$, such that $\exists c \in Ar$, $(c, b) \in att$, but $\nexists d \in E$ such that d defends b, i.e. d attacks c. If c attacks b, then c mediated-attacks a, there is no d attacks c, and then $E \notin \sigma_6^c$. This is a contradiction. □

Proposition 10.102. *σ_6^x satisfy P4.7.*

Proof. From Definition 9, it is obvious that $\sigma_6^x(Ar, att, \varnothing) = \sigma_0^x(Ar, att, \varnothing)$. □

Proposition 10.103. *σ_6^x satisfy P4.8.*

Proof. Obviously. □

Proposition 10.104. *σ_6^x do not satisfy P4.9.*

Proof. We use a counterexample from Figure 10.33. We have the initial BAF $\mathscr{F}$ on the left, which is $\sigma_6^x = \{\{a, c\}\}$, $Sk(\mathscr{F}) = \{a, c\}$. However, if we add the support from c to b such that we have a new BAF $\mathscr{F}'$, then $\sigma_6^x(\mathscr{F}') = \{\{a\}\}$, $Sk(\mathscr{F}') = \{a\}$. Thus, $Sk(\mathscr{F}) \nsubseteq Sk(\mathscr{F}')$.

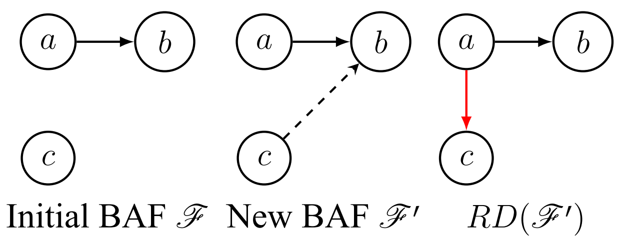


Figure 10.33: σ_6^x do not satisfy P4.9

□

Proposition 10.105. *σ_7^x do not satisfy P4.2.*

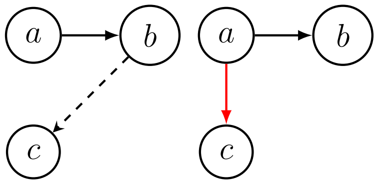


Figure 10.34: σ_7^x do not satisfy P4.2.

Proof. We use a counterexample from Figure 10.34, $\sigma_7^x(Ar, att, sup) = \{\{a\}$. However, if we do not take support into account, we have $\sigma_7^x(Ar, att, \emptyset) = \{a, c\}$ $\sigma_7^c(Ar, att, sup) \notin \sigma_7^c(Ar, att, \emptyset)$. ☐

Proposition 10.106. *σ_7^x satisfy P4.3.*

Proof. For all BAFs $\mathscr{F}$, for all extensions $E \in \sigma_7^x(\mathscr{F})$, for all arguments $a, b \in E$, for all the arguments c attacking a, b, there exists an argument $d \in E$ attacking c. If we remove the support from a to b, only the secondary attacks on b will be removed, or those extended attacks from b to other arguments will be removed that are still attacked by a. Neither will not change E. ☐

Proposition 10.107. *σ_7^x satisfy P4.4.*

Proof. For all BAFs $\mathscr{F}$, for all extensions $E \in \sigma_7^x(\mathscr{F})$, there are no arguments $a, b \in Ar$ such that $a \in E$ and $b \notin E$. ☐

Proposition 10.108. *σ_7^x do not satisfy P4.5.*

Proof. We use a counterexample from Figure 10.22. There is a unique complete/grounded/preferred/stable extension, $E = \{a, c\}$. So even if $(a, b) \in sup$, we have $b \notin E$. ☐

Proposition 10.109. *σ_7^x satisfy P4.6.*

Proof. We use proof by contradiction. Let $\langle Ar, att, sup\rangle$ be a BAF, let there be an extension $E \in \sigma_7^c$. Assume that σ_7^c does not satisfy P4.6, such that $\exists\, a \in E$ and $b \in Ar \setminus E$, such that $(b, a) \in sup$. As $b \notin E$, such that $\exists\, c \in Ar$ and $(c, b) \in att$, but $\nexists\, d \in E$ such that d defends b, i.e. d attacks c. If c attacks b, then c secondary attacks a and there is no d that attacks c, and so $E \notin \sigma_7^c$. This is a contradiction. ☐

Proposition 10.110. *σ_0^s does not satisfy P4.10.*

Proof. In [35], the authors have showed that σ_0^s violates directionality, so σ_0^s violates also BAF directionality. We use their counterexample from Figure 10.35 Let $U = \{a, b\} \in US(BAF)$, $\sigma_0^s(BAF_{\downarrow U}) = \{\{a\}, \{b\}\}$, $\sigma_0^s(BAF) = \{\{b\}\} = \{E \cap U | E \in \sigma_0^s(BAF)\}$. $\sigma_0^s(BAF_{\downarrow U}) \neq \{E \cap U | E \in \sigma_0^s(BAF)\}$

☐

Proposition 10.111. *σ_1^s does not satisfy P4.10.*

Proof. We use a counterexample from Figure 10.36 Let $U = \{a, b\} \in US(BAF)$, $\sigma_1^s(BAF_{\downarrow U}) = \{\{a\}, \{b\}\}$, $\sigma_1^s(BAF) = \{\{b\}\} = \{E \cap U | E \in \sigma_1^s(BAF)\}$. $\sigma_1^s(BAF_{\downarrow U}) \neq \{E \cap U | E \in \sigma_1^s(BAF)\}$

☐

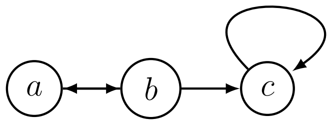


Figure 10.35: σ_0^s does not satisfy P4.10

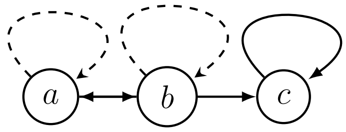


Figure 10.36: σ_1^s does not satisfy P4.10

Proposition 10.112. *σ_2^s does not satisfy P4.10.*

Proof. We use a counterexample from Figure 10.37 Let $U = \{a, b, d\} \in US(BAF)$, $\sigma_2^s(BAF_{\downarrow U}) = \{\{a, d\}, \{b, d\}\}$, $\sigma_2^s(BAF) = \{\{b, d\}\} = \{E \cap U | E \in \sigma_2^s(BAF)\}$. $\sigma_2^s(BAF_{\downarrow U}) \neq \{E \cap U | E \in \sigma_2^s(BAF)\}$

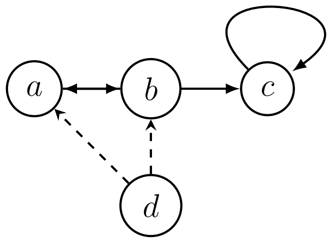


Figure 10.37: σ_2^s does not satisfy P4.10

□

Proposition 10.113. *σ_3^s does not satisfy P4.10.*

Proof. We use the same proof of proposition 10.110 □

Lemma 10.14. *If a set of arguments $E' \subseteq U$ is admissible$_i$ in $BAF_{\downarrow U}$, then E' is admissible$_i$ in BAF, for $i \in \{0, 1, 2, 3\}$.*

Proof. E' is admissible$_i$ in $BAF_{\downarrow U}$, E' can not be attacked by any argument in $Ar \setminus U$, so E' defends$_i$ itself in BAF. □

Proposition 10.114. *σ_i^p satisfy P4.10, for $i \in \{0, 1\}$.*

Proof. The authors in [35], have proved that σ_0^p satisfies P4.10. We only give the proof for σ_1^p.

We need to prove that $\sigma_1^p\ (BAF_{\downarrow U}) = \{E \cap U | E \in \sigma_1^p(BAF)\}$.

- **Part 1: Prove that** $\{E \cap U | E \in \sigma_1^p(BAF)\} \subseteq \sigma_1^p\ (BAF_{\downarrow U})$. Suppose that $E' \in \{E \cap U | E \in \sigma_1^p(BAF)\}$, we need to prove that $E' \in \sigma_1^p\ (BAF_{\downarrow U})$. $E' \in \{E \cap U | E \in \sigma_1^p(BAF)\}$, that means that there exists $E \in \sigma_1^p(BAF)$ such that $E' = E \cap U$.

1. E' is conflict-free$_1$ because E is conflict-free$_1$.
2. We prove that E' is admissible$_1$ in $BAF_{\downarrow_U}$. For each argument $y \in E'$, y can be unattacked, then y is defended$_1$ by E'. If y is attacked, then there exists $x \in U \setminus E'$ such that x attacks y. Since E is admissible$_1$ in BAF, that means that there exists $z \in E$ such that z attacks x and z supports y. z must belong to U, so $z \in (E \cap U = E')$, so E' is admissible$_1$ in $BAF_{\downarrow_U}$.
3. We prove that E' is a maximal admissible$_1$ set in $BAF_{\downarrow_U}$ by contradiction. Suppose there exists E_2, such that $E' \subset E_2$ and $E_2 \subseteq U$ and E_2 is admissible$_1$ in $BAF_{\downarrow_U}$. E_2 defends$_1$ all its arguments in $BAF_{\downarrow_U}$. Define $E_3 = E2 \cup E$. If we prove that E_3 is admissible$_1$ in BAF, that means that E is not a preferred$_i$ extension in BAF which is a contradiction, so we prove that E' is a preferred$_i$ extension of $BAF_{\downarrow_U}$.
   (a) E_3 is conflict-free$_1$ because:
      i. E_2 is conflict-free$_1$ because E_2 is admissible$_1$ in $BAF_{\downarrow_U}$, also E is conflict-free$_1$ because $E \in \sigma_1^p(BAF)$.
      ii. E can not attack E_2 because $E_2 \subseteq U$.
      iii. To prove that $E_2 \setminus E'$ cannot attack $E \setminus E'$, we do it by contradiction. Suppose that there exists $x \in E_2 \setminus E'$ that attacks an argument $y \in E \setminus E'$, since E is admissible$_1$ in BAF, there exists $z \in E'$ such that z attacks x and z supports y. If z attacks x, then E_2 is not conflict-free$_1$. Contradiction.
   (b) E_3 is admissible$_1$ in BAF, because E is admissible$_1$ in BAF, and from Lemma 10.14, if E_2 is admissible$_1$ in $BAF_{\downarrow_U}$, then E_2 is admissible$_1$ in BAF.

- **Part 2: Prove that σ_1^p $(BAF_{\downarrow_U}) \subseteq \{E \cap U | E \in \sigma_1^p(BAF)\}$.** Suppose that $E' \in \sigma_1^p$ $(BAF_{\downarrow_U})$, we need to prove that $E' \in \{E \cap U | E \in \sigma_1^p(BAF)\}$, so we need to prove that there exists $E \in \sigma_1^p(BAF)\}$ such that $E' = E \cap U$. E' is admissible$_1$ in $(BAF_{\downarrow_U})$, from Lemma 10.14, E' is admissible$_1$ in BAF. That means that there exists $E_2 \in \sigma_1^p(BAF)\}$ such that $E' \subseteq E_2$ and $E_2 \cap U = E'$. $E' \subseteq E_2$ because we can always have a preferred$_i$ extension that contains E'. $E_2 \cap U \not\supset E'$ because then E' would not be a maximal admissible$_1$ set in $(BAF_{\downarrow_U})$.

□

Proposition 10.115. *σ_2^p satisfies P4.10.*

Proof. The proof is the same as the proof of proposition 10.114, except for **Part 1 (2) and Part 1 (3) (a) iii**. We will only prove these two parts.

- **Part 1 (2):** We prove that E' is admissible$_2$ in $BAF_{\downarrow_U}$. For each argument $y \in E'$, y can be unattacked, then y is defended$_2$ by E'. If y is attacked, then there exists $x \in U \setminus E'$ such that x attacks y. Since E is admissible$_2$ in BAF, that means that there exists $z, w \in E$ such that z attacks x and w supports z. z must belong to U, so $z \in (E \cap U = E')$, and z cannot be supported from outside of U so $w \in E'$ so E' is admissible$_2$ in $BAF_{\downarrow_U}$.

- **Part 1 (3) (a) iii:** To prove that $E_2 \setminus E'$ cannot attack $E \setminus E'$, we do it by contradiction. Suppose that there exists $x \in E_2 \setminus E'$ that attacks an argument $y \in E \setminus E'$, since E is admissible$_2$ in BAF, there exists $z, w \in E'$ such that z attacks x and w supports z. If z attacks x, then E_2 is not conflict-free$_2$. Contradiction.

□

Proposition 10.116. *σ_3^p satisfies P4.10.*

Proof. The proof is the same as the proof of proposition 10.114, except for **Part 1 (2) and Part 1 (3) (a) iii**. We will only prove these two parts.

- **Part 1 (2):** We prove that E' is admissible$_3$ in $BAF_{\downarrow U}$. For each argument $y \in E'$, y can be unattacked, then y is defended$_3$ by E'. If y is attacked, then there exists $x \in U \setminus E'$ such that x attacks y. Since E is admissible$_3$ in BAF, that means that there exists $z, w \in E$ such that z attacks x and for any supporter b of x, w attacks b. z must belong to U, so $z \in E \cap U = E'$, $b \in U$, so $w \in U$, so $w \in E'$ so E' is admissible$_3$ in $BAF_{\downarrow U}$.
- **Part 1 (3) (a) iii:** To prove that $E_2 \setminus E'$ cannot attack $E \setminus E'$, we do it by contradiction. Suppose that there exists $x \in E_2 \setminus E'$ that attacks an argument $y \in E \setminus E'$, since E is admissible$_3$ in BAF, there exists $z, w \in E'$ such that z attacks x and for any supporter b of x, w attacks b. If z attacks x, then E_2 is not conflict-free$_3$. Contradiction.

□

Lemma 10.15. *If a set of arguments $E \subseteq Ar$ is admissible$_i$ in BAF, then the set of arguments $E' = (E \cap U)$ is admissible$_i$ in $BAF_{\downarrow U}$, for $i \in \{0, 1, 2, 3\}$.*

Proof. We find the proofs in Parts 1 (2) of proofs of propositions 10.114, 10.115 and 10.116.

□

Proposition 10.117. *σ_i^c satisfy P4.10, for $i \in \{0, 1\}$.*

Proof.
- **Part 1:** Suppose that $E' = E \cap U$ such that $E \in \sigma_1^c(BAF)$, we need to prove that $E' \in \sigma_1^c(BAF_{\downarrow U})$.
  - From Lemma 10.15, if E is admissible$_1$ in BAF, then E' is admissible$_1$ in $BAF_{\downarrow U}$.
  - We need to prove that there does not exist $x \in U \setminus E'$ such that E' defends$_1$ x. We prove this by contradiction. Suppose that E' defends$_1$ arguments in $U \setminus E'$. We have two cases.
    1. Case 1: Suppose that there exists $x \in U \setminus E'$ such that x is not attacked. Then x is defended$_1$ by E which is not possible because E is a complete$_i$ extension in BAF.
    2. Case 2: Suppose that there exists $x, y \in U \setminus E'$, such that y attacks x. Suppose that E' defends$_1$ x, so there exists $z \in E'$ such that z attacks y, and z supports x. Since $z \in E'$, so $z \in E$, so E defends$_1$ an argument outside of E, but $E \in \sigma_1^c(BAF)$. Contradiction.

- **Part 2:** Suppose that $E' \in \sigma_1^c(BAF_{\downarrow U})$, we need to prove that there exists $E \in \sigma_1^c(BAF)$ such that $E' = E \cap U$.
  For each set of arguments S in BAF, we denote by $\text{Def}_i(S)$ the set of arguments defended$_i$ by S. Let $d_1(\mathscr{F}, E') = E_1$, $d_1(\mathscr{F}, E_1) = E_2$, ..., $d_1(\mathscr{F}, E_{i-1}) = E_i$. $(E_1 \cap U) = E'$ because $E' \in \sigma_1^c(BAF_{\downarrow U})$, which means that E' cannot defend$_1$ arguments in $(U \setminus E')$. Thus, $(E_1 \setminus E') \subseteq (Ar \setminus U)$.
  Also, for each i, $E_i = d_1(\mathscr{F}, E_{i-1})$, E_{i-1} cannot defend$_1$ the arguments in U, so $(E_i \setminus E') \subseteq (Ar \setminus U)$. That is because if U is attacked, it is going to be attacked from the arguments in U, U cannot be defended$_1$ from the arguments outside of U.
  Let us now prove by induction, that for each i, $E_i = d_1(\mathscr{F}, E_{i-1})$ is admissible$_1$ in BAF.

  - **Base:** Let us prove that E' is admissible$_1$ in BAF.
    From Lemma 10.14, if E' is admissible$_1$ in $BAF_{\downarrow U}$, then E' is admissible$_1$ in BAF.
  - **Step:** Let us suppose that it holds that E_i is admissible$_1$ in BAF, and let us prove that $E_{i+1} = d_1(\mathscr{F}, E_i)$ is admissible$_1$ in BAF.

    1. We prove that E_{i+1} is conflict-free$_1$.
       * E_i is conflict-free$_1$ because E_i is admissible$_1$ in BAF.
       * E_i cannot attack $(E_{i+1} \setminus E_i)$, because since E_i defends$_1$ E_{i+1}, if an arguments $x \in E_i$ attacks E_{i+1}, there exists an argument $y \in E_i$ attacking x, which is not possible since E_i is conflict-free$_1$.
       * $(E_{i+1} \setminus E_i)$ is conflict-free$_1$ because if an argument $x \in (E_{i+1} \setminus E_i)$ attacks an argument $y \in (E_{i+1} \setminus E_i)$, E_i must attack x, which is not possible because E_i defends$_1$ E_{i+1}.
       * $(E_{i+1} \setminus E_i)$ cannot attack E_i because since E_i is admissible$_1$ in BAF, if $(E_{i+1} \setminus E_i)$ attacks E_i, E_i defends$_1$ itself so E_i must attack $(E_{i+1} \setminus E_i)$ which is not possible.
    2. We prove that E_{i+1} is admissible$_1$ in BAF.
       * E_i is admissible$_1$ in BAF.
       * E_i defends$_1$ E_{i+1} in BAF.

  Hence, by induction, we conclude that for each i, $E_i = d_1(\mathscr{F}, E_{i-1})$ is admissible$_1$ in BAF.

  Having that at each step i, $d_1(\mathscr{F}, E_i)$ is admissible$_1$, also having a finite number of arguments in Ar, we conclude that there is a fix point of the above defined sequence, that is, there exists j such that $d_1(\mathscr{F}, E_j) = d_1(\mathscr{F}, d_1(\mathscr{F}, d_1(\mathscr{F}, d_1(\mathscr{F}, ..., d_1(\mathscr{F}, E'))))) = E_j$. Hence, E_j is a complete$_i$ extension in BAF such that $E_j \cap U = E'$.

□

Proposition 10.118. *σ_2^c satisfies P4.10.*

Proof. The proof is the same as the proof of proposition 10.117 except for **Part 1 (2)**, we will only prove this part.

- **Part 1 (2)**: Suppose that there exist $x, y \in U \setminus E'$, such that y attacks x. Suppose that E' defends$_2$ x, so there exist $z, w \in E'$ such that z attacks y, and w supports z. Since $z, w \in E'$, so $z, w \in E$, so E defends$_2$ an argument outside of E, but $E \in \sigma_2^c(BAF)$. Contradiction.

□

Proposition 10.119. *σ_3^c satisfies P4.10.*

Proof. The proof is the same as the proof of proposition 10.117 except for **Part 1 (2)**, we will only prove this part.

- **Part 1 (2):** Suppose that there exist $x, y \in U \setminus E'$, such that y attacks x. Suppose that E' defends$_3$ x, so there exist $z, w \in E'$ such that z attacks y, and for any supporter b of y, w attacks b. Since $z, w \in E'$, so $z, w \in E$, so E defends$_3$ an argument outside of E, but $E \in \sigma_3^c(BAF)$. Contradiction.

□

Lemma 10.16. *For any set of arguments $E \subseteq Ar$, if $E \in \sigma_i^c(BAF)$, then $E' = (E \cap U) \in \sigma_i^c(BAF_{\downarrow U})$, for $i \in \{0,1,2,3\}$.*

Proof. We find the proof in Parts 1 of proofs of propositions 10.117, 10.118, and 10.119. □

Lemma 10.17. *For any set of arguments $E' \subseteq U$, if $E' \in \sigma_i^c(BAF_{\downarrow U})$, for $i \in \{0,1,2,3\}$, then there exists a fix point E_j of the sequence $d_i(\mathscr{F}, E_j) = d_i(\mathscr{F}, d_i(\mathscr{F}, d_i(\mathscr{F}, d_i(\mathscr{F}, ..., d_i(\mathscr{F}, E'))))) = E_j$ where $E_j \in \sigma_i^c(BAF)$ such that $E_j \cap U = E'$.*

Proof. We find the proof in Part 2 of the proof of proposition 10.117. □

Proposition 10.120. *σ_i^g satisfy P4.10, for $i \in \{0,1,2,3\}$.*

Proof. The authors in [35], have proved that σ_0^g satisfies P4.10. The proof we will give below, is the same for each σ_i^g, with $i \in \{0,1,2,3\}$.

- **Part 1:** Suppose that E is the grounded$_i$ extension of BAF, let $E' = E \cap U$, we prove by contradiction that E' is the grounded$_i$ extension of $BAF_{\downarrow U}$.
  From Lemma 10.16, if $E \in \sigma_i^c(BAF)$, then $E' \in \sigma_i^c(BAF_{\downarrow U})$ such that $E' = E \cap U$.
  Suppose there exists $E_1 \subset E'$such that $E_1 \in \sigma_i^c(BAF_{\downarrow U})$.
  Let $E_2 = d_i(\mathscr{F}, E_1)$ be the set of arguments defended$_i$ by E_1, $E_3 = d_i(\mathscr{F}, E_2)$, ..., $E_u = d_i(\mathscr{F}, E_{u-1}) = d_i(\mathscr{F}, d_i(\mathscr{F}, ..., d_i(\mathscr{F}, E_1)))$. Let us prove now by induction that, for each u, $(E_u = d_i(\mathscr{F}, E_{u-1})) \subset E$.
  - **Base:** $E_1 \subset E'$, so $E_1 \subset E$.
  - **Step:** Let us suppose that it holds that $E_{u-1} \subset E$, and let us prove that $(E_u = d_i(\mathscr{F}, E_{u-1})) \subset E$. Since $E \in \sigma_i^c(BAF)$, there does not exist any argument x in E, such that x defends$_i$ an argument in $(Ar \setminus E)$. Hence, $E_{u-1} \subset E$, so E_{u-1} cannot defend$_i$ arguments in $(Ar \setminus E)$, so $(E_u = d_i(, \mathscr{F}, E_{u-1})) \subset E$.

From Lemma 10.17, if $E_1 \in \sigma_i^c(BAF_{\downarrow_U})$, for $i \in \{0,1,2,3\}$, then there exists a fix point E_j of the sequence $d_i(\mathscr{F}, E_j) = d_i(\mathscr{F}, d_i(\mathscr{F}, d_i(\mathscr{F}, d_i(\mathscr{F}, ..., d_i(\mathscr{F}, E'))))) = E_j$ where $E_j \in \sigma_i^c(BAF)$ such that $E_j \cap U = E_1$. We have proved that $E_j \subset E$. Also, $E_j \subsetneq E$ because $E_1 \subsetneq E'$. But E is the grounded$_i$ extension of BAF. Contradiction.

- **Part 2:** Suppose E' is the grounded$_i$ extension of $BAF_{\downarrow_U}$, we need to prove that $E' = E \cap U$ such that E is the grounded$_i$ extension of BAF. From Lemma 10.17, if $E' \in \sigma_i^c(BAF_{\downarrow_U})$, for $i \in \{0,1,2,3\}$, then there exists $E \in \sigma_i^c(BAF)$ such that $E' = E \cap U$ and E is the fix point obtained by successively applying the function Def$_i$, so $E = d_i(\mathscr{F}, d_i(\mathscr{F}, d_i...(\mathscr{F}, d_i(E')))) = d_i(\mathscr{F}, E)$. We want to prove that E is the grounded$_i$ extension of BAF by contradiction. Suppose that there exists $E_1 \subset E$ such that E_1 is the grounded$_i$ extension of BAF. Then, from **Part 1** of the proof of proposition 10.120, $E_1' = E_1 \cap U$ is the grounded$_i$ extension of $BAF_{\downarrow_U}$. But E' is the grounded$_i$ extension of $BAF_{\downarrow_U}$. Contradiction. So E is the grounded$_i$ extension of BAF.

□

Proposition 10.121. *$\sigma_4^{c,p,s}$ do not satisfy P4.10.*

Proof. We use a counterexample from Figure 10.38 Let $U = \{a, b\} \in US(BAF)$, $\sigma_4^{c,p,s}(BAF_{\downarrow_U}) = \{\{a\},\{b\}\}$, $\sigma_4^{c,p,s}(BAF) = \{\{a,c\}\}.\{E \cap U | E \in \sigma_4^{c,p,s}(BAF)\} = \{\{a\}\}$. $\sigma_4^{c,p,s}(BAF_{\downarrow_U}) \neq \{E \cap U | E \in \sigma_4^{c,p,s}(BAF)\}$.

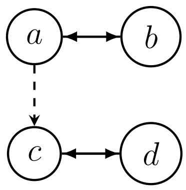


Figure 10.38: $\sigma_{4,5}^{c,p,s}$ do not satisfy P4.10

□

Proposition 10.122. *$\sigma_5^{c,p,s}$ do not satisfy P4.10.*

Proof. We use a counterexample from Figure 10.38 Let $U = \{a, b\} \in US(BAF)$, $\sigma_5^{c,p,s}(BAF_{\downarrow_U}) = \{\{a\},\{b\}\}$, $\sigma_5^c(BAF) = \{\{c\},\{b,c\}\}$, $\sigma_5^{p,s}(BAF) = \{\{b,c\}\}.\{E \cap U | E \in \sigma_5^{c,p,s}(BAF)\} = \{\{b\}\}$. $\sigma_5^{c,p,s}(BAF_{\downarrow_U}) \neq \{E \cap U | E \in \sigma_5^{c,p,s}(BAF)\}$. □

Proposition 10.123. *$\sigma_6^{c,p,g}$ do not satisfy P4.10.*

Proof. We use a counterexample to prove $\sigma_6^{c,p,g}$ does not satisfy P4.10 which is showed in Figure 10.39. Let $U = \{a, b\} \in US(BAF)$ shown in the righthand sub-figure, $\sigma_6^{c,p,g}(BAF_{\downarrow_U}) = \{\{b\}\}$, $\sigma_6^{c,p}(BAF) = \{\{a\},\{b,c\}\}$. $\{E \cap U | E \in \sigma_6^{c,p}(BAF)\} = \{\{a\},\{b\}\}$. $\sigma_6^{c,p}(BAF_{\downarrow_U}) \neq \{E \cap U | E \in \sigma_6^{c,p}(BAF)\}$. $\sigma_6^g(BAF) = \{\varnothing\}$, $\{E \cap U | E \in \sigma_6^g(BAF)\} = \{\varnothing\}$, $\sigma_6^g(BAF_{\downarrow_U}) \neq \{E \cap U | E \in \sigma_6^g(BAF)\}$. □

Proposition 10.124. *σ_7^g satisfy P4.10.*

For the sake of conciseness, to prove Proposition 10.124, we use $\rightarrow$ and $\dashrightarrow$ to represent attack and support relation respectively, i.e. a BAF is $\mathscr{F} = \langle Ar, att, sup \rangle$, a attacks b (rep. a supports b) can written as $a \rightarrow b$ ($a \dashrightarrow b$). Recall that $\sigma_6^x(\mathscr{F}) = \sigma_0^x(RD(\mathscr{F}))$, and $\sigma_7^x(\mathscr{F}) = \sigma_0^x(RN(\mathscr{F}))$.

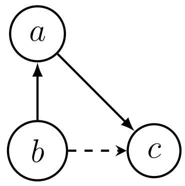


Figure 10.39: $\sigma_6^{c,p,g}$ do not satisfy P4.10

Definition 10.6. *Given a bipolar BAF $\mathscr{F} = (Ar, \rightarrow, \dashrightarrow)$, for every $j \in \mathbb{N}$, we define the set $\rightarrow^j \subseteq Ar \times Ar$ inductively as follows:*

1. $\rightarrow^0 = \rightarrow$;
2. $\rightarrow^{j+1} = \rightarrow^j \cup \{(c,b) \in Ar \times Ar \mid \exists a \in Ar : a \dashrightarrow b \,\&\, c \rightarrow^j a\} \cup \{(b,c) \in Ar \times Ar \mid \exists a \in Ar : a \dashrightarrow b \,\&\, a \rightarrow^j c\}$

Definition 10.7. *Given a bipolar BAF $\mathscr{F} = (Ar, \rightarrow, \dashrightarrow)$, for every $X \subseteq Ar$:*

- *let* $\mathscr{F}\downarrow_X = (X, \rightarrow \cap (X \times X), \dashrightarrow \cap (X \times X))$.
- *for every $a, b \in X$, write $a \rightarrow_X b$ if $(a,b) \in (\mathscr{F}\downarrow_X)^+$.*

Lemma 10.18. *For every $x, y \in U$, $x \rightarrow_{Ar} y$ iff $x \rightarrow_U y$.*

Proof. From left to right: Since $x \rightarrow_{Ar} y$, there must be $i \in \mathbb{N}$ such that $x \rightarrow^i_{Ar} y$. The proof is carried out by induction on the value i. If $x \rightarrow^0_{Ar} y$, then $x \rightarrow y$, thus $x \rightarrow_U y$. Suppose the lemma holds for $i = n$. If $x \rightarrow^{n+1}_{Ar} y$ and $x \not\rightarrow^n_{Ar} y$, then, by Definition 10.6, it must be either

- $\exists a \in Ar : a \dashrightarrow y \,\&\, x \rightarrow^n_{Ar} a$ or
- $\exists a \in Ar : a \dashrightarrow x \,\&\, a \rightarrow^n_{Ar} y$.

In the former case, we have $a \in U$ (by the definition of U) and $x \rightarrow_U a$ (by the IH). Thus $x \rightarrow_U y$ be the definition of $\rightarrow_U$. Similarly for the latter case.

The direction from right to left is trivial. □

Lemma 10.19. *For any $i \in \mathbb{N}$, $x \in d_0^i(\mathscr{F}, \varnothing)$ and $y \in U$, if $x \rightarrow_{Ar} y$, then there must be $z \in U \cap d_0^i(\mathscr{F}, \varnothing)$ such that $z \rightarrow_{Ar} y$.*

Proof. If $x \in U$ then we can set $z = x$. Thus we assume $x \notin U$ in the following proof.

Since $x \rightarrow_{Ar} y$, there must be $j \in \mathbb{N}$ such that $x \rightarrow^j_{Ar} y$. The proof is carried out by induction on the value j.

If $j = 0$, it means that $x \rightarrow y$. Since we assume $x \notin U$, this is impossible. Thus the lemma holds vacuously.

Suppose the lemma holds for $j = n$. Let $x \rightarrow^{n+1}_{Ar} y$ and $x \not\rightarrow^n_{Ar} y$, then, by Definition 10.6, it must be either

- $\exists a \in Ar : a \dashrightarrow y \,\&\, x \rightarrow^n_{Ar} a$ or
- $\exists a \in Ar : a \dashrightarrow x \,\&\, a \rightarrow^n_{Ar} y$.

In the former case, it must be $a \in U$. Since $x \rightarrow^n_{Ar} a$, by IH, we have there must be $z \in U \cap d^i_0(\mathscr{F}, \varnothing)$ such that $z \rightarrow_{Ar} a$. Thus $z \rightarrow_{Ar} y$ by Definition 10.6. In the latter case, note that $a \in d^i_0(\mathscr{F}, \varnothing)$ since $x \in d^i_0(\mathscr{F}, \varnothing)$ and $a \dashrightarrow x$. Since $a \rightarrow^n_{Ar} y$, apply the IH, we have there must be $z \in U \cap d^i_0(\mathscr{F}, \varnothing)$ such that $z \rightarrow_{Ar} y$. □

Lemma 10.20. *If there is $x \in Ar$ and $y \in U$ such that $x \rightarrow_{Ar} y$, then there must be $z \in U$ such that $z \rightarrow_U y$.*

Proof. Since $x \rightarrow_{Ar} y$, there must be $i \in \mathbb{N}$ such that $x \rightarrow^i_{Ar} y$. The proof is carried out by induction on the value i. If $x \rightarrow^0_{Ar} y$, then $x \rightarrow y$. By the definition of U, $x \in U$. Thus we can set $z = x$. Suppose the lemma holds for $i = n$. If $x \rightarrow^{n+1}_{Ar} y$ and $x \not\rightarrow^n_{Ar} y$, then, by Definition 10.6, it must be either

- $\exists a \in Ar : a \multimap y \;\&\; x \rightarrow^n_{Ar} a$ or
- $\exists a \in Ar : a \multimap x \;\&\; a \rightarrow^n_{Ar} y$.

In the former case, it must be $a \in U$ by the definition of U. Since $x \rightarrow^n_{Ar} a$, by the IH, there exists $z \in U$ with $z \rightarrow_U a$. Thus $z \rightarrow_U y$ by the definition of $\rightarrow_U$. In the latter case, since $a \rightarrow^n_{Ar} y$, by IH, there must be a $z \in U$ with $z \rightarrow_U y$. □

We now ready to give the proof of Proposition 10.124 based on Lemma 10.18 to 10.20.

Proof. We first show that

$$\text{for every } i \in \mathbb{N}^+,\ c \in U,\ c \in d^i_0(\mathscr{F}, \varnothing) \text{ iff } c \in d^i_0(U, \varnothing). \tag{10.2}$$

Base: The direction from left to right is straightforward (since $(\mathscr{F}\downarrow_U)^+ \subseteq \mathscr{F}^+$). For the converse, suppose $c \in d^1_0(U, \varnothing)$ and $c \notin d^1_0(\mathscr{F}, \varnothing)$. It follows that there is $a \in Ar$ with $a \rightarrow_{Ar} c$. If $a \in U$, then $a \rightarrow_U c$ by Lemma 10.18, contradiction! If $a \notin U$, there is also $z \in U$ with $z \rightarrow_U c$, contradiction!

IH: For every $c \in U$, $c \in d^n_0(\mathscr{F}, \varnothing)$ iff $c \in d^n_0(U, \varnothing)$.

Induction Step: **From left to right.** Suppose $c \in d^{n+1}_0(\mathscr{F}, \varnothing)$ and let $x \in U$ be an arbitrary argument such that $x \rightarrow_U c$. Then $x \rightarrow_{Ar} c$. Thus there must be $y \in d^n_0(\mathscr{F}, \varnothing)$ such that $y \rightarrow_{Ar} x$. By Lemma 10.19, there must be $y' \in U \cap d^n_0(\mathscr{F}, \varnothing)$ such that $y' \rightarrow_{Ar} x$. Since $y', x \in U$, by Lemma 10.18 we have $y' \rightarrow_U x$. By IH, we also have $y' \in d^n_0(U, \varnothing)$. Since x is arbitrary, we have $c \in d^{n+1}_0(U, \varnothing)$.

From right to left. We need to show that

$$\begin{array}{l}\text{For every } c \in D^{n+1}_U(\varnothing) \text{ and every } x \in Ar,\\ \text{if } x \rightarrow_{Ar} c \text{ then there is } y \in D^n_{Ar}(\varnothing) \text{ such}\\ \text{that } y \rightarrow_{Ar} x\end{array} \tag{10.3}$$

Since $x \rightarrow_{Ar} c$, there must be $j \in \mathbb{N}$ such that $x \rightarrow^j_{Ar} c$. We prove (10.3) by induction on the value j. If $j = 0$, then $x \rightarrow c$, thus it must be that $x \in U$ by the definition of U. Since $c \in d^{n+1}_0(U, \varnothing)$, there must be $y \in d^n_0(U, \varnothing)$ such that $y \rightarrow_U x$. By IH, we have $y \in d^n_0(\mathscr{F}, \varnothing)$. It also holds that $y \rightarrow_{Ar} x$.

Suppose (2) holds for $j = n$, we are going to show the case $j = n+1$. We consider two cases: (a) If $x \in U$, then $x \rightarrow_U c$ by Lemma 10.18. Thus there must be $y \in d^n_0(U, \varnothing)$ such that $y \rightarrow_u x$. Note that $y \rightarrow_{Ar} x$ and $y \in d^n_0(\mathscr{F}, \varnothing)$ (by the IH). (b) If $x \notin U$, then, by Definition 10.6, it must be that either

- $\exists a \in Ar : a \multimap c \,\&\, x \to^{n}_{Ar} a$ or
- $\exists a \in Ar : a \multimap x \,\&\, a \to^{n}_{Ar} c$.

In the former case, it must be that $a \in U$ by the definition of U. Note also that $a \in d_0^{n+1}(U, \varnothing)$ since $c \in d_0^{n+1}(U, \varnothing)$. Thus, together with the fact that $x \to^{n}_{Ar} a$, it implies that there is $y \in d_0^{n}(\mathscr{F}, \varnothing)$ such that $y \to_{Ar} x$ (because we assume that (2) holds for $j = n$). In the latter case, since $a \to^{n}_{Ar} c$, there must be $y \in d_0^{n}(\mathscr{F}, \varnothing)$ such that $y \to_{Ar} a$ (because we assume that (2) holds for $j = n$). Note that $y \to_{Ar} x$ by Definition 10.6.

□

Proposition 10.125. *σ_7^c satisfy P4.10.*

Lemma 10.21. *Let $V \subseteq Ar$ be given. For any $X \in \sigma_7^c(V)$ and any $a, b \in V$ with $a \multimap b$, if $b \in X$, then $a \in X$.*

Proof. Let $x \in V$ be an arbitrary argument such that $x \to_V a$. Then $x \to_V b$ by Definition 10.6. Since $b \in X$ and $X \subseteq d_0(V, X)$, then there must be $y \in X$ such that $y \to_V x$. Since a is arbitrary, we have $a \in d_0(V, X)$. Since $X \supseteq d_0(V, X)$, we have $a \in X$. □

Lemma 10.22. *Let $X \in \sigma_7^c(\mathscr{F})$. For any $y \in X$ and $x \in U$, if $y \to_{Ar} x$, then there is $z \in X \cap U$ such that $z \to_{Ar} x$.*

Proof. Since $y \to_{Ar} x$, there must be $i \in \mathbb{N}$ such that $y \to^{i}_{Ar} x$. The proof is carried out by induction on i. If $i = 0$, it means that $y \to x$, thus $y \in U$ by the definition of U. Since we assume $y \in X$, we have $y \in X \cap U$. Hence we can set $z = y$.

Suppose the lemma holds for $i = n$. If $y \to^{n+1}_{Ar} x$ and $y \not\to^{n}_{Ar} x$, then, by Definition 10.6, it must be either

- $\exists a \in Ar : a \multimap x \,\&\, y \to^{n}_{Ar} a$ or
- $\exists a \in Ar : a \multimap y \,\&\, a \to^{n}_{Ar} x$.

In the former case, since $x \in X$ we have $a \in X$. Since $y \to^{n}_{Ar} a$, apply the IH, we have that there must be $z \in X \cap U$ such that $z \to_{Ar} a$. Note that $a \multimap x$, thus $z \to_{Ar} x$ by Definition 10.6. In the latter case, we first have $a \in X$ since $y \in X$ (by Lemma 10.21). Since $a \to^{n}_{Ar} x$, apply the IH, there must be $z \in X \cap U$ such that $z \to_{Ar} x$.

□

Lemma 10.23. *For every $X \in \sigma_7^c(\mathscr{F})$, $X \cap U \in \sigma_7^c(U)$.*

Proof. By Lemma 10.18, it is easy to know that $X \cap U$ is conflict-free in $\mathscr{F}\downarrow^{+}_{U}$. It remains to show $X \cap U = d_0(U, X \cap U)$.

For the direction $\subseteq$, let $c \in X \cap U$ be arbitrary and $x \in U$ be an arbitrary argument such that $x \to_U c$. By Lemma 10.18 we know that it must be $x \to_{Ar} c$. Since $c \in d_0(\mathscr{F}, X)$, there must be some $y \in X$ such that $y \to_{Ar} x$. By Lemma 10.20, there must be $z \in X \cap U$ such that $z \to_{Ar} x$. By Lemma 10.18, $z \to_U x$. Since x is arbitrary, we have $c \in d_0(U, X \cap U)$. Since c is arbitrary, $X \cap U \subseteq d_0(U, X \cap U)$.

For the direction $\supseteq$, we prove a stronger claim:

$$\text{for every } c \in U \text{, if } c \in d_0(U, X \cap U) \text{, then } c \in d_0(\mathscr{F}, X). \tag{10.4}$$

Namely,

$$\text{for every } c \in U, \text{ if } c \in d_0(U, X \cap U), \text{ then for every } x \in Ar \text{ with } x \to_{Ar} c, \text{ there is } y \in X \text{ such that } y \to_{Ar} x. \tag{10.5}$$

That is,

$$\text{for every } x \in Ar \text{ and } c \in d_0(U, X \cap U), \text{ if } x \to_{Ar} c, \text{ then there is } y \in X \text{ such that } y \to_{Ar} x. \tag{10.6}$$

Since $x \to_{Ar} c$, there must be $i \in \mathbb{N}$ such that $x \to^i_{Ar} c$. We prove (10.6) by induction on the value of i. If $i = 0$, then $x \to c$. Thus $x \in U$ by the definition of U. Since $c \in d_0(U, X \cap U)$, there must be $y \in X \cap U$ such that $y \to_U x$. By Lemma 10.18, we know that $y \to_{Ar} x$. Suppose (3) holds for $i = n$. If $x \to^{n+1}_{Ar} c$ and $x \not\to^n_{Ar} c$, then, by Definition 10.6, it must be either

- $\exists a \in Ar : a \multimap c \;\&\; x \to^n_{Ar} a$ or
- $\exists a \in Ar : a \multimap x \;\&\; a \to^n_{Ar} c$.

In the former case, since $a \multimap c$, $a \in U$ (by the definition of U) and $c \in d_0(U, X \cap U)$, $a \in d_0(U, X \cap U)$. Since $x \to^n_{Ar} a$, apply the IH, there must be $y \in X$ such that $y \to_{Ar} x$. In the latter case, since $a \to^n_{Ar} c$, apply the IH, there is also $y \in X$ such that $y \to_{Ar} x$. □

Lemma 10.24. *For any $X \in \sigma^c_7(U)$, X is admissible in $\mathscr{F}^+$.*

Proof. We first show that X is admissible in $\mathscr{F}^+$: By Lemma 10.18, it is easy to see that X is conflict-free in $\mathscr{F}^+$. To show $X \subseteq d_0(\mathscr{F}, X)$, it suffices to show

$$\text{for every } c \in X, \text{ if } c \in X, \text{ then } c \in d_0(\mathscr{F}, X). \tag{10.7}$$

Namely,

$$\text{for every } c \in X \text{ and } x \in Ar, \text{ if } x \to_{Ar} c \text{ then there is } y \in X \text{ such that } y \to_{Ar} x. \tag{10.8}$$

Since $x \to_{Ar} c$, there must be $i \in \mathbb{N}$ such that $x \to^i_{Ar} c$. We prove (10.8) by induction on the value of i. If $i = 0$, then $x \to c$. Thus $x \in U$ by the definition of U. Since $c \in X \subseteq d_0(U, X)$, there must be $y \in X$ such that $y \to_U x$. By Lemma 10.18, $y \to_{Ar} x$. Suppose (10.8) holds for $i = n$. If $x \to^{n+1}_{Ar} c$ and $x \not\to^n_{Ar} c$, then, by Definition 10.6, it must be either

- $\exists a \in Ar : a \multimap c \;\&\; x \to^n_{Ar} a$ or
- $\exists a \in Ar : a \multimap x \;\&\; a \to^n_{Ar} c$.

In the former case, since $X \in \sigma^c_7(U)$, $c \in X$ and $a \in U$ (by the definition of U), it follows from Lemma 10.21 that $a \in X$. Since $x \to^n_{Ar} a$, apply the IH, there must be $y \in X$ such that $y \to_{Ar} x$. In the latter case, since $a \to^n_{Ar} c$, by the IH, there is $y \in X$ such that $y \to_{Ar} a$. It follows that $y \to_{Ar} x$ by Definition 10.6. We have shown (10.8) and, therefore, (10.7). □

Lemma 10.25. *For every $X \subseteq Ar$, we say X is* inverse closed under suppport *if for every $x, y \in Ar$ if $x \multimap y$ and $y \in X$, then $x \in X$. Let X be a set inverse closed under support. Then for every $x \in X$ and $y \in U$, if $x \to_{Ar} y$ then there must be $z \in X \cap U$ such that $z \to_{Ar} y$ (thus, by Lemma 10.18, $z \to_U y$).*

Proof. Since $x \to_{Ar} y$, there must be $i \in \mathbb{N}$ such that $x \to^{i}_{Ar} y$. The proof is carried out by induction on the value of i. If $i = 0$, then $x \to y$. By the definition of U, $x \in U$. Thus we can set $z = x$. Suppose the lemma holds for $i = n$. If $x \to^{n+1}_{Ar} y$ and $x \not\to^{n}_{Ar} y$, then, by Definition 10.6, it must be either

- $\exists a \in Ar : a \multimap y \,\&\, x \to^{n}_{Ar} a$ or
- $\exists a \in Ar : a \multimap x \,\&\, a \to^{n}_{Ar} y$.

In the former case, it must be $a \in U$ by the definition of U. Since $x \to^{n}_{Ar} a$, by the IH, there exists $z \in X \cap U$ with $z \to_{Ar} a$. Thus $z \to_{Ar} y$ by the definition of $\to_{Ar}$. In the latter case, we have $a \in X$ since X is inverse closed under support. Since $a \to^{n}_{Ar} y$, apply the IH, there must be $z \in X \cap U$ such that $z \to_{Ar} y$.

□

Lemma 10.26. *For every $X \in \sigma^{c}_{7}(U)$, there is $Y \in \sigma^{c}_{7}(\mathscr{F})$ such that $Y \cap U = X$.*

Proof. From Lemma 10.24 we know that X is admissible in $\mathscr{F}^{+}$. For every $i \in \mathbb{N}$, we inductively define Y_i as follows:

1. $Y_0 = X$.
2. $Y_{n+1} = Y_n \cup d_0(\mathscr{F}, Y_n)$.

It is easy to see that $Y = \bigcup_{i \in \mathbb{N}} Y_i \in \sigma^{c}_{7}(\mathscr{F})$. It remains to show $Y \cap U = X$. We first show Y_i is inverse closed under support for every $i \in \mathbb{N}$ inductively: If $i = 0$, let $x, y \in Ar$ be such that $x \multimap y$ and $y \in X$. By the definition of U, $x \in U$. It follows from Lemma 10.21 that $x \in X$. Suppose Y_n is inverse closed under support. If $i = n + 1$, let $x, y \in Ar$ be such that $x \multimap y$ and $y \in Y_{n+1} = Y_n \cup d_0(\mathscr{F}, Y_n)$. If $y \in Y_n$, it follows from IH that $x \in Y_n \subseteq Y_{n+1}$. If $y \in d_0(\mathscr{F}, Y_n)$, it is easy to know $x \in d_0(\mathscr{F}, Y_n) \subseteq Y_{n+1}$.

We then show $Y_i \cap U = X$ for every $i \in \mathbb{N}$ inductively: The case $i = 0$ is trivial. Suppose $Y_n \cap U = X$, we need to show $Y_{n+1} \cap U = X$. That is, $(Y_n \cup d_0(\mathscr{F}, Y_n)) \cap U = X$. Since $Y_n \cap U = X$, it suffices to show $d_0(\mathscr{F}, Y_n) \cap U \subseteq X$. Since $X \supseteq d_0(U, X)$. We only need to $d_0(\mathscr{F}, Y_n) \cap U \subseteq d_0(U, X)$. Let $c \in d_0(\mathscr{F}, Y_n) \cap U$ be arbitrary and $x \in U$ be an arbitrary argument such that $x \to_U c$. Since $x \to_{Ar} c$ (by Lemma 10.18), there must be $y \in Y_n$ such that $y \to_{Ar} x$. Note that Y_n is inverse closed under support, it follows from Lemma 10.25 that there is also $z \in Y_n \cap U = X$ with $z \to_U x$. Since x is arbitrary, $c \in d_0(\mathscr{F}, Y_n) \cap U$.

□

Theorem 10.1. *Let a bipolar AF $F = (Ar, \to, \multimap)$ be given. Let $U \subseteq Ar$ be such that for all $a \in Ar \backslash U$ and $b \in U$, $a \not\to b$ and $a \not\multimap b$. $\sigma^{c}_{7}(U) = \{X \cap U \mid X \in \sigma^{c}_{7}(\mathscr{F})\}$.*

Proof. Immediately from Lemma 10.23 and Lemma 10.26. □

Proposition 10.126. *σ^{p}_{7} satisfy P4.10.*

Lemma 10.27. *For every $X \in \sigma^{p}_{7}(\mathscr{F})$, $X \cap U \in \sigma^{p}_{7}(U)$.*

Proof. Since preferred extensions are maximal complete extensions which has been proved by Dung, $X \in \sigma^{c}_{7}(\mathscr{F})$. Thus, by Lemma 10.23, $X \cap U \in \sigma^{c}_{7}(\mathscr{F})$. It remains to show $X \cap U$ is maximal among all complete extensions of $(\mathscr{F} \downarrow_U)^{+}$. Suppose there is $Y \in \sigma^{p}_{7}(U)$ such that

$Y \supset X \cap U$. Consider the set $Y \cup X$. It suffices to show $Y \cup X$ is admissible, because this will contradict the fact that $X \in \sigma_7^p(\mathscr{F})$.

We first show $Y \cup X$ is conflict-free. Suppose not. Since both X and Y are conflict-free in F^+. It can only be that there are $a \in Y \backslash X$ and $b \in X \backslash Y$ such that $a \rightarrow_{Ar} b$ or $b \rightarrow_{Ar} a$. In the former case, since X is admissible in F^+ and $b \in X$, there must be $c \in X$ such that $c \rightarrow_{Ar} a$. Since $X \in \sigma_7^c(\mathscr{F})$, if follows from Lemma 10.21 that X is closed under support. Thus, by Lemma 10.25, there must be $z \in X \cap U$ such that $z \rightarrow_{Ar} a$, contradicting with that Y is conflict-free in F^+. In the latter case, by Lemma 10.25, there is also $z \in X \cap U$ such that $z \rightarrow_{Ar} a$, contradiction!

We then show $Y \cup X \subseteq d_0(\mathscr{F}, Y \cup X)$. Since $X \in \sigma_7^p(\mathscr{F})$, $X \subseteq d_0(\mathscr{F}, X) \subseteq d_0(\mathscr{F}, Y \cup X)$. Since $Y \in \sigma_7^p(U) \subseteq \sigma_7^c(U)$, by Lemma 10.24, Y is admissible in F^+. Thus $Y \subseteq d_0(\mathscr{F}, Y) \subseteq d_0(\mathscr{F}, Y \cup X)$. □

Lemma 10.28. *For every $X \in \sigma_7^p(U)$, there is $Y \in \sigma_7^p(\mathscr{F})$ such that $Y \cap U = X$.*

Proof. Since $X \in \sigma_7^p(U)$, $X \in \sigma_7^c(U)$. By Lemma 10.26, there must be $X' \in \sigma_7^c(\mathscr{F})$ such that $X' \cap U = X$. Since Ar is finite, there must be $Y \in \sigma_7^p(\mathscr{F})$ such that $Y \supseteq X'$. Consider the set $Y \cap U$. By Lemma 10.27, $Y \cap U \in \sigma_7^p(U)$. We have $Y \cap U \supseteq X$ since $X \subseteq Y$ and $X \subseteq U$. Because $X \in \sigma_7^p(U)$, we have $Y \cap U = X$. □

Theorem 10.2. *Let a bipolar AF $F = (Ar, \rightarrow, \multimap)$ be given. Let $U \subseteq Ar$ be such that for all $a \in Ar \backslash U$ and $b \in U$, $a \not\rightarrow b$ and $a \not\multimap b$. $\sigma_7^p(U) = \{X \cap U \mid X \in \sigma_7^p(\mathscr{F})\}$.*

Proof. Immediately from Lemma 10.27 and Lemma 10.28. □