Graph Models for Rational Social Interaction

Dissertation

submitted towards the degree
Doctor of Natural Science
of the Faculty of Mathematics and Computer Science of
Saarland University

by

Cosmina Croitoru

Saarbrücken
2017

Day of Colloquium: 11. Oktober 2017
Dean of of the Faculty:
Prof. Dr. Frank-Olaf Schreyer

Prüfungsausschuss:
Prof. Dr. Jörg Hoffmann (Chair of the Committee)
Prof. Dr. Dr. h.c. mult. Kurt Mehlhorn (First Reviewer)
Research Director Dr. Leila Amgoud (Second Reviewer)
Dr. Marvin Künnemann (Academic Assistant)

Abstract

A fundamental issue in multi-agent systems is to extract a consensus from a group of agents with different perspectives. Even if the bilateral relationships (reflecting the outcomes of disputes, product comparisons, or evaluation of political candidates) are rational, the collective output may be irrational (e.g., intransitivity of group preferences). This motivates AI's research for devising social outcomes compatible with individual positions. Frequently, such situations are modeled as graphs. While the preponderance of formal theoretical studies of such graph based-models has addressed semantic concerns for defining a desirable output in order to formalize some high-level intuition, results relating to algorithmic and computational complexity are also of great significance from the computational point of view.

The first Part of this thesis is devoted to combinatorial aspects of *Argumentation Frameworks* related to computational issues. These abstract frameworks, introduced by Dung in 1995, are directed graphs with nodes interpreted as *arguments* and the directed edges as *attacks* between the arguments. By designing a *conflict-resolution* formalism to make distinction among *acceptable* and *unacceptable* arguments, Dung initiated an important area of research in Artificial Intelligence. I prove that any argumentation framework can be syntactically augmented into a *normal form* preserving the semantic properties of the original arguments, by using a cubic time rewriting technique. I introduce *polyhedral labellings* for an argumentation frameworks, which is a polytope with the property that its integral points are exactly the incidence vectors of specific types of Dung's outcome. Also, a new notion of acceptability of arguments is considered – *deliberative acceptability* – and I provide it's time computational complexity analysis. This part extends and improves some of the results from the my Master thesis.

In the second Part, I introduce a novel graph-based model for aggregating preferences. By using graph operations to describe properties of the aggregators, axiomatic characterizations of aggregators corresponding to usual majority or approval & disapproval rule are given. Integrating Dung's semantics into our model provides a novel qualitative approach to classical social choice: *argumentative aggregation of individual preferences*. Also, a functional framework abstracting many-to-many two-sided markets is considered: the study of the existence of a *Stable Choice Matching* in a *Bipartite Choice System* is reduced to the study of the existence of *Stable Common Fixed Points* of two choice functions. A generalization of the Gale-Shapley algorithm is designed and, in order to prove its correctness, a new characterization of path independence choice functions is given.

Finally, in the third Part, we extend Dung's Argumentation Frameworks to *Opposition Frameworks*, reducing the gap between *Structured* and *Abstract Argumentation*. A *guarded attack calculus* is developed, giving proper generalizations of Dung's extensions.

Zusammenfassung

Ein grundlegendes Problem von Multiagentensystemen ist, eine Gruppe von Agenten mit unterschiedlichen Perspektiven zum Konsens zu bringen. Während die bilaterale Ergebnisse von Rechtsstreitigkeiten, Produktvergleichen sowie die Bewertung von politischen Kandidaten wiederspiegelnden Beziehungen rational sein sollten, könnte der kollektive Ausgang irrational sein z.B. durch die Intransitivität von Präferenzen der Gruppe. Das motiviert die KI-Forschung zur Entwicklung von sozialen Ergebnissen, welche mit individuellen Einstellungen kompatibel sind. Häufig werden solche Situationen als Graphen modelliert. Während die meisten formalen theoretischen Studien von Graphmodellen sich mit semantischen Aspekten für die Definition eines wünschenswerten Ausgangs zur Formalisierung auf hohem Intuitionsniveau beschäftigen, ist es ebenfalls von großer Bedeutung, die Komplexität von Algorithmen und Berechnungen zu verstehen.

Der erste Teil der vorliegenden Arbeit widmet sich den kombinatorischen Aspekten von *Argumentation Frameworks* im Zusammenhang mit rechnerischen Fragen. Diese von Dung in 1995 eingeführten abstrakten Frameworks sind gerichtete Graphen mit als *Argumenten* zu interpretierenden Knoten, wobei die gerichteten Kanten *Angriffe* zwischen den Argumenten sind. Somit hat Dung mit seiner Gestaltung eines *Konfliktlösungsformalismus* zur Unterscheidung zwischen *akzeptablen* und *inakzeptablen* Argumenten für einen wichtigen Bereich von Forschung in KI den Grundstein gelegt. Die Verfasserin hat bewiesen, dass jedes Argumentation Framework sich in einer die semantischen Eigenschaften der originalen Argumente bewahrenden *normalen Form* syntaktisch erweitern lässt, indem man eine mit kubischer Laufzeit umwandelnde Technik verwendet. Neu eingefürt werden hier *Polyhedrische Etiketten* für Argumentation Frameworks. Dabei handelt es sich um einen Polytop, wessen ganze Punkte genau die Inzidenzvektoren von bestimmten Arten von Dungs Ausgabe sind. Weiterhin wird ein neuer Begriff der Akzeptanz von Argumenten geprägt, nämlich - *deliberative Akzeptanz* - und dessen Komplexität analysiert. Dieser Teil erweitert und verfeinert einige ihrer Ergebnisse aus der Masterarbeit.

Im zweiten Teil wurde ein neuartiges graphenbasiertes Modell für die Aggregation von Präferenzen erarbeitet. Hier werden axiomatische Charakterisierungen von Aggregatoren neu eingeführt, und zwar durch die Verwendung von Graphoperationen zur Beschreibung der Eigenschaften von Aggregatoren. Sie entsprechen dem üblichen Mehrheitsprinzip bzw. der Genehmigungs- & Ablehnungsregel. Einen neuartigen, qualitativen Ansatz im Vergleich zu der klassischen Sozialwahltheorie bietet die Integration der Semantik von Dung in dem neuen Modell, und zwar *argumentative Aggregation individueller Präferenzen*. Desweiteren wird ein funktionales many to many zweiseitige Märkte abstrahierendes Framework untersucht, indem statt die Existenz einer *Stabilen Wahl Matching* in einem *Bipartite Wahlsystem* zu studieren, wird die Existenz von *Stable Common Fixed Points* auf zwei Wahlfunktionen erforscht. Im nächsten Schritt wird eine neue Verallgemeinerung des Gale-Shapley Algorithmus entworfen und eine neue Charakterisierung der Wegunabhängigkeitsfunktion gegeben, die einen Korrektheitsbeweis für den Algorithmus ermöglicht.

Im dritten Teil werden schließlich Dungs Argumentation Frameworks auf *Opposition Frameworks* erweitert und dadurch die in der gegenwärtigen Forschung bestehende Lücke zwischen *strukturierter* und *abstrakter Argumentation* verringert. Dafür wird ein *bewachter Angriffskalkül* entwickelt, welches strikten Verallgemeinerungen von Dungs echten Erweiterungen führt.

Acknowledgements

I would like to express my sincere gratitude to all my colleagues from the group Algorithms and Complexity (Department 1 from MPII Saarbrücken) who have assured me a very pleasant and stimulating working environment throughout my PhD program.

I want also to express my sincere thanks to my good friends from IMPRS-CS (International Max Planck Research School for Computer Science) staff, and the funny guys from IST, without which the everyday life in MPII wouldn't be that fun.

I am all gratitude to professor Jörg Hoffmann, and professor Leila Amgoud for serving as my committee members. Thank you for letting my defense be such an enjoyable moment, for your brilliant comments, suggestions and questions.

I am deeply indebted (for all the rest of my life) to my supervisor Prof. Dr. Kurt Mehlhorn whose enlightening discussions, patience, motivation, enthusiasm and deep knowledge influenced me in all the time of research to finally complete this thesis.

I am very grateful to my parents and my sister, who always supported me during my studies and who actively watched my life way.

Thanks!

Contents

List of Figures

Introduction

This thesis covers various areas in *Social Reasoning* having as central hub *Abstract Argumentation*, viewed as a **graph-based modeling** of the fundamental issues that arise in defeasible domains. My own view of the well-known *Argumentation Frameworks*, introduced by Dung [Dun95], is:

> If the edges of a given directed graph are seen as attacks, how can a satisfactory set of winner nodes be rationally selected and justified?

The nodes of such directed graphs are interpreted as *arguments* and in order to make distinction among *acceptable* and *unacceptable* arguments, Dung defined some families of sets of nodes (called extesions) representing sets of collectively accepted arguments. General network reasoning models investigating the informal logic structure of many social and economic problems instantiate Dung's argumentation frameworks, and therefore can be implemented based on an unifying principle.

A natural defeasible domain is provided by the classical *Social Choice Theory* (SCT) [Arr63], where on a shared *reasons space*, a *society* expresses a set of possibly shared forms of *subjectivity*, whose deeper interactions enable new consistent collective judgments, creating *social inference relations*. The corresponding graph models are bipartite digraphs in which one part represents the set of *alternatives* and the other part the set of *individuals*, the society. It follows a simple graph-based model for aggregating dichotomous preferences, paving the way of using graph operations to describe *normative* properties of the aggregators.

One interesting application discussed by Dung in his original paper [Dun95], is to view the Gale and Shapley *stable matchings* ([GS62]) as an instance of the *stable* output in an appropriate argumentation framework. Generalizations of stable matchings (motivated by many-to-many matchings in two-sided markets applications) can be also approached by considering special graph-based models. We consider *Bipartite Choice Systems* as a family of choice functions defined on the edge set of a bipartite multi-graph and indexed by the vertices of this multi-graph. Then, a *Stable Choice Matching* is defined as a stable common fixed point of two associated collective choice functions.

Dung's argumentation frameworks have been generalized and extended in order to obtain better models (see Brewka *et al.* [BPW14], or Modgil [Mod13]). We propose a new such generalization, by using edge-labeled multi-digraphs.

Structure of the Thesis

After a preliminary chapter, the thesis contains six chapters (organized in three parts as described below) and a short concluding discussion chapter.

Preliminaries Basic notions and results of Dung's *Theory of Argumentation* are presented together with time-complexity results of the main related decision problems. An entire section, entitled *Graph-theoretic Digression*, discusses related work (already) done in *Graph Theory*, as a collection of small remarks. A nice treatment of quasi-kernels using a two-way scan of the vertices of a graph is also presented.

Part 1 **Combinatorial Aspects of Argumentation Frameworks**

Chapter 1 A new type of acceptability of an argument, *deliberative acceptance* – in which its attacking and defending sets of arguments are uniformly treated – is introduced. We discuss how this and the classical acceptance notions interrelate and analyze its computational properties.

Chapter 2 We say that an argumentation framework is in *normal form* if no argument attacks a conflicting pair of arguments. An *augmentation* of an argumentation framework is obtained by adding new arguments and changing the attack relation such that the acceptability status of original arguments is maintained in the new framework. Furthermore, we define join-normal semantics leading to augmentations of the joined argumentation frameworks. Also, a rewriting technique which transforms in cubic time a given argumentation framework into a normal form is devised.

Chapter 3 We define a *polyhedral labelling* for an argumentation framework $AF = (A, D)$ as a set of solutions $x \in \mathbb{R}^A$ (x_a is the label of the argument $a \in A$), to a system of linear constraints, such that the integral solutions are exactly the incidence vectors of some specific type of Dung's extensions. The linear constraints vary from the obvious $x_a = 1$ for each non attacked argument a, or $x_a + x_b \leq 1$ for each attack $(a, b) \in D$ (in order to assure Dung's conflict-free condition), to more deep inequalities of the form "*the sum of the label of an argument and the labels of all its attackers is at least 1*" or if (b, a) is an attack then "*the label of a is not greater than the sum of the labels of all attackers of b*".

Part 2 **Graphs and Social Choice Theory**

Chapter 4 A novel graph-based model for aggregating dichotomous preferences is introduced. We are not only interested into positive collective position (as is intensively done in social choice theory) but also into the negative collective position (which can be used also as an explanation of the selected positive facts). The set of individual opinions (called "profile" in social choice theory) can be represented using a bipartite directed graph, which we call *debate*. In this way, well-established graph theoretical notions (digraph isomorphism, in-degree and out-degree of a vertex, induced sub-digraphs, digraph operations, etc.) can be used in describing normative conditions on the aggregating rules. The aggregate opinion can also be viewed as a debate (bipartite digraph) in which each individual has the same sets of in-neighbors and out-neighbors (called *consensual debate*). This is used to obtain new axiomatic characterizations of aggregators corresponding to usual majority or approval & disapproval rule. We introduce *argumentative aggregation of individual opinions* in which *collective opinions* are obtained by merging the opinions of non-conflicting coalitions of individuals.

Chapter 5 In the *Economics Matching*, broad generalizations of the classical model of Gale and Shapley were considered by allowing centralized matching schemes where there could be multiple partners on both sides (many-to-many) of the market and the agent's preferences were given by choice functions on the set of its neighbors. We introduce a functional framework abstracting these models, propose a new generalization of the Gale-Shapley algorithm and, in order to prove its correctness, a new characterization of path independence choice functions is obtained.

Part 3 Opposition Frameworks

Chapter 6 In this chapter we introduced a new generalization of Dung's argumentation framework which is conceptually different from other generalized abstract argumentation frameworks. It formally exploits the link that exists between "attacks" and "node's positions" capturing some high-level intuition, not addressed by other proposals. The "nodes" of our *Opposition Frameworks* have a minimal *content* expressed as finite non-empty sets of facts (the node's position), which are used to relate the "attacks" between two nodes to their positions. We introduced a simple recursive definition of acceptance: a node (the position expressed by a node) is accepted if either it can counterattack all attacks targeting it or there is another "compatible" node such that in the opposition framework obtained by "contracting" these two nodes in a single "supernode", this supernode is accepted. The use of the set of attacks instead of the set of parents in the study of the acceptability of a node, simplifies the description of a novel DPLL type backtracking acceptance algorithm. The characterization of the basic outputs (admissible sets) by the models of a propositional logic formula is also given.

Concluding Discussion A summary of the contributions of this work and a short discussion on further possibilities of improving it.

Preliminaries

0.1. Introduction

Abstract argumentation frameworks, introduced by Dung [Dun95], constitute a common mechanism for studying reasoning in defeasible domains and for relating different non-monotonic formalisms. General network reasoning models investigating the informal logic structure of many social and economic problems instantiate Dung's argumentation frameworks, and therefore can be implemented based on an unifying principle.

This graph-theoretic model of argumentation frameworks focuses on the manner in which a specified set A of abstract arguments interact via an attack (defeat) binary relation D on A. If $(a,b) \in D$ (argument *a attacks* argument b) we have a conflict. A *conflict-free* set of arguments is a set $T \subseteq A$ such that there are no $a,b \in T$ with $(a,b) \in D$. An *admissible* set of arguments is a conflict-free set $T \subseteq A$ such that the arguments in T defend themselves "collectively" against any attack: for each $(a,b) \in D$ with $b \in T$, there is $c \in T$ such that $(c,a) \in D$.

In this model, the main aim of argumentation is deciding the status of arguments. The acceptability of an argument a is defined based on its membership in an admissible set of arguments satisfying certain properties (formalizing different intuitions about which arguments to accept on the basis of the given framework) called *semantics*. The attack graph is given in advance – abstracting on the underlying logic and structure of arguments, as well on the reason and nature of the attacks – and provides a defeasible-based conceptualization of commonsense reasoning.

0.2. Dung's Theory of Argumentation

In this section we present the basic concepts used for defining classical semantics in abstract argumentation frameworks introduced by Dung in 1995, [Dun95]. All notions and results, if not otherwise cited, are from this paper (even some of them are not literally the same). We consider U a fixed countable *universe* of arguments.

Definition 1 An *Argumentation Framework* is a digraph $AF = (A,D)$, where $A \subset U$ is finite and nonempty, the vertices in A are called *arguments*, and if $(a,b) \in D$ is a directed edge, then *argument a defeats (attacks) argument b*. A, the argument set of AF, is referred as $Arg(AF)$ and its attack set D is referred as $Def(AF)$. The set of all argumentation frameworks (over U) is denoted by $\mathbb{AF}$.

All definitions and concepts from graph theory are adapted implicitly or explicitly.

If $AF = (A, D)$ is an argumentation framework and $A_1 \subseteq A$, then the **argumentation framework induced by** A_1 **in** AF is $AF[A_1] = (A_1, D \cap (A_1 \times A_1))$.

Two argumentation frameworks AF_1 and AF_2 are *isomorphic* (denoted $AF_1 \cong AF_2$) if there is a bijection $h : Arg(AF_1) \to Arg(AF_2)$ such that

$$(a,b) \in Def(AF_1) \text{ if and only if } (h(a),h(b)) \in Def(AF_2).$$

h is called an *argumentation framework isomorphism*, and it is emphasized by the notation $AF_1 \cong_h AF_2$. If $S \subseteq Arg(AF_1)$ then $h(S) \subseteq Arg(AF_2)$ is the set $\{h(a)|a \in S\}$. Similarly, if $M \subseteq 2^{Arg(AF_1)}$, then $h(M) \subseteq 2^{Arg(AF_2)}$ is $h(M) = \{h(S)|S \in M\}$.

The extension-based acceptability semantics is a central notion in Dung's argumentation frameworks, which we define as follows (see also [BG07b]).

Definition 2 An *extension-based acceptability semantics* is a function σ that assigns to every argumentation framework $AF \in \mathbb{AF}$ a family of sets $\sigma(AF) \subseteq 2^{Arg(AF)}$ such that

$$\forall AF_1, AF_2 \in \mathbb{AF}, \text{ if } AF_1 \cong_h AF_2 \text{ then } \sigma(AF_2) = h(\sigma(AF_1)).$$

A member $E \in \sigma(AF)$ is called a σ*-extension* in AF.

If a semantics σ satisfies the condition $|\sigma(AF)| = 1$ for any argumentation framework AF, then σ is said to belong to the *unique-status approach*, otherwise to the *multiple-status approach* (Prakken and Vreeswijk, [PV02]).

The main types of argument's acceptability status with respect to a given semantics are defined as follows.

Definition 3 Let $AF = (A, D)$ be an argumentation framework, $a \in A$ be an argument, and σ be a semantics.
a is σ*-credulously accepted* if and only if $a \in \bigcup_{S \in \sigma(AF)} S$.
a is σ*-sceptically accepted* if and only if $a \in \bigcap_{S \in \sigma(AF)} S$.

Let $AF = (A, D)$ be an argumentation framework. For each $a \in A$ we denote $a^+ = \{b \in A|\ (a,b) \in D\}$ the set of all arguments *attacked* by a, and $a^- = \{b \in A|\ (b,a) \in D\}$ the set of all arguments *attacking* a. These notations can be extended to sets of arguments. The set of all arguments *attacked by* (the arguments in) $S \subseteq A$ is $S^+ = \bigcup_{a \in S} a^+$, and the set of all arguments *attacking* (the arguments in) S is $S^- = \bigcup_{a \in S} a^-$. We also have $\emptyset^+ = \emptyset^- = \emptyset$.

The set S of arguments *defends* an argument $a \in A$ if $a^- \subseteq S^+$ (i.e., any a's attacker is attacked by an argument in S). The set of *all arguments defended by a set* S of arguments is denoted by $F(S)$:

$$F(S) = \{a \in A | a^- \subseteq S^+\}.$$

If $\mathbb{M}_{AF}$ is a non-empty set of sets of arguments in AF, then $\mathbf{max}(\mathbb{M}_{\mathbf{AF}})$ denotes the set of maximal (with respect to set inclusion) members of $\mathbb{M}_{AF}$ and $\mathbf{min}(\mathbb{M}_{\mathbf{AF}})$ denotes the set of its minimal (with respect to set inclusion) members.

We now define the main admissibility extension-based acceptability semantics.

Definition 4 Let $AF = (A, D)$ be an argumentation framework.

- A *conflict-free set* in AF is a set $S \subseteq A$ with property $S \cap S^+ = \emptyset$ (i.e., there are no attacking arguments in S). The family of all conflict free-sets is

$$\mathbf{cf}(AF) = \{S \subseteq A | S \text{ is conflict-free set }\}.$$

- An *admissible set* in AF is a set $S \in \mathbf{cf}(AF)$ with property $S^- \subseteq S^+$ (i.e., defends its elements). The family of all admissible sets is

$$\mathbf{adm}(AF) = \{S \subseteq A | S \text{ is admissible set }\}.$$

- A *complete extension* in AF is a set $S \in \mathbf{cf}(AF)$ with property $S = F(S)$. The family of all complete extensions is

$$\mathbf{comp}(AF) = \{S \subseteq A | S \text{ is complete extension }\}.$$

- A *preferred extension* in AF is a set $S \in \mathbf{max}(\mathbf{comp}(AF))$. The family of all preferred extensions is

$$\mathbf{pref}(AF) := \mathbf{max}(\mathbf{comp}(AF)).$$

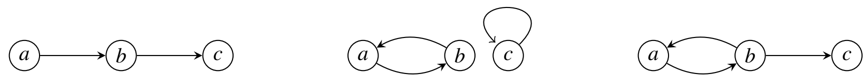
- A *grounded extension* in AF is a set $S \in \mathbf{min}(\mathbf{comp}(AF))$. The family of all grounded extensions is

$$\mathbf{gr}(AF) := \mathbf{min}(\mathbf{comp}(AF)).$$

- A *stable extension* in AF is a set $S \in \mathbf{cf}(AF)$ with the property $S^+ = A - S$. The family of all stable extensions is

$$\mathbf{stb}(AF) = \{S \subseteq A | S \text{ is stable extension }\}.$$

Examples:

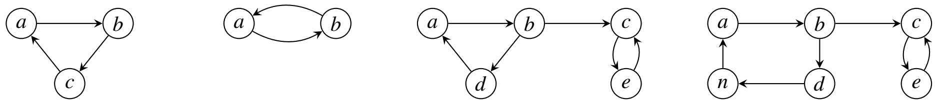


AF_1	AF_2	AF_3
$\mathbf{cf}(AF_1) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a,c\}\}$	$\mathbf{cf}(AF_2) = \{\emptyset, \{a\}, \{b\}\}$	$\mathbf{cf}(AF_3) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a,c\}\}$
$\mathbf{adm}(AF_1) = \{\emptyset, \{a\}, \{a,c\}\}$	$\mathbf{adm}(AF_2) = \{\emptyset, \{a\}, \{b\}\}$	$\mathbf{adm}(AF_3) = \{\emptyset, \{a\}, \{b\}, \{a,c\}\}$
$\mathbf{comp}(AF_1) = \{\{a,c\}\}$	$\mathbf{comp}(AF_2) = \{\emptyset, \{a\}, \{b\}\}$	$\mathbf{comp}(AF_3) = \{\emptyset, \{b\}, \{a,c\}\}$
$\mathbf{pref}(AF_1) = \{\{a,c\}\}$	$\mathbf{pref}(AF_2) = \{\{a\}, \{b\}\}$	$\mathbf{pref}(AF_3) = \{\{b\}, \{a,c\}\}$
$\mathbf{gr}(AF_1) = \{\{a,c\}\}$	$\mathbf{gr}(AF_2) = \{\emptyset\}$	$\mathbf{gr}(AF_3) = \{\emptyset\}$

Figure 0.1.: Different extensions of AF's.

Note that $\emptyset \in \mathbf{adm}(AF)$ for any AF (hence $\mathbf{adm}(AF) \neq \emptyset$) and if $a \in A$ is a self-attacking argument (i.e., $(a,a) \in D$), then a is not contained in an admissible set. It is not difficult to see that any admissible set is contained in a preferred extension, which exists in any AF; the preferred extension is unique if AF has no directed cycle of even length (Bench-Capon, [BC03], Baroni and Giacomin, [BG03]).

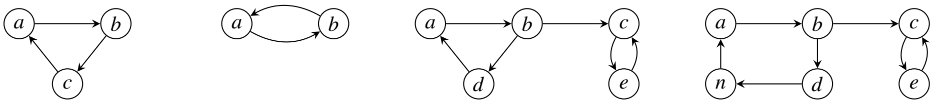
The examples in the Figure 0.2 show that $\mathbf{pref}(AF)$ is strongly influenced by the existence of circuits in AF and by the even or odd length of these circuits.



AF_1	AF_2	AF_3	AF_4
$\mathbf{pref}(AF_1) = \{\emptyset\}$	$\mathbf{pref}(AF_2)=$ $=\{\{a\}, \{b\}\}$	$\mathbf{pref}(AF_3)=\{\{e\}\}$	$\mathbf{pref}(AF_4)=\{\{n,b,e\},$ $\{a,d,c\}, \{a,d,e\}\}$

Figure 0.2.: $\mathbf{pref}(AF)$ vs the existence of odd and even circuits in AF.

In [Cro12], is proved that a complete extension $S \in \mathbf{comp}(AF)$ is a preferred extension if and only if S is either a stable extension or, in the argumentation framework $AF[A-(S \cup S^+)]$ the only admissible set is the empty set, as in the figure below.



AF_1	AF_2	AF_3	AF_4
$\mathbf{st}(AF_1) = \emptyset$	$\mathbf{st}(AF_2)=$ $=\{\{a\}, \{b\}\}$	$\mathbf{st}(AF_3)=\emptyset$	$\mathbf{st}(AF_4)=\{\{n,b,e\},$ $\{a,d,c\}, \{a,d,e\}\}$

Figure 0.3.: $\mathbf{st}(AF)$.

Let us note that $\mathbf{st}(AF)$ can be empty. However, when in AF there are no circuits, the following theorem holds.

Theorem 5 *[Dun95] If the argumentation framework* $AF = (A, D)$ *has no circuits (it is a DAG) then*
$$\boldsymbol{st}(AF) = \boldsymbol{pref}(AF) = \boldsymbol{gr}(AF) \neq \emptyset.$$

The grounded extension exists and it is unique in any argumentation framework. It can be constructed by a very simple algorithm:

> *consider all non-attacked arguments, delete these arguments and those attacked by them from the digraph, and repeat these two steps for the digraph obtained until no node remains.*

An equivalent way to express Dung's extension-based semantics is using argument labellings as proposed by Caminada [Cam06a] (originally introduced in Pollock [Pol95]). The idea underlying the labellings-based approach is to assign to each argument a label from the set $\{I, O, U\}$. The label I (i.e., In) means the argument is accepted, the label O (i.e., Out) means the argument is rejected, and the label U (i.e., Undecided) means one abstains from an opinion on whether the argument is accepted or rejected.

Definition 6 *[Cam06a]* Let $AF = (A, D)$ be an argumentation framework. An *admissible labelling* of AF is a function $Lab : A \to \{I, O, U\}$ such that $\forall a \in A$:
• $Lab(a) = I$ if and only if $a^- \subseteq Lab^{-1}(O)$,
• $Lab(a) = O$ if and only if $a^- \cap Lab^{-1}(I) \neq \emptyset$.
A *complete labelling* of AF is an admissible labelling Lab such that $\forall a \in A$: $Lab(a) = U$ if and only if $a^- \cap Lab^{-1}(I) = \emptyset$ and $a^- \cap Lab^{-1}(U) \neq \emptyset$. A *grounded labelling* of AF is a complete labelling Lab such that there is no complete labelling Lab_1 with $Lab_1^{-1}(I) \subset Lab^{-1}(I)$. A *preferred labelling* of AF is a complete labelling Lab such that there is no complete labelling Lab_1 with $Lab^{-1}(I) \subset Lab_1^{-1}(I)$. A *stable labelling* of AF is a complete labelling Lab such that $Lab^{-1}(U) = \emptyset$.

In [Cam06a] it was proved that, for any argumentation framework $AF = (A, D)$ and any semantics $\sigma \in \{\mathbf{adm}, \mathbf{comp}, \mathbf{gr}, \mathbf{pref}, \mathbf{stb}\}$, a set $S \subseteq A$ satisfies $S \in \sigma(AF)$ if and only if there is a σ-labelling Lab of AF such that $S = Lab^{-1}(I)$.

In the context of combinatorial games, Fraenkel [Fra97] considered a special type of partitions, *P,N,D -partitions*, which are closely related to Caminada's labellings. Based on these, I observed that the above algorithm that constructs a ground extension in an AF can be used to give an intrinsic characterization of grounded labellings.

Observation 7 *Let* $AF = (A, D)$ *be an argumentation framework. A complete labelling Lab of AF is a grounded labelling if and only if there is a linear order* $<$ *on* $Lab^{-1}(I)$ *such that the following condition holds:*

if $a \in Lab^{-1}(I)$ *and* $b \in a^-$ *then there is* $a' \in Lab^{-1}(I) \cap b^-$ *such that* $a' < a$.

0.3. Complexity-theoretic Issues

Depending on the semantics $\sigma \in \{\mathbf{adm}, \mathbf{comp}, \mathbf{gr}, \mathbf{pref}, \mathbf{stb}\}$ considered, the basic decision problems related to the acceptability of arguments in a given argumentation framework are listed below.

$\mathbf{VER}_\sigma$ (Verification)
Instance : $AF = (A, D)$ and $S \subseteq A$.
Question : Is $S \in \sigma(AF)$?

$\mathbf{CA}_\sigma$ (Credulous Acceptance)
Instance : $AF = (A, D)$ and $a \in A$.
Question : Is there $S \in \sigma(AF)$ such that $a \in S$?

$\mathbf{SA}_\sigma$ (Skeptical Acceptance)
Instance : $AF = (A, D)$ and $a \in A$.
Question : Is a a member of each $S \in \sigma(AF)$?

$\mathbf{EX}_\sigma$ (Existence)
Instance : $AF = (A, D)$.
Question : Is $\sigma(AF) \neq \emptyset$?

$\mathbf{NE}_\sigma$ (Non-Emptiness)
Instance : $AF = (A, D)$.
Question : Is $\sigma(AF) = \{\emptyset\}$?

The problem $\mathbf{Ver}_\sigma$ is in $\mathbf{P}$ for each $\sigma \in \{\mathbf{adm}, \mathbf{gr}, \mathbf{stb}\}$, since all constraints defining a member of $\sigma(AF)$ can be verified in polynomial time with respect to $|Arg(AF)| + |Def(AF)|$.

Despite of the fact that $\sigma(AF) = \emptyset$ and $\sigma(AF) = \{\emptyset\}$ give the same answer to an acceptance query, the problems $\mathbf{EX}_\sigma$ and $\mathbf{NE}_\sigma$ are different. The problem $\mathbf{EX}_\sigma$ is trivial (each istance is an yes-instance) for any $\sigma \in \{\mathbf{adm}, \mathbf{comp}, \mathbf{gr}, \mathbf{pref}\}$ but $\mathbf{EX}_{\mathbf{stb}}$ is $\mathbf{NP}$-complete (this follows easily from Chvátal [Chv73]). Clearly, all the decision problems listed above are in $\mathbf{P}$ for $\sigma = \mathbf{gr}$.

From the work of Dimopoulos and Torres [DT96] it follows that $\mathbf{CA}_{\mathbf{adm}}$ is $\mathbf{NP}$-complete and $\mathbf{NE}_{\mathbf{adm}}$ is $\mathbf{coNP}$-complete. Since there is a simple polynomial time reduction from $\mathbf{NE}_{\mathbf{adm}}$ to $\mathbf{VER}_{\mathbf{pref}}$, it follows that $\mathbf{VER}_{\mathbf{pref}}$ is $\mathbf{coNP}$-complete.

The problem $\mathbf{SA}_{\mathbf{pref}}$ is difficult. This has been proved by Dunne and Bench-Capon [DBC02], by showing that this problem is at the third-level of the polynomial-time hierarchy. If the complexity class Π_2^{P} comprises those problems decidable by co-NP computations given (unit cost) access to an NP complete oracle, then $\mathbf{SA_{pref}}$ is Π_2^{P}-complete [DBC02].

We close this section by noting that every $AF = (A, D)$ induces a propositional theory (a set of propositional formula) $\mathbb{T}_{AF} = \{x_a \leftrightarrow \wedge_{b \in a^-} \neg x_b | a \in A\}$, where, as usual, $\wedge_{b \in \emptyset} \neg x_b = true$. Using logical equivalences this can be transformed in CNF (by denoting A_0 the set of arguments a with $a^- = \emptyset$):

$$\mathbb{T}_{AF} = \bigwedge_{a \in A - A_0} \big((x_a \vee \bigvee_{b \in a^-} x_b) \wedge \bigwedge_{b \in a^-} (\neg x_a \vee \neg x_b)\big) \wedge \bigwedge_{a \in A_0} x_a.$$

This can be written equivalently as:

$$\mathbb{T}_{AF} = \bigwedge_{a \in A_0} x_a \wedge \bigwedge_{a \in A - A_0} (x_a \vee \bigvee_{b \in a^-} x_b) \wedge \bigwedge_{(b,a) \in D} (\neg x_a \vee \neg x_b).$$

It is not difficult to prove that AF has a stable extension if and only if $\mathbb{T}_{AF}$ is satisfiable. Similar approaches (with important complexity corollaries, or practical implementations) can be found in Creignou [Cre95], Besnard and Doutre [BD04], Walicki and Dyrkolbotn [WD12], Bezem, Grabmayer, and Walicki [BGW12].

0.4. Graph-theoretic Digression

If $D = (V, E)$ is a digraph, a *kernel* in D is a stable set $S \subseteq V$ (that is, $E \cap S \times S = \emptyset$) with the property that for any $v \in V - S$ there is $w \in S$ such that $(v, w) \in E$. Clearly, S is a kernel in D if and only if S is a stable extension in $AF_D = (V, \tilde{E})$, where $\tilde{E} = \{(v, w) | (w, v) \in E\}$.

It seems that von Neumann and Morgenstern [NM44] were the first to introduce kernels when describing their concept of *solution* of a n-person game. Dung [Dun95] explained that for any n-person game we can associate an argumentation framework (having as arguments *imputations* of the game and attack relation given by a suitable *domination* relation between imputations) such that a set of imputations is a vNM-solution in the n-person game if and only if it is a stable extension in the associated argumentation framework.

There are digraphs without kernels, for example $\vec{C}_3$. On the other hand, in a symmetric digraph (undirected graphs; each edge is viewed as a symmetric pair of arcs) any maximal (w.r.t. inclusion) stable set is a kernel (for argumentation frameworks this has been discussed in Coste-Marquis *et al.* [CMDM05b]). It follows that the number of kernels in a digraph with n vertices is between 0 and $O(3^{\frac{n}{3}})$ (the upper bound follows from the work of Moon and Moser [MM65] that determined the maximum number of maximal stable sets in a graph with n vertices).

The following example is from Fraenkel [Fra97], where kernels are studied in the context of combinatorial games. Let $D = (V, E)$ be the digraph with $V =$

$\{a_0,\ldots,a_{2n}\}$ and $E = \{(a_0,a_i)|i \in \{1,\ldots,n\} \cup \{(a_i,a_{n+i}),(a_{n+i},a_i)|i \in \{1,\ldots,n\}$ (see Figure 0.4).

Then D has 2^n kernels, namely $\{a_0,a_{n+1},\ldots,a_{2n}\}$ and $A \cup \{a_{n+1},\ldots,a_{2n}\} - \{a_{n+i}|a_i \in A\}$, for each $A \subseteq \{a_1,\ldots,a_n\}$, $A \neq \emptyset$.

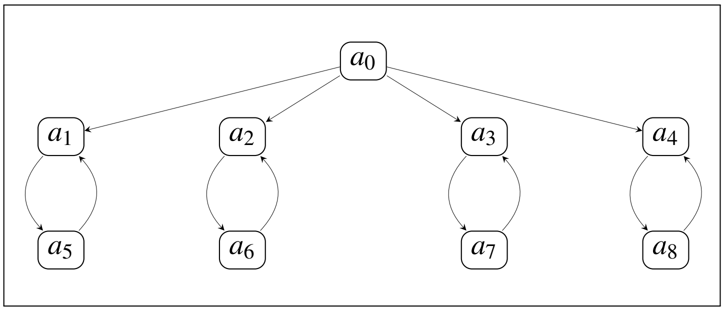


Figure 0.4.: $n = 4$: digraph with $2 \cdot n + 1$ vertices and 2^n kernels.

Note that in the corresponding argumentation framework, AF_D, any argument is credulous stable accepted (each argument belongs to a stable extension) but no argument is sceptically stable accepted: despite of the large number of stable extensions, each argument does not belong to at least one stable extension.

$\diamondsuit$

An acyclic directed graph D has a unique kernel S. This has been observed and proved firstly by von Neumann and Morgenstern [NM44] in the context of the theory of *combinatorial games*:

> *Given an acyclic digraph $D = (V,A)$ and a token in one of its vertices, let two players move this token along the arcs of D, alternating. In each step the token is moved from the current vertex v to some $v' \in v^+$. Since the graph is finite and acyclic, the token eventually will arrive to a dead end (i.e., to a vertex v for which $v^+ = \emptyset$). The player who's turn would be to move from a dead end is the looser of the game.*
>
> *Let S be the unique kernel of D. It is easy to see that the player who can start in a vertex $v \notin S$ has a winning strategy. Indeed, the player who moves from a vertex $v \notin S$ can always move the token into S. His opponent then either cannot move (being in a dead end, and hence loosing the game) or is forced to leave S again, since S is a stable set.*

The unique kernel S in an acyclic digraph D can be constructed efficiently, by recursively adding all dead ends to S and deleting those vertices and their neighbors from D. Translated into the converse digraph (the argumentation framework AF_D), this is exactly the algorithm to find the grounded extension, described after the Theorem 5, and this is the way we made the Observation 7.

Richardson [Ric53] proved that a digraph without odd directed cycles, has at least one kernel which can be determined in polynomial time. The algorithm can be described as follows:

> *Let $D = (V,E)$ such a digraph. Initially $K = \emptyset$. Repeat the following until no vertex remains in D: find the strongly connected components of D; since there are no odd directed cycles, each strongly connected component C_i is a bipartite digraph, say $C_i = (N_{i_1};N_{i_2},E_i)$; for each such component C_i with no incoming edges, select $j \in \{1,2\}$, set $K := K \cup N_{i_j}$, and delete from the digraph D the vertices in $N_{i_1} \cup N_{i_2} \cup N^{+}_{i_j}$.*

Translated into the converse digraph (the argumentation framework AF_D), the above algorithm can be considered the precursor of the general **SCC**-recursive schema for argumentation semantics introduced by Baroni, Giacomin, and Guida [BGG05].

If the digraph D is strongly connected then AF_D is strongly connected too, and the above algorithm gives two stable extensions. Figure 0.5 shows a strongly connected digraph without odd circuits (adapted from Dimopoulos, Magirou, and Papadimitriou [DMP97]) having another stable extension ($\{a_1,a_4,a_7\}$) besides the two stable extensions given by the above algorithm ($\{a_1,a_3,a_5,a_7\},\{a_2,a_4,a_6,a_8\}$).

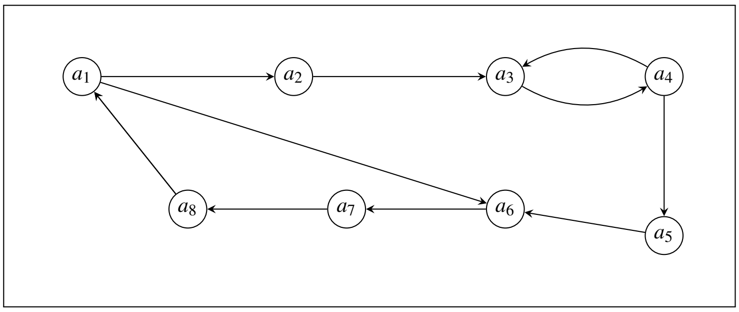


Figure 0.5.: A strongly connected even digraph with 3 stable extensions.

♢

A *semikernel* of a digraph D is any non-empty stable set S of D such that for any $z \in V(D) - S$ for which there exists an Sz-arc, there is also a zS-arc. Clearly, in the argumentation frameworks language, a semikernel is an admissible set in the converse digraph AF_D. Duchet [Duc80] defined a *kernel-perfect digraph* as a digraph D with the property that every induced subdigraph of D has a kernel. For example the circuit with 4 vertices, $\vec{C}_4$, is kernel perfect. Neumann-Lara [NL71] proved that *a digraph D is kernel perfect if and only if every induced subdigraph of D has a **semikernel***. Translating into the argumentation frameworks language, it follows that *an argumentation framework $AF = (A,D)$ has the property that each* *induced argumentation framework* $AF[A_1]$, *for* $A_1 \subseteq A$, *has a stable extension if and only if* $\mathbf{adm}(AF[A_1]) \neq \{\emptyset\}$, *for every* $A_1 \subseteq A$.

$\diamondsuit$

Competition graphs were introduced by Cohen [Coh68] in connection with problems of ecology. Let $D = (V, E)$ be a digraph. The *competition graph* of D is the undirected graph $C(D)$ with $V(C(D)) = V$ and $\{u, v\} \in E(C(D))$ if and only if for some $w \in V$, arcs (u, w) and (v, w) are in E. Hence two vertices are linked in $C(D)$ if and only if they have a common prey in D. The *common enemy graph* of D is $CE(D) = C(\tilde{D})$, ($\tilde{D}$ is the converse of D) that is, two vertices are linked in $CE(D)$ if and only if they have a common enemy in D.

We are interested in the partitions of the vertex set of D induced by the connected components of the graphs $C(D)$ and $CE(D)$. Some interesting facts arise. We discuss them in the argumentation frameworks language.

Definition 8 *Let* $AF = (A, D)$ *be an argumentation framework,* $A_0 = \{a \in A | a^- = \emptyset\}$ *and* $A_{fin} = \{a \in A | a^+ = \emptyset\}$. *We say that* $a, b \in A - A_0$ *are* **attacked-related** *if there are an integer* $p \geq 1$ *and arguments* $a_0, a_1, \ldots, a_p \in A - A_0$ *such that* $a_0 = a$, $a_p = b$ *and* $a_{i-1}^- \cap a_i^- \neq \emptyset$ *for all* $i \in \{1, \ldots, p\}$. *We denote by* $\rho^- \subseteq (A - A_0)^2$ *the binary relation* $\rho^- = \{(a, b) | a, b \in A - A_0, a \text{ and } b \text{ are attacked-related}\}$. *We say that* $a, b \in A - A_{fin}$ *are* **attacking-related** *if there are an integer* $p \geq 1$ *and arguments* $a_0, a_1, \ldots, a_p \in A - A_{fin}$ *such that* $a_0 = a$, $a_p = b$ *and* $a_{i-1}^+ \cap a_i^+ \neq \emptyset$ *for all* $i \in \{1, \ldots, p\}$. *We denote by* $\rho^+ \subseteq (A - A_{fin})^2$ *the binary relation* $\rho^+ = \{(a, b) | a, b \in A - A_{fin}, a \text{ and } b \text{ are attacking-related}\}$.

It is not difficult to see that ρ^- and ρ^+ are equivalent relations. Let $(A - A_0)/\rho^- = \{A_1, \ldots, A_k\}$ be the partition of $A - A_0$ into ρ^-- equivalent classes, and similarly $(A - A_{fin})/\rho^+ = \{B_1, \ldots, B_l\}$ be the partition of $A - A_{fin}$ into ρ^+- equivalent classes.

A simple way to compute these equivalent classes is to consider the undirected graphs $G^- = (A - A_0, E^-)$ and $G^+ = (A - A_{fin}, E^+)$, where $\{a, b\} \in E^-$ if and only if $a^- \cap b^- \neq \emptyset$ and $\{a, b\} \in E^+$ if and only if $a^+ \cap b^+ \neq \emptyset$. Then, the vertex sets of the connected components of G^- (G^+) are the ρ^-- equivalent classes $A_1, \ldots, A_k$ (ρ^+- equivalent classes $B_1, \ldots, B_l$).

Figure 0.6 depicts ρ- equivalent classes for a particular AF. Note that in this example each argument is attacked and each argument is attacking, hence $A - A_0 = A - A_{fin} = A$.

Proposition 9 *Let* $AF = (A, D)$ *be an argumentation framework with* $(A - A_0)/\rho^- = \{A_1, \ldots, A_k\}$ *and* $(A - A_{fin})/\rho^+ = \{B_1, \ldots, B_l\}$. *Then*

i) $(A_i^-)^+ = A_i$ *for each* $i \in \{1, \ldots, k\}$.

i') $(B_i^+)^- = B_i$ *for each* $i \in \{1, \ldots, l\}$.

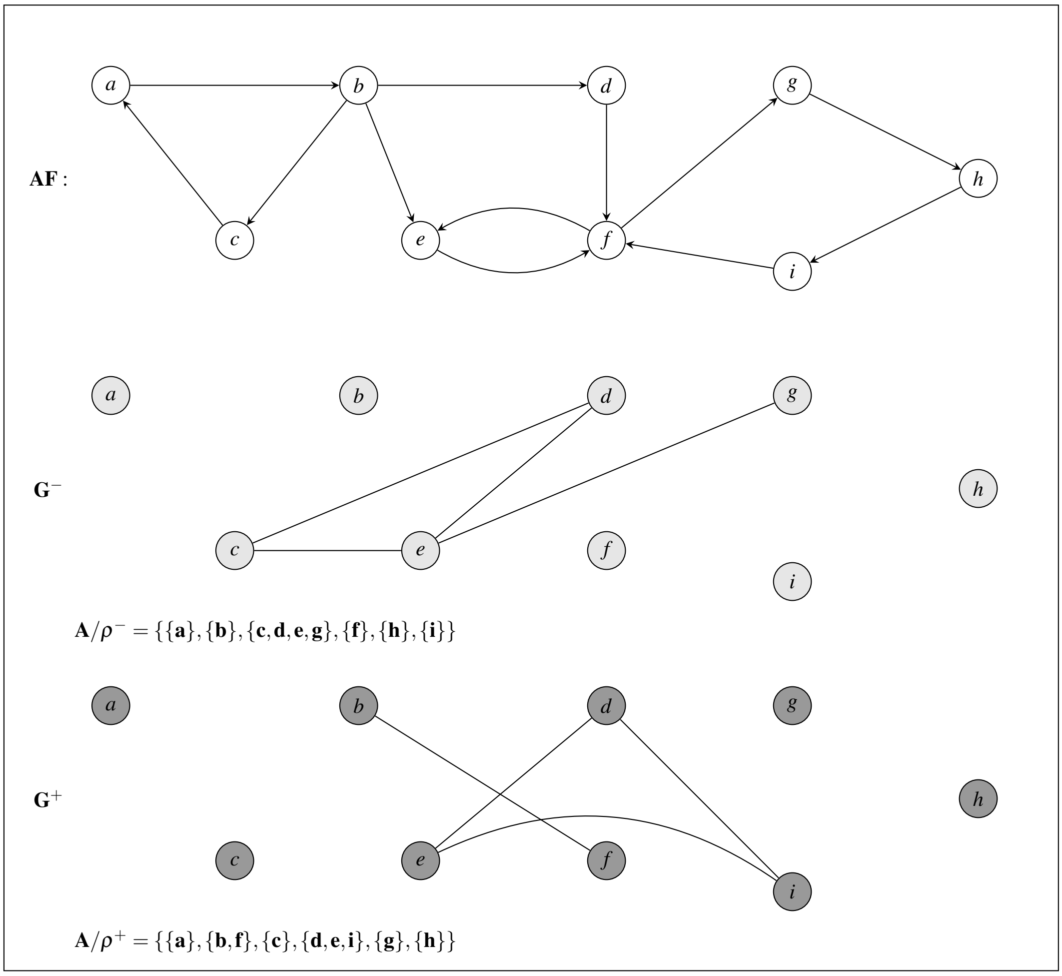


Figure 0.6.: An argumentation framework and its ρ-equivalent classes.

ii) $(A_1^-, \ldots, A_k^-)$ *is a partition of* $A - A_{fin}$.

ii') $(B_1^+, \ldots, B_l^+)$ *is a partition of* $A - A_0$.

iii) $k = l$ *and* $(A - A_{fin})/\rho^+ = \{A_1^-, \ldots, A_k^-\}$.

iii') $k = l$ *and* $(A - A_0)/\rho^- = \{B_1^+, \ldots, B_l^+\}$.

iv) *If* $D_i = \{(b,a) \in D | a \in A_i\}$, *then* $(D_1, \ldots, D_k)$ *is a partition of* D.

iv') *If* $D'_i = \{(b,a) \in D | b \in B_i\}$, *then* $(D'_1, \ldots, D'_k)$ *is a partition of* D.

v) $(D_1, \ldots, D_k) = (D'_1, \ldots, D'_k)$.

Proof. We prove only i)-iv), the proof of i')-iv') is similar. Also, v) is an obvious consequence of iii) and iii').

i) If $a \in A_i$ then $a^- \neq \emptyset$, hence $a \in (a^-)^+ \subseteq (A_i^-)^+$. It follows that $A_i \subseteq (A_i^-)^+$. Conversely, if $b \in (A_i^-)^+$ there is $c \in A_i^-$ such that $c \in b^-$. Since $c \in A_i^-$, there is $a \in A_i$ such that $c \in a^-$. Hence $c \in a^- \cap b^-$, that is $(a,b) \in \rho^-$. Since A_i, is a ρ^--equivalent class, it follows that $b \in A_i$. Therefore $(A_i^-)^+ \subseteq A_i$.

ii) Since $A_i \subseteq A - A_0$ it follows that $A_i^- \neq \emptyset$ for each $i \in \{1,\ldots,k\}$. If $a \in A - A_{fin}$, then $a^+ \neq \emptyset$, hence there is $b \in A - A_0$ such that $(a,b) \in D$. Since $(A_1,\ldots,A_k)$ is a partition of $A - A_0$, there is $i \in \{1,\ldots,k\}$ such that $b \in A_i$, that is $a \in A_i^-$. Consequently $\bigcup_{i=1,k} A_i^- = A - A_{fin}$. Suppose that there are $i,j \in \{1,\ldots,k\}$ such that $i \neq j$ and $A_i^- \cap A_j^- \neq \emptyset$. Then there are $a_i \in A_i$, $a_j \in A_j$ and $b \in A_i^- \cap A_j^-$ such that $b \in a_i^- \cap a_j^-$. It follows that $a_i \rho^- a_j$, hence $A_i = A_j$, contradicting the hypothesis.

iii) From ii') we have that $(A_1,\ldots,A_k)$ and $(B_1^+,\ldots,B_l^+)$ are two partitions of $A - A_0$. Suppose that there is $i \in \{1,\ldots,l\}$ such that $B_i^+ \cap A_{j_1} \neq \emptyset$ and $B_i^+ \cap A_{j_2} \neq \emptyset$ for some $j_1, j_2 \in \{1,\ldots,k\}$, $j_1 \neq j_2$. Let $b_{i_1} \in B_i \cap A_{j_1}^-$ and $b_{i_2} \in B_i \cap A_{j_2}^-$. Since $b_{i_1}, b_{i_2} \in B_i$, there are $c_0,\ldots,c_p$ such that $c_0 = b_{i_1}$, $c_p = b_{i_2}$ and $c_{i-1}^+ \cap c_i^+ \neq \emptyset$ for each $i \in \{1,\ldots,p\}$. If for some $i \in \{1,\ldots,p\}$ we have $c_{i-1} \in A_{j_1}^-$ then, since $(A_{j_1}^-)^+ = A_{j_1}$ (by i)), and $c_{i-1}^+ \cap c_i^+ \neq \emptyset$, we obtain that $c_i \in A_{j_1}^-$. Because $c_0 = b_{i_1} \in A_{j_1}^-$, it follows that $b_{i_2} \in A_{j_1}^-$. But then, $b_{i_2} \in A_{j_1}^- \cap A_{j_2}^-$, contradicting ii).
Hence, we proved that for each $i \in \{1,\ldots,l\}$ there is $j \in \{1,\ldots,k\}$ such that $B_i^+ \subseteq A_j$. Similarly, we can prove that for each $j \in \{1,\ldots,k\}$ there is $i \in \{1,\ldots,l\}$ such that $A_j \subseteq B_i^+$. It follows that iii) holds.

iv) Let $(b,a) \in D$. Since $(A_1,\ldots,A_k)$ is a partition of $A - A_0$ it follows that there is $i \in \{1,\ldots,k\}$ such that $a \in A_i$, therefore $(b,a) \in D_i$. Hence $\bigcup_{i=1,k} D_i = D$. Since $A_i^- \neq \emptyset$ for each $i \in \{1,\ldots,k\}$, it follows that $D_i \neq \emptyset$ for each $i \in \{1,\ldots,k\}$. If there are $i,j \in \{1,\ldots,k\}$, $i \neq j$ such that $D_i \cap D_j \neq \emptyset$ it follows that $A_i^- \cap A_j^- \neq \emptyset$ contradicting ii). □

This is a nice and somewhat unexpected property of the connected components of the competition and common enemy graphs associated to an argumentation framework. Note that, essentially, the same decomposition is obtained by Liu and West [LW98], where *coreflexive* vertex sets are considered (a *coreflexive set* in $D = (V,E)$ is either the set of sinks in D (A_{fin} above) or a minimal nonempty set U such that $(U^+)^- = U$ (as in i') above).

◇

A *quasi-kernel* in the digraph $D = (V,E)$ is a stable set Q in D such that for every $v \in V - Q$ there is $x \in Q$ such that $(v,x) \in E$ or there is $w \in V - Q$ such that $(v,w),(w,x) \in E$, i.e. Q is non-empty, there are no edges inside Q, and every vertex outside Q can reach Q in at most two hops. Clearly, any kernel is a quasi-kernel. For example, if D is the path $a \to b \to c$, then its non-empty stable sets are

$\{a\}, \{b\}, \{c\}$, and $\{a,c\}$. $\{a\}$, and $\{b\}$ are neither kernels nor quasi-kernels, $\{c\}$ is a quasi-kernel but not a kernel, and $\{a,c\}$ is a kernel (and also a quasi-kernel).

Note that the concept of quasi-kernel has no significance in an argumentation framework: in AF_D, a quasi-kernel is a non-empty conflict-free set of arguments Q, such that any vertex outside Q is in $Q^+ \cup Q^{++}$. However we present here some interesting properties of quasi-kernels from [Cro15b] since a quasi-kernel corresponds to a well-known solution in *Social Choice Theory* (uncovered sets and "two-step principle", see, e.g., Duggan [Dug12]).

Chvátal and Lovász [CL74] observed that every digraph has a quasi-kernel. We describe a simple algorithmic proof of this (by making explicit the construction of the two kernels in the acyclic digraphs used in the proof given by Thomassé, cf. Bondy [Bon03]).

Let n be the cardinality of V. For an arbitrary ordering π of V, i.e. an injective mapping $\pi : [1..n] \rightarrow V$, and $i \in [1..n]$, let

$$L_i(\pi) = \{\pi(j) | (\pi(i), \pi(j)) \in E \text{ and } j < i\}$$

$$H_i(\pi) = \{\pi(j) | (\pi(i), \pi(j)) \in E \text{ and } j > i\}$$

be the sets of lower and higher numbered out-neighbours of the i-th vertex of V under π. The following algorithm constructs a quasi-kernel in D.

Function $QK(D,\pi)$
$Q := \emptyset$
Forward: **for** $i = 1$ **to** n **do** **if** $Q \cap L_i(\pi) = \emptyset$ **then** $Q := Q \cup \{\pi(i)\}$
Backward: **for** $i = n$ **to** 1 **do** **if** $[\pi(i) \in Q \text{ and } Q \cap H_i(\pi) \neq \emptyset]$ **then** $Q := Q - \{\pi(i)\}$
Return(Q)

In the forward scan, $\pi(1)$ is added to Q, and vertices which are not added to Q have a smaller (with respect to π) out-neighbor in Q. In the backward scan, $last(Q,\pi)$, the last vertex (with respect to π) of Q, remains in Q and vertices of Q having a larger (with respect to π) out-neighbour in Q are deleted from Q.
It follows that, after the two scans, Q is a non-empty stable set in D and $V - Q$ is the disjoint union of two sets $R_1(\pi) \dot\cup R_2(\pi)$, where each vertex in $R_1(\pi)$ has (at least) one out-neighbour in Q and each vertex in $R_2(\pi)$ has no out-neighbours in Q. $R_1(\pi)$ is the set of vertices deleted from Q in the backward scan or not added to Q in the forward scan and having an out-neighbour in Q after the backward scan. A vertex v is in $R_2(\pi)$ if it was not added to Q in the forward scan (when it has at least one out-neighbour, w, in the current Q) but after the backward scan all its out-neighbours are out of Q. Clearly, the vertex w is in $R_1(\pi)$ after the two scans, hence there is $x \in Q$ such that $(v,w),(w,x) \in E$. Therefore Q is a quasi-kernel in D. Note that if $R_2(\pi) = \emptyset$ then Q is a kernel.

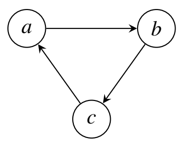
A natural question is if every quasi-kernel Q in D can be constructed using the above algorithm. The answer is no, as the following example shows: let D be the digraph obtained from the undirected path $a — b — c$ by replacing each undirected edge with a symmetric pair of directed edges; then $\{b\}$ and $\{a,c\}$ are kernels in D and $\{a\}$ and $\{c\}$ are quasi-kernels; however, it is easy to check that $QK(D,\pi) \neq \{a\}$ (or $\{c\}$), for every ordering π of $\{a,b,c\}$.

We will next show that if D has no kernel then D has at least three quasi-kernels. Assume that D has no kernel and let π_1 be an ordering of V and $Q_1 := QK(D,\pi_1)$. Since D has no kernel, $R_2(\pi_1) \neq \emptyset$.

Let π_2 be the ordering of V starting with vertices in Q_1, followed by vertices in $R_1(\pi_1)$, and ending with vertices in $R_2(\pi_1)$ (vertices in the same set are ordered arbitrarily). Let $Q_2 := QK(D,\pi_2)$. In the forward scan of the construction of Q_2, all vertices in Q_1 are added to Q_2, vertices in $R_1(\pi_1)$ are not added, and the vertex $last(Q_2,\pi_2)$ is from $R_2(\pi_1)$. Therefore $last(Q_2,\pi_2) \in Q_2 - Q_1$, and Q_2 is a quasi-kernel in D with the property that $Q_2 \neq Q_1$. Since D has no kernel, $R_2(\pi_2) \neq \emptyset$.

Let π_3 be the ordering of V starting with vertices in Q_2, followed by vertices in $R_1(\pi_2)$, and ending with vertices in $R_2(\pi_2)$. Let $Q_3 := QK(D,\pi_3)$. As above, the last vertex in Q_3 does not belong to Q_2 and therefore Q_3 is a quasi-kernel in D with the property that $Q_3 \neq Q_2$. We have also $Q_3 \neq Q_1$ since the last vertex in Q_3 is from $R_2(\pi_2)$ which contains only vertices from $R_1(\pi_1)$ hence not from Q_1.

We have obtained that if D has no kernel then D has at least three quasi-kernels, a result discovered by Jacob and Meyniel ([JM96]). The existence of four quasi-kernels cannot be established in this way since another application of the argument may reproduce Q_1, as the Figure 0.7 shows.



D

$\pi_1 : a,b,c$

After Forward scan: $Q_1 = \{a,b\}$

After Backward scan: $\mathbf{Q_1} = \{\mathbf{b}\}$; $R_1(\pi_1) = \{a\}$; $R_2(\pi_1) = \{c\}$

$\pi_2 : b,a,c$

After Forward scan: $Q_2 = \{b,c\}$

After Backward scan: $\mathbf{Q_2} = \{\mathbf{c}\}$; $R_1(\pi_2) = \{b\}$; $R_2(\pi_2) = \{a\}$

$\pi_3 : c,b,a$

After Forward scan: $Q_3 = \{c,a\}$

After Backward scan: $\mathbf{Q_3} = \{\mathbf{a}\}$; $R_1(\pi_3) = \{c\}$; $R_2(\pi_3) = \{b\}$

$\pi_4 : a,c,b$

After Forward scan: $Q_4 = \{a,b\}$

After Backward scan: $\mathbf{Q_4} = \{\mathbf{b}\} = \mathbf{Q_1}$!!

Figure 0.7.: Applying successively the algorithm on a digraph with no kernel.

If a vertex v of the digraph D has no out-neighbour (it is a sink) then, clearly, v belongs to every quasi-kernel of D. If v has an out-neighbour w (that is, $(v,w) \in E$) then, considering any ordering π starting with w, $Q = QK(D,\pi)$ is a quasi-kernel not containing v (v is not selected in Q in the Forward scan).

However, deciding if there exists a quasi-kernel in D containing the vertex v is a NP-complete problem. Since the membership to NP is obvious, we prove the hardness of this decision problem by exhibiting a polynomial time reduction from the CNF satisfiability problem **SAT**.

Let $F = C_1 \wedge \ldots \wedge C_m$ be an arbitrary instance of **SAT**, where for each $i \in \{1,\ldots,m\}$, the clause C_i is a a disjunction of literals, $C_i = l_{i_1} \vee \ldots \vee l_{i_{k_i}}$, and a literal l is either a variable x_j or its negation $\bar{x}_j$, where $\{x_1,\ldots,x_n\}$ is the set of boolean variables occurring in F. We construct in polynomial time the digraph $D_F = (V_F, E_F)$ and a vertex $v_F \in V_F$, such that F is satisfiable if and only if D_F has a quasi-kernel containing v_F.

Let $V_F = \{v_F\} \dot{\bigcup}_{i=1,m} \{u_{C_i}, v_{C_i}\} \dot{\bigcup}_{i=1,n} \{v_{x_i}, v_{\bar{x}_i}\}$, $\quad$ $(2n+2m+1$ vertices) and
$E_F = \bigcup_{i=1,m} \{(v_F,u_{C_i}),(v_F,v_{C_i}),(u_{C_i},v_{C_i})\} \bigcup_{i=1,n} \{(v_{x_i},v_{\bar{x}_i}),(v_{\bar{x}_i},v_{x_i})\} \bigcup$
$\bigcup_{i=1,n} \bigcup_{j=1,m} \{(v_{C_j},v_{x_i}) | x_i \text{ is literal in } C_j\} \bigcup$
$\bigcup_{i=1,n} \bigcup_{j=1,m} \{(v_{C_j},v_{\bar{x}_i}) | \bar{x}_i \text{ is literal in } C_j\}$, $\quad$ $(3m+2n+3size(F)$ edges).
Clearly, D_F can be constructed in polynomial time in the size of the instance of **SAT**. An example of this construction is illustrated in Figure 0.8.

It is not difficult to see that any quasi-kernel in D_F containing v_F is of the form $Q = \{v_f\} \bigcup S$, where S selects for each variable x_i exactly one of the vertices v_{x_i} or $v_{\bar{x}_i}$ such that for each clause C_j, the vertex v_{C_j} has at least one out-neighbour in S. But this is equivalent to the fact that F is satisfiable.

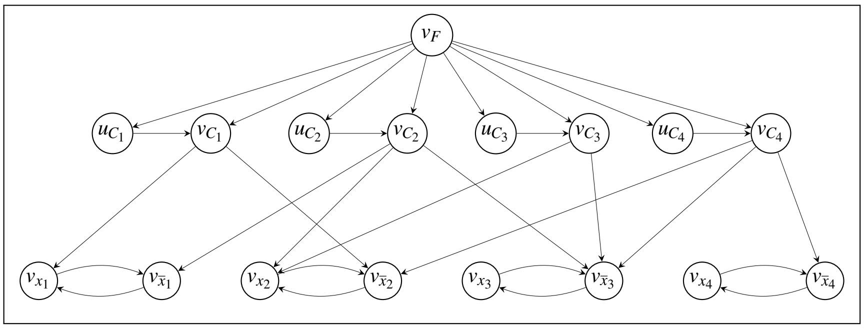


Figure 0.8.: The digraph $D(F)$ associated to instance $F = C_1 \wedge C_2 \wedge C_3 \wedge C_4$, where $C_1 = x_1 \vee \bar{x}_2$, $C_2 = \bar{x}_1 \vee x_2 \vee \bar{x}_3$, $C_3 = x_2 \vee \bar{x}_3$, $C_4 = \bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4$.

It follows that there is no polynomial time algorithm to list all quasi-kernels of a given digraph, unless **P=NP** (otherwise, we can decide in polynomial time if there is a quasi-kernel containing a specified vertex, by testing its membership to each quasi-kernel produced).

If a digraph $D = (V,E)$ has exactly one quasi-kernel Q, then (as noted above) Q contains all the sinks of D and each vertex of Q is a sink. Moreover Q is a kernel, otherwise taking the ordering π of V starting with Q, followed by a vertex with no out-neighbour in Q, and ending with the the remaining vertices, $QK(D,\pi)$ is quasi-kernel having at least one vertex not in Q, so D has at least two quasi-kernels.

Hence

> *a digraph has exactly one quasi-kernel if and only if its set of sinks is a kernel.*

Let $D=(V,E)$ be a digraph with exactly two quasi-kernels Q_1 and Q_2. Let π_1 be the ordering of V starting with vertices in Q_1-Q_2, followed by the vertices in $Q_1\cap Q_2$, the vertices of Q_2-Q_1, and ending with vertices in $V-(Q_1\cup Q_2)$. Then $Q:=QK(D,\pi_1)\in\{Q_1,Q_2\}$. This means that each vertex $v\in V-(Q_1\cup Q_2)$ has an out-neighbour in Q (otherwise, $last(Q,\pi_1)\in V-(Q_1\cup Q_2)$). Also, each vertex of $\{Q_1,Q_2\}-\{Q\}$ which is not in $Q_1\cap Q_2$ has an out-neighbour in Q (since it is not added in the forward scan or has been eliminated in the backward scan). Hence *if a digraph has exactly two quasi-kernels then at least one of them is a kernel.*

If D has exactly two quasi-kernels Q_1 and Q_2 then $Q_1\cap Q_2$ is exactly the set S of sinks in D. Clearly, $S\subseteq Q_1\cap Q_2$. If there is $v\in(Q_1\cap Q_2)-S$, let w an out-neighbour of v, and consider any ordering π of V having w and v as first two vertices. Then $Q:=QK(D,\pi)\notin\{Q_1,Q_2\}$ ($v\in Q_1\cap Q_2$ is not added to Q in the forward scan).

To simplify our discussion suppose that $D=(V,E)$ has exactly two quasi-kernels Q_1 and Q_2 and has no sinks. By the above remarks, we can suppose that Q_1 is a kernel and $Q_1\cap Q_2=\emptyset$.

Let $w_2\in Q_2$ be an arbitrary vertex of Q_2. Since Q_1 is a kernel, w_2 has an out-neighbour $v_2\in Q_1$. Then, the set of out-neighbours of each vertex in $Q_1-\{v_2\}$ is $\{w_2\}$. Indeed, suppose that there is $w_1\in Q_1-\{v_2\}$ with $(w_1,v_1)\in E$ and $v_1\neq w_2$. If $(v_2,v_1)\notin E$, then let π be the ordering of V starting with v_1,v_2,w_1,w_2. Clearly, $Q:=QK(D,\pi)\notin\{Q_1,Q_2\}$ (w_1 and w_2 are not added to Q in the forward scan). Similarly, if $(v_1,v_2)\notin E$ then the ordering of V starting with v_2,v_1,w_1,w_2 will give a contradiction. Hence $(v_1,v_2),(v_2,v_1)\in E$ ($\{v_1,v_2\}$ is a 2-cycle in D). Let $D'=(V,E')$ the digraph obtained from D by deleting (v_1,v_2). Then, as above, $Q':=QK(D',\pi)\notin\{Q_1,Q_2\}$. However, since $(v_2,v_1)\in E'$, Q' is also a quasi-kernel in D, a contradiction. Since the vertex w_2 was arbitrary in Q_2, in order that the set of out-neighbours of each vertex in $Q_1-\{v_2\}$ is $\{w_2\}$, it is necessary that either $|Q_2|=1$ or $|Q_1|=|Q_2|=2$ and $Q_1\cup Q_2$ induces a 4-cycle in D. In both cases Q_2 is a kernel. Indeed, if $|Q_1|=2$ then there is $v_1\in Q_1$ (in the 4-cycle) having an out-neighbour in Q_2. If $|Q_1|=1$, let $Q_1=\{v_1\}$ and $(v_1,z)\in E$ (v_1 is not a sink). Then each vertex $v\in V-\{v_1,z\}$ reaches z in two hops (via v_1). Hence $\{z\}$ is a quasi-kernel in D and therefore $Q_2=\{z\}$ and again there is $v_1\in Q_1$ having an out-neighbour in Q_2. If Q_2 is not a kernel then there is $v_2\in V-Q_2$ having no out-neighbour in Q_2. Let π the ordering of V starting with vertices in Q_2, followed by v_1,v_2, and ending with the remaining vertices. Then $Q:=QK(D,\pi)$ does not contain v_1 (hence is not Q_1) and has at least a vertex not in Q_2 (since v_2 is added to Q in the forward scan, it follows that $last(Q,\pi)\notin Q_2$). Hence $Q\notin\{Q_1,Q_2\}$, contradiction.

Since Q_2 is a kernel, it follows (by repeating the argument above for Q_1) that $Q_1\cup Q_2$ induces a 2-cycle or a 4-cycle in D, having no out-neighbors in $V-(Q_1\cup$

Q_2), and each vertex in $V-(Q_1 \cup Q_2)$ has at least two adjacent out-neighbours in $Q_1 \cup Q_2$. Since a strongly connected digraph has no sinks, it follows that

> *a strongly connected digraph D of order at least three has at least three quasi-kernels, unless D is* $\vec{C}_4$,

a result proved by Gutin, Koh, Tay and Yeo [GKTY04].

Part I.

Combinatorial Aspects of Argumentation Frameworks

This part presents my approaches to the combinatorial structure of argumentation frameworks for better modeling the acceptability of arguments, for simplifying the structure of attacks, and for introducing new argument labellings. Preliminary results in this direction have been discussed in my Master Thesis. While Chapter 1 makes only a review of a new type of acceptability of arguments on the basis of a given argumentation framework (based on [Cro12] and [CK12]), the next two chapters contain improvements and extensions of the work initiated in my Master Thesis as these appears in [CK13] and [Cro14b].

1. Deliberative Acceptability of Arguments

In this chapter we introduce a new type of acceptability of an argument, in which its attacking and defending sets of arguments are uniformly treated. We call it *deliberative acceptance*, discuss how this and the classical acceptance notions interrelate and analyze its computational properties. In particular, we prove that the corresponding decision problem is Π_2^{P}-complete, but its restrictions on bipartite or co-chordal argumentation frameworks are in P.

1.1. Introduction

Discussions on how to select the appropriate semantics for a given application context or how to compare the different semantics for argumentation frameworks, have been the subject of several papers (see Baroni and Giacomin [BG07a], [BG07b], Dunne and Bench-Capon [DBC01], Cayrol, Doutre, Lagasquie-Schiex, and Mengin [CDLSM02], Prakken and Vreeswijk [PV02], etc.). The motivation of considering different semantics (*semi-stable*, Caminada [Cam06b], *prudent*, Coste-Marquis, Devred, and Marquis [CMDM05a], *ideal*, Dung, Mancarella, and Toni [DMT06], *SCC-recursive*, Baroni and Giacomin. [BG03], *evidence-based* Oren and Norman [ON08], etc.) is given by the need to formalize "everyday reasoning" in order to design mechanisms expressing a "legitimate" argumentation in support of an argument.

The common principle of argument acceptance in a given argumentation framework is to identify collectively acceptable arguments, that is, arguments a belonging to a σ-extension. Each σ-extension has at least the properties of being conflict-free (that is, they are internally logically coherent) and defend themselves from any external attack (that is, they survive the attacks together).

This type of argument acceptance is biased toward attacks: while the defending extension of an argument a is internally coherent, no such requirement is imposed on its attacking set. In many cases in the real world, when argument a is attacked by a coherent set of arguments T, a is defended by giving a coherent set S including a, which defends against T. Having only one defending extension for all these attacking sets would contradict the deliberative nature of argumentation in the real world disputes and debates, where only the coherent sets of attacks matter and the defending sets of arguments depend on the former.

In this chapter we follow this intuition and introduce a notion of acceptability at the level of justification states of arguments rather than of extensions. We call it *deliberative acceptability* due to the uniform treatment of both attacking and defending sets in the definition of the acceptability of an argument. More precisely, an argument a is *deliberatively acceptable* in a given argumentation framework if, for each conflict-free set of arguments attacking a, there is a conflict-free set of arguments containing a and defending itself against the former set. We analyze this type of acceptability and investigate it from a complexity point of view. We prove that if an argument is credulously grounded, preferred or stable accepted then it is deliberatively accepted, but the converse is not true. This is illustrated in Figure 1.1.

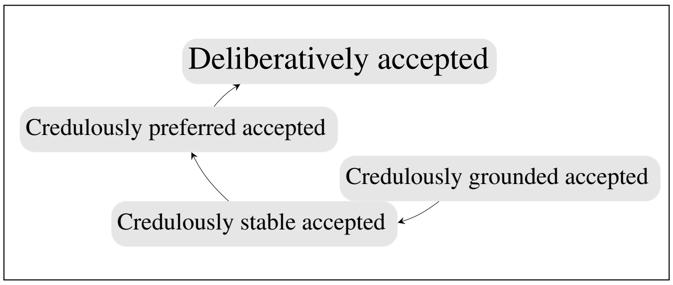


Figure 1.1.: Acceptability implications (edges implied by transitivity are omitted).

We give a sufficient condition in which a deliberatively accepted argument belongs to an admissible set (see Proposition 14). Also, we show that in bipartite or symmetric argumentation frameworks the deliberative acceptability is equivalent to credulous preferred acceptability (Proposition 15). In Proposition 16, we prove that the problem of deciding if an argument is deliberatively accepted in a given argumentation framework is Π_2^{P}-complete but its restrictions on bipartite or co-chordal argumentation frameworks are polynomial time solvable (Propositions 17 and 18).

We close this section by considering the following illustrative example.

Example 10 *Let us consider the argumentation framework in the Figure 1.2 below.*

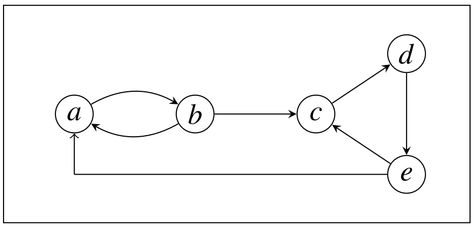


Figure 1.2.: Argumentation Framework in Example 10

Since each argument has at least one attacker, the grounded extension is empty. The conflict-free sets containing the argument a are $\{a\}$, $\{a,c\}$, and $\{a,d\}$. Since a

cannot defend against the attacker e, $\{a\}$ is not admissible. The set $\{a,c\}$ cannot defend against the attacker e, and the set $\{a,d\}$ cannot defend against the attacker c. It follows that a does not belong to a preferred extension. The unique preferred extension is $\{b,d\}$ (its attackers are defeated by b). Therefore we could accept b and not a, because b can be extended to a maximal conflict-free set of arguments $\{b,d\}$ defending itself from all attacks.

Despite of the fact that a is not credulously preferred accepted, its "acceptability" could be argued as follows. The conflict-free sets attacking a are: $\{b\}$, $\{b,d\}$ and $\{b,e\}$. Clearly, a defends itself against $\{b\}$. Against $\{b,d\}$, we can defend a by considering the conflict-free set $\{a,c\}$, that is choosing a coherent sets of arguments on the same idea induced by the attacking set $\{b,d\}$. The set $\{a,c\}$ is not appropriate for the attacking set $\{b,e\}$ but, in this case, $\{a,d\}$ does the job. Hence we can defend a against all conflict-free sets of arguments attacking it, therefore we can "accept" a. We will study this type of acceptability in the next section.

1.2. Deliberative Acceptability of Arguments

Definition 11 *Let $AF = (A,D)$ an argumentation framework. An argument $a \in A$ is* **deliberative acceptable** *if for any conflict-free set T attacking a (i.e. $T \cap a^- \neq \emptyset$), there is a conflict-free set $S \subset A$ such that $a \in S$ and S defends itself against T (i.e. $T \cap S^- \subseteq S^+$).*

This type of acceptability is closer to the intuition about real life debate-type argumentation. If the argument a is attacked by a conflict-free set T of arguments, there is a conflict-free set of arguments S containing a, depending on T such that S defends a. S also defends its arguments that are attacked by T. Figure 1.3 illustrates two conflict-free sets T_i attacking a and the corresponding two answers to these attacks, two conflict-free sets S_i containing a. Each attack of a from T_i is counterattacked by an argument from S_i. If T_i attacks also this defender, then S_i has also a counterattack to this. And so on.

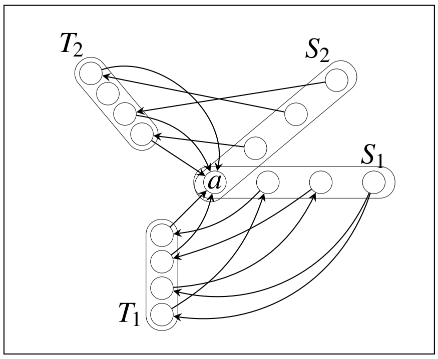


Figure 1.3.: Deliberative Acceptability Example.

Note that we assumed that there are no self-attacking arguments in the argumentation framework. If a would be a self-attacking argument, then a is not contained in any conflict-free set. Hence, if a is attacked by some other argument, then a cannot be deliberatively accepted. Moreover, since a does not belong to any conflict-free set, the deliberative acceptance of other arguments is not influenced. Thus, without loss of generality, we restrict our attention to argumentation frameworks without self-attacking arguments.

It is possible to introduce a weaker form of the deliberative acceptance by considering in the above definition only the conflict-free sets T contained in a^-. This could be too tolerant as the following example shows.

Example 12 *Let us consider the argumentation framework in the Figure 1.4 below.*

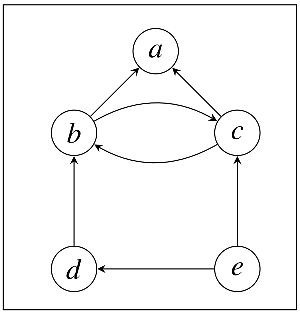


Figure 1.4.: Argument a is not deliberative acceptable, despite it can be defended against conflict-free sets contained in a^-.

The grounded extension is $\{e,b\}$. It follows that the argument a is not credulously accepted in any classical extension semantics. The conflict-free sets contained in a^- are $\{b\}$ and $\{c\}$ which can be defeated by the conflict-free sets containing a: $\{d,a\}$, respectively, $\{e,a\}$. If we consider a (weak) deliberative acceptance this will correspond to a superficial analysis. Indeed, there are also two conflict-free sets attacking a, namely $\{c,d\}$ and $\{b,e\}$. The set $\{e,a\}$ defeats $\{c,d\}$, but $\{e,b\}$ is a conflict-free set attacking a, which could not be defended. It follows that the argument a is not deliberatively accepted.

The following proposition shows that deliberative acceptance is strictly more liberal than credulous acceptance with respect to the classical extension-based semantics.

Proposition 13 *Let $AF = (A,D)$ be an argumentation framework and $a \in A$. If a is σ-credulously accepted for $\sigma \in \{$ grounded, preferred, stable$\}$, then a is deliberatively accepted. For each of the above semantics σ, there are argumentation frameworks in which a deliberatively accepted argument is not σ-credulously accepted.*

Proof. If $a \in A$ is σ-credulously accepted for $\sigma \in \{\text{grounded, preferred, stable}\}$, then there is an admissible set S_0 containing a. It follows that for any $b \in S_0^-$ there is $s \in S_0$ such that $(s,b) \in D$. Let T be a conflict-free set of arguments attacking a. Each argument $b \in T \cap S_0^-$ is attacked by S_0. Hence $T \cap S_0^- \subseteq S_0^+$, and therefore a is deliberative accepted.
In the argumentation framework in Figure 1.2 the grounded extension is $\emptyset$. Hence a is not grounded-credulous accepted. However, as we argued in the end of Section 1.2, a is deliberative accepted. Also, a is not preferred accepted because the unique preferred extension $\{b,d\}$ does not contain a. Since any stable extension is a preferred extension, it follows that a is also not stable-credulously accepted. □

Depending on the combinatorial structure of the argumentation framework, the deliberative acceptance could agree to that based on extensions. The following proposition gives an easy sufficient condition for this.

Proposition 14 *Let $AF = (A,D)$ be an argumentation framework and $a \in A$. If AF does not contain the induced subdigraphs F_1 and F_2 in Figure 1.5 (dotted directed edges are optional), then, if a is deliberatively accepted, it follows that a is credulously preferred accepted.*

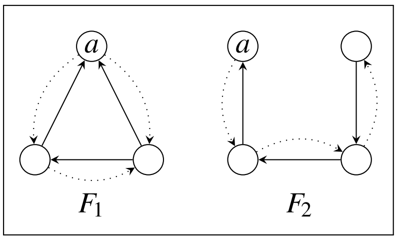


Figure 1.5.: Forbidden induced subdigraphs F_1 and F_2.

Proof. See [CK12].

Using similar arguments as in the above proof, it is not difficult to show that if AF is a *directed acyclic graph* (DAG) or if the underlying undirected graph of AF is bipartite, then an argument a is deliberatively accepted if and only if a is credulously preferred accepted. However, for bipartite graphs we can do better as follows.

Let $G = (A,E)$ be the underlying undirected graph of AF ($\{a,b\} \in E$ if and only if $(a,b) \in D$ or $(b,a) \in D$). Since G is bipartite, then A can be partitioned $A = U \cup V$, $U,V \neq \emptyset$, $U \cap V = \emptyset$, and if $\{a,b\} \in E$ then $|\{a,b\} \cap V| = 1$ and $|\{a,b\} \cap U| = 1$.

In [Dun07], Dunne proved that the following algorithm applied to a bipartite argumentation framework $AF = (U \cup V, D)$ (below, $[X]_{AF}$ denotes the subdigraph induced by $X \subseteq A = U \cup V$ in AF: $[X]_{AF} = (X, D \cap X \times X)$).

input $AF = (U \cup V, D)$
while $\exists v \in V - U^+$ s.t. $v^+ \cap U \neq \emptyset$ **do**
  $AF := [(U - v^+) \cup V]_{AF}$
return U

returns a final set U, denoted U_0, which satisfies: U_0 is conflict-free (it is a subset of U), and $U_0^- \subseteq U_0^+$ (in the given argumentation framework AF). Similarly, the algorithm:

input $AF = (U \cup V, D)$
while $\exists u \in U - V^+$ s.t. $u^+ \cap V \neq \emptyset$ **do**
  $AF := [U \cup (V - u^+)]_{AF}$
return V

returns a final set V, denoted V_0, which satisfies: V_0 is conflict-free (it is a subset of V), and $V_0^- \subseteq V_0^+$. Moreover, $a \in A$ is credulous preferred accepted if and only if $a \in U_0$ or $a \in V_0$ (Dunne, [Dun07]). It is not difficult to prove that the following proposition holds.

Proposition 15 *Let $AF = (A, D)$ be an argumentation framework and $a \in A$. If the underlying undirected graph of AF is bipartite with bipartition $A = U \cup V$, and U_0 and V_0 are the sets of arguments constructed by the Dunne's algorithm above, then a is deliberatively accepted if and only if $a \in U_0 \cup V_0$.*

Proof. See [CK12].

We note that if the argumentation framework is symmetric (that is, AF is obtained from an undirected graph by replacing each undirected edge $\{a, b\}$ by the pair (a, b) and (b, a) of directed edges), then each argument is deliberative accepted. This is obvious since, in this case, a set of arguments is admissible if and only if it is a maximal conflict-free set, [CMDM05a].

1.3. Complexity

Let us consider the following decision problem:

Deliberative Acceptability
Instance: $AF = (A, D)$ argumentation framework, $a \in A$.
Question: Is a deliberatively acceptable?

The complexity class Π_2^P comprises those problems decidable by co-NP computations given (unit cost) access to an NP complete oracle. Alternatively, Π_2^P can be viewed as the class of languages L whose membership is certified by a polynomial-time testable ternary relation $R_L \subseteq W \times X \times Y$: there is a polynomial p such that, for all w, $w \in L$ if and only if $(\forall x \in X : |x| \leq p(|w|))(\exists y \in Y : |y| \leq p(|w|))\ (w, x, y) \in R_L$.

Theorem 16 **Deliberative acceptability** *is* Π_2^{P}*-complete.*

Proof. It is easy to see that **Deliberative acceptability** is in Π_2^{P}, since it corresponds to the language $L = \{w|\forall x \exists y R(w,x,y)\}$, where w encodes an instance (AF,a) of the problem and $(w,x,y) \in R$ if and only if x encodes a conflict-free set T attacking a, and y encodes a conflict-free set S containing a such that S defends itself against T.

We prove Π_2^{P}-hardness for **Deliberative acceptability** by a reduction from the decision problem $\forall\exists$SAT.
An instance of $\forall\exists$SAT is a formula $F' = \forall x_1 \ldots \forall x_n \exists y_1 \ldots \exists y_n F(x_1,\ldots,x_n,y_1,\ldots,y_n)$, where $F(x_1,\ldots,x_n,y_1,\ldots,y_n)$ is a CNF formula over the disjoint sets of variables $X = \{x_1,\ldots,x_n\}$ and $Y = \{y_1,\ldots,y_n\}$. The instance $F' = \forall X \exists Y F(X,Y)$ is accepted if and only if for any truth assignment α_X of the variables in X, there is a truth assignment α_Y of the variables in Y such that (α_X, α_Y) satisfies the formula $F(X,Y)$. It is well known that $\forall\exists$SAT is Π_2^p-complete (Stockmeyer and Meyer [SM73], Wrathall [Wra76]).

We will construct in polynomial time an argumentation framework $AF_{F'}$ for each instance F' of $\forall\exists$SAT. Let $F' = \forall X \exists Y F(X,Y)$ and $F(X,Y) = C_1 \wedge \ldots \wedge C_m$, where each clause C_i is a disjunction of literals. A literal is a variable $x_i \in X = \{x_1,\ldots,x_n\}$, $y_i \in Y = \{y_1,\ldots,y_n\}$, or their negations $\overline{x}_i \in \overline{X} = \{\overline{x_1},\ldots,\overline{x}_n\}$, $\overline{y}_i \in \overline{Y} = \{\overline{y_1},\ldots,\overline{y}_n\}$.
The argumentation framework associated to F' is $AF_{F'} = (A,D)$, where:
- $A = \{F\} \cup \{C_1,\ldots,C_m\} \cup X \cup \overline{X} \cup Y \cup \overline{Y}$, and
- $D = \cup_{i=1}^m \{(C_i,F)\} \bigcup$

$$\cup_{i=1}^m \cup_{j=1}^n \{(x_j,C_i)|x_j \text{ occurs in } C_i\} \bigcup \cup_{i=1}^m \cup_{j=1}^n \{(\overline{x}_j,C_i)|\overline{x}_j \text{ occurs in } C_i\} \bigcup$$
$$\cup_{i=1}^m \cup_{j=1}^n \{(y_j,C_i)|y_j \text{ occurs in } C_i\} \bigcup \cup_{i=1}^m \cup_{j=1}^n \{(\overline{y}_j,C_i)|\overline{y}_j \text{ occurs in } C_i\} \bigcup$$
$$\cup_{j=1}^n \{(x_j,\overline{x}_j),(\overline{x}_j,x_j),(y_j,\overline{y}_j),(\overline{y}_j,y_j)\} \bigcup$$
$$\cup_{i=1}^n \cup_{j=1}^n \{(x_i,y_j),(x_i,\overline{y}_j)\} \bigcup \cup_{i=1}^n \cup_{j=1}^n \{(y_i,x_j),(y_i,\overline{x}_j)\} \bigcup$$
$$\cup_{i=1}^n \cup_{j=1}^m \{(C_j,y_i),(C_j,\overline{y}_i)\} \bigcup$$
$$\cup_{i=1}^n \{(F,x_i),(F,\overline{x}_i)\}.$$

Clearly, $AF_{F'}$ can be constructed in polynomial time from F'. Its structure is visualized in Figure 1.6 below. Note that the only conflict-free sets of arguments containing F are subsets of the form $\{F\} \cup S$, where $S \subset Y \cup \overline{Y}$. Also, the only conflict-free sets of arguments meeting C are of the form $C' \cup T$, where $\emptyset \neq C' \subseteq C$ and $T \subset X \cup \overline{X}$.

We prove that $F' = \forall X \exists Y F(X,Y)$ is an accepted instance of $\forall\exists$SAT if and only if F is deliberatively accepted in $AF_{F'}$.

Suppose that F is deliberatively accepted in $AF_{F'}$. Let α_X be any truth assignment for the variables in X. If all clauses from C are satisfied by α_X then (α_X, α_Y) is a satisfying assignment for $F(X,Y)$, for any α_Y. Suppose that there is a nonempty subset $C' \subseteq C$ such that no clause C_i of C' is satisfied by α_X. This means that for every literal $l \in (X \cup \overline{X}) \cap C_i$, we have $\alpha(l) = false$. It follows that $T = \{l \in X \cup \overline{X} | \alpha(l) = true\} \cup C'$ is a conflict-free set in $AF_{F'}$ attacking the argument F (by

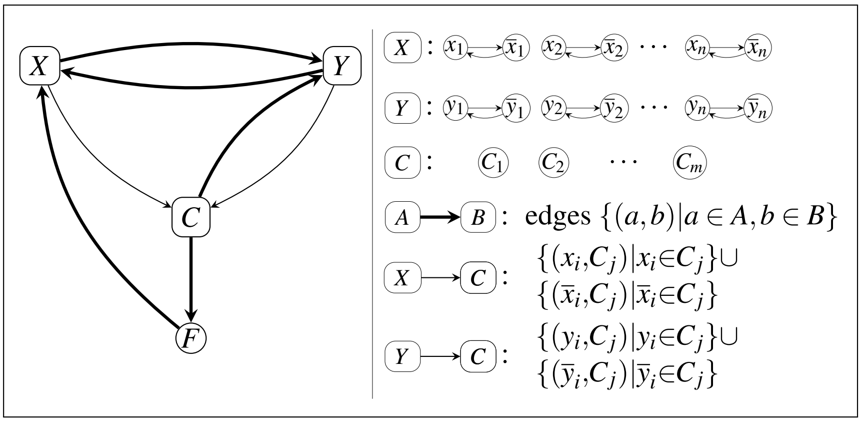


Figure 1.6.: The argumentation framework $AF_{F'}$ associated to instance $F' = \forall X \exists Y F(X,Y)$.

all arguments in C'). Since F is deliberatively accepted, there is a conflict-free set of argument S such that $F \in S$ and $T \cap S^- \subseteq S^+$. It follows that $T - \{F\} \subseteq Y \cup \overline{Y}$ and for each $C_i \in C'$ there is $l \in Y \cup \overline{Y}$ such that $(l, C_i) \in D$, that is $l \in C_i$. If we consider α_Y, the assignment with $\alpha_Y(l) = true$ for all these literals l, we obtain that (α_X, α_Y) is a satisfying assignment for $F(X,Y)$. Hence F' is a positive instance of $\forall\exists$SAT.

Conversely, let $F' = \forall X \exists Y F(X,Y)$ be a positive instance of $\forall\exists$SAT. Let T be a conflict-free set of arguments attacking F (that is, $C' = T \cap C \neq \emptyset$). Let α_X a truth assignment for variables in X such that for each $l \in (X \cup \overline{X}) \cap T$ we have $\alpha_X(l) = false$ (such assignment exists since T is conflict-free). Since F' is a positive instance of $\forall\exists$SAT, there is a truth assignment α_Y for the variables in Y such that for each $C_i \in C'$ there is a literal $l_i \in (Y \cup \overline{Y}) \cap C_i$ such that $\alpha_Y(l_i) = true$. Taking S the set of arguments containing F and all such literals l_i, we obtain a conflict-free set of arguments with the property that $T \cap S^- \subseteq S^+$. It follows that F is deliberatively accepted in $AF_{F'}$. $\square$

It is interesting to identify graph-theoretic constraints for an argumentation framework under which the problem **Deliberative acceptability** becomes polynomial-time solvable. A first such constraint follows from Proposition 15 and the polynomial runtime of Dunne's algorithm.

Proposition 17 *If the underlying undirected graph associated to the argumentation framework $AF = (A, D)$ is bipartite, then* **Deliberative acceptability** *is in* P.

A more interesting restriction of the **Deliberative acceptability** problem can be obtained by imposing that the underlying undirected graph associated to the argumentation framework to be co-chordal. A chordal (triangulated) graph is an undirected graph which does not contain as an induced subgraph the cycle graph C_k, for any $k \geq 4$ (equivalently, in such a graph every cycle of length at least 4 has a chord).

The complement of a chordal graph is a co-chordal graph. Chordal graphs can be recognized in $O(n+m)$ time, where n is its number of vertices and m is its number of edges, Rose [Ros70], Tarjan and Yannakakis, [TY84]. Also, the number of cliques (maximal complete subgraphs) of a chordal graph is $O(n)$ and all these cliques can be found in $O(n+m)$, using *Maximum Cardinality Search*, [TY84]. It follows that in a co-chordal graph the number of maximal independent sets of vertices (maximal stable sets) is $O(n)$ and all these maximal stable sets can be found in linear time.) Hence, for the argumentation frameworks with an underlying co-chordal graph, the number of maximal conflict-free sets is linear.

Proposition 18 *If the underlying undirected graph associated to the argumentation framework $AF = (A, D)$ is a co-chordal graph, then* **Deliberative acceptability** *is in* P.

Proof. Firstly, let us remark that if the condition in the definition of deliberative acceptance of an argument a in an argumentation framework $AF = (A, D)$ is fulfilled only by maximal conflict-free sets of arguments, then it is fulfilled by arbitrary conflict-free sets. Indeed, if T is a conflict-free set attacking a, let T_0 be a maximal conflict-free set containing T. Clearly, $T \cap a^- \subseteq T_0 \cap a^-$, hence T_0 attacks a. Since T_0 is maximal, there is the conflict-free S such that $a \in S$ and $T_0 \cap S^- \subseteq S^+$. Let S_0 a maximal conflict-free set containing S. Clearly, $a \in S_0$ and $T_0 \cap S_0^- \subseteq S_0^+$. Hence $T \cap S^- \subseteq S_0^+$.

Since the underlying undirected graph G of AF is co-chordal, there are only $O(|A|)$ maximal conflict-free sets Tattacking a. The subgraph of G induced by $A - a^-$ is co-chordal, hence there are $O(|A|)$ maximal conflict-free sets containing a. Each such set S can be tested in polynomial time if $T \cap S^- \subseteq S^+$. It follows that in polynomial time we can decide if a is deliberatively accepted. □

2. A Normal Form for Argumentation Frameworks

We show that any argumentation framework can be syntactically augmented into a normal form (having a simplified attack relation), preserving the semantic properties of original arguments.
An argumentation framework is in *normal form* if no argument attacks a conflicting pair of arguments. An *augmentation* of an argumentation framework is obtained by adding new arguments and changing the attack relation such that the acceptability status of original arguments is maintained in the new framework. Furthermore, we define join-normal semantics leading to augmentations of the joined argumentation frameworks. Also, a rewriting technique which transforms in cubic time a given argumentation framework into a normal form is devised.

2.1. Introduction

It is well-known that the syntactical structure of argumentation frameworks directly influences the output (e.g., Dung [Dun95], Dunne and Bench-Capon [DBC02], or Baroni and Giacomin [BG03]) and the complexity of algorithms for deciding acceptability questions, Dunne [Dun07]. In Prakken and Sartor [PS96], a four-layers succession for any AI-argumentation process was proposed. First we have the *logical layer* in which arguments are defined. Second, in the *dialectical layer*, the attacks are defined. Next, in the *procedural layer*, are defined rules that control the way arguments are introduced and challenged. Last layer, the *heuristics layer*, contains the remaining parts of the process, including methods for deciding the justification status of arguments.

In this chapter, keeping the abstract character of arguments and attacks, we are interested in understanding the syntactical properties of argumentation frameworks related to the *procedural layer*. We prove in a formal way that a discipline policy can be adopted in forming of an argumentation framework, without changing the semantic properties. It follows from our result that if the output of a dispute is obtained using an extension based reasoning engine, then it will be not influenced if we impose the following rule: any new argument added by an agent attacks no existing pair of conflicting arguments and, at the same time, at most one argument from any existing pair of conflicting arguments can attack the new argument.

We formalize this by considering σ-extensions (for σ a classical semantics), and introducing the notion of σ-*augmentation of an argumentation framework* AF. An argumentation framework AF' is a σ-*augmentation* of AF if it contains all arguments of AF, and the attacks of AF' are such that, for any set S of arguments of AF, S is contained in a σ-extension of AF if and only if S is contained in a σ-extension of AF'. We show that for suitable *join-normal* semantics the join of two argumentation frameworks gives rise to a common σ-augmentation of the joined argumentation frameworks. In the main result of this chapter, we prove that *for any argumentation framework* AF *there is a* σ-*augmentation* AF' in *normal form*, where σ is any Dung's classical semantics. An argumentation framework is in normal form if the set of arguments attacked by any argument contains no two attacking arguments. We prove that an argumentation framework is in normal form if and only if it can be constructed by adding its arguments one after one (the order does not matter), such that each new argument cannot attack two attacking arguments already added, and cannot be attacked by a pair of two attacking arguments already added.

2.2. The σ-Augmentations

We introduce the following binary relation between argumentation frameworks.

Definition 19 Let $AF, AF' \in \mathbb{AF}$ and σ be a semantics.
We say that AF' is a σ-*augmentation* of AF, denoted $AF \sqsubseteq_\sigma AF'$, if

- $Arg(AF) \subseteq Arg(AF')$,
- for any $S \in \sigma(AF)$ there is $S' \in \sigma(AF')$ such that $S \subseteq S'$, and
- for any $S' \in \sigma(AF')$ there is $S \in \sigma(AF)$ such that $S' \cap Arg(AF) \subseteq S$.

The binary relation $\sqsubseteq_\sigma$ between argumentation frameworks is a preorder : clearly $\sqsubseteq_\sigma$ is reflexive, and it is transitive as the following proposition shows.

Proposition 20 *If* $AF \sqsubseteq_\sigma AF'$ *and* $AF' \sqsubseteq_\sigma AF''$, *then* $AF \sqsubseteq_\sigma AF''$.

Proof. Clearly, $Arg(AF) \subseteq Arg(AF'')$.
Let $S \in \sigma(AF)$. Since $AF \sqsubseteq_\sigma AF'$, there is $S' \in \sigma(AF')$ such that $S \subseteq S'$, and since $AF' \sqsubseteq_\sigma AF''$, there is $S'' \in \sigma(AF'')$ such that $S' \subseteq S''$. Hence for any $S \in \sigma(AF)$ there exists $S'' \in \sigma(AF'')$ such that $S \subseteq S''$.

Let $S'' \in \sigma(AF'')$. Since $AF' \sqsubseteq_\sigma AF''$, there is $S' \in \sigma(AF')$ such that $S'' \cap Arg(AF') \subseteq S'$. Since $AF \sqsubseteq_\sigma AF'$, there is $S \in \sigma(AF)$ such that $S' \cap Arg(AF) \subseteq S$. Since $Arg(AF) \subseteq Arg(AF')$ it follows that $S'' \cap Arg(AF) \subseteq S'' \cap Arg(AF') \subseteq S'$, hence $S'' \cap Arg(AF) \subseteq S' \cap Arg(AF) \subseteq S$. □

It follows that if we define $AF \equiv_\sigma AF'$ if and only if $AF \sqsubseteq_\sigma AF'$ and $AF' \sqsubseteq_\sigma AF$ then, we obtain an equivalence relation on $\mathbb{AF}$. Two $\equiv_\sigma$-equivalent argumentation frameworks have the same set of arguments, but they are not isomorphic in general. For example, $AF = (\{a,b\},\{(a,a),(a,b)\})$ and $AF' = (\{a,b\},\{(a,a),(b,b)\})$ are $\equiv_{\mathbf{adm}}$ (since $\mathbf{adm}(AF) = \{\emptyset\} = \mathbf{adm}(AF')$), but, clearly, they are not isomorphic.

It is not necessary that the attack set of the σ-augmentation to be a superset of the attack set of the initial argumentation framework, as the following example shows.

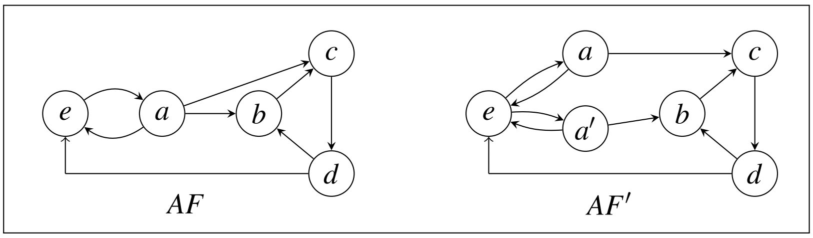


Figure 2.1.: AF' is an admissible augmentation of AF

Example 21 *Let us consider the two argumentation frameworks in the Figure 2.1. We have $A' = A \cup \{a'\}$ and $D' = (D - \{(a,b)\}) \cup \{(e,a'),(a',e),(a',b)\}$, hence $D \not\subseteq D'$. However, $AF \sqsubseteq_{adm} AF'$. Indeed, the admissible sets in AF are $\emptyset$, $\{a\}$, and $\{a,d\}$ (no conflict-free set containing b defends the attack (d,b), no conflict-free set containing c defends the attack (b,c)), which remain admissible sets in AF'. The admissible sets in AF' are $\emptyset$, $\{a\}$, $\{a'\}$, $\{a,a'\}$, $\{a,d\}$, $\{a',d\}$, and $\{a,a',d\}$ (the "new" conflict-free sets $\{a,b\}$ and $\{a',c\}$ can not be extended to admissible sets in AF' due to the attacks (a',b), respectively (a,c)), and their intersections with A are contained in admissible sets of AF.*

The next proposition follows easily from the definition.

Proposition 22

(i) If $\sigma(AF) = \emptyset$, then we have $AF \sqsubseteq_\sigma AF'$ if and only if $Arg(AF) \subseteq Arg(AF')$ and $\sigma(AF') = \emptyset$.

(ii) If $\sigma(AF) = \{\emptyset\}$, then we have $AF \sqsubseteq_\sigma AF'$ if and only if $Arg(AF) \subseteq Arg(AF')$, $\sigma(AF') \neq \emptyset$, and $S' \cap Arg(AF) = \emptyset$ for all $S' \in \sigma(AF')$.

It is easy to prove that the σ-credulous acceptability of an argument in a given AF is not changed in a σ-augmentation AF' of AF. More precisely, the following proposition holds.

Proposition 23 *If $AF \sqsubseteq_\sigma AF'$ and $a \in Arg(AF)$ then a is σ-credulously accepted in AF if and only if a is σ-credulously accepted in AF'.*

Proof. If there is $S \in \sigma(AF)$ such that $a \in S$, then since $AF \sqsubseteq_\sigma AF'$ it follows that there is $S' \in \sigma(AF')$ such that $S \subseteq S'$, hence there is $S' \in \sigma(AF')$ such that $a \in S'$, that is a is σ-credulously accepted in AF'. Conversely, if there is $S' \in \sigma(AF')$ such that $a \in S'$, then since $AF \sqsubseteq_\sigma AF'$ and $a \in Arg(AF)$, it follows that there is $S \in \sigma(AF)$ such that $S' \cap Arg(AF) \subseteq S$, hence there is $S \in \sigma(AF)$ such that $a \in S$, that is a is σ-credulously accepted in AF. □

The converse of Proposition 23 does not hold: if $AF = (\{a,b\},\{(a,b),(b,a)\})$, $AF' = (\{a,b,a',b'\},\{(a,a'),(a',b),(b,b'),(b',a)\})$, and $\sigma = \mathbf{adm}$, then a and b are **adm**-credulously accepted in AF and AF'. However, $AF \not\sqsubseteq_{\mathbf{adm}} AF'$, since the admissible set $\{a,b\}$ in AF' is not contained in an admissible set in AF.

If σ is an admissibility-based semantics, the σ-sceptical acceptance is not preserved in general by the σ-augmentations. Indeed, let the argument a be **adm**-sceptically accepted in the argumentation framework AF and let AF' be the argumentation framework obtained from AF by adding a new copy a' of a, each attack (a,x) or (x,a) giving rise to a new attack (a',x) or (x,a'), and adding the attacks (a,a') and (a',a). It is not difficult to see that $\mathbf{adm}(AF') = \mathbf{adm}(AF) \cup \{S - \{a\} \cup \{a'\} | S \in \mathbf{adm}(AF), a \in S\}$. It follows that $AF \sqsubseteq_{\mathbf{adm}} AF'$ but a is not **adm**-sceptically accepted in the argumentation framework AF'.

A simple way of constructing σ-augmentations is given by the join of two argumentation frameworks.

Definition 24 Let AF_1 and AF_2 be disjoint argumentation frameworks, that is $Arg(AF_1) \cap Arg(AF_2) = \emptyset$.

- The *disjoint union* of AF_1 and AF_2 is the argumentation framework $AF' = AF_1 \dot\cup AF_2$, where $Arg(AF') = Arg(AF_1) \cup Arg(AF_2)$ and $Def(AF') = Def(AF_1) \cup Def(AF_2)$.
- The *sum* of AF_1 and AF_2 is the argumentation framework $AF'' = AF_1 + AF_2$, where $Arg(AF'') = Arg(AF_1) \cup Arg(AF_2)$ and $Def(AF'') = Def(AF_1) \cup Def(AF_2) \cup \{(a_1,a_2), (a_2,a_1) | a_i \in Arg(AF_i), i = 1,2\}$.
- If σ is a semantics then it is *join-normal* if $\sigma(AF_1 \dot\cup AF_2) = \{S \cup S' | S \in \sigma(AF_1), S' \in \sigma(AF_2)\}$ and $\sigma(AF_1 + AF_2) = \sigma(AF_1) \cup \sigma(AF_2)$.

If $\sigma \in \{\mathbf{adm, comp, pref, gr, stb}\}$ then σ is join-normal.
Indeed, S is a conflict-free set in $AF_1 \dot\cup AF_2$ if and only if $S_i = S \cap Arg(AF_i)$ is a conflict-free set in AF_i ($i \in \{1,2\}$). Also, $S^+ = S_1^+ \dot\cup S_2^+$.
Similarly, S is a conflict-free set in $AF_1 + AF_2$ if and only if $S \in \mathbf{cf}(AF_1)$ or $S \in \mathbf{cf}(AF_2)$. If $S \in \mathbf{cf}(AF_1)$ then $S^+ = Arg(AF_2) \cup S^+ \cap Arg(AF_1)$ and if $S \in \mathbf{cf}(AF_2)$ then $S^+ = Arg(AF_1) \cup S^+ \cap Arg(AF_2)$.
The next proposition follows easily from the definition above.

Proposition 25 *Let AF_1 and AF_2 be disjoint argumentation frameworks, and σ a join-normal semantics. Then $AF_1, AF_2 \sqsubseteq_\sigma AF_1 \dot\cup AF_2$, and $AF_1, AF_2 \sqsubseteq_\sigma AF_1 + AF_2$.*

We close this section by noting that σ-augmentations can be defined equivalently, for $\sigma \in \{$**adm, comp, pref, gr, stb**$\}$, using Caminada's labellings. More precisely, the following proposition is easy to prove from Caminada's characterizations ([Cam06a]) of σ-extensions, where the extension of a labelling from a subset to a larger set is the usual function extension.

Proposition 26 *Let $\sigma \in \{$**adm, comp, pref, gr, stb**$\}$. AF' is a σ-augmentation of the argumentation framework AF if and only if i) $Arg(AF) \subseteq Arg(AF')$, ii) any σ-labelling of AF can be extended to a σ-labelling of AF', and iii) the restriction of any σ-labelling of AF' to $Arg(AF)$ can be extended to a σ-labelling of AF.*

2.3. Normal Forms

In this section we confine ourselves only to $\sigma =$ **adm** and we refer to an **adm**-augmentation as an admissible augmentation.

The results obtained for admissible augmentations can be easily adapted for σ-augmentations, where $\sigma \in \{$**comp, pref, gr, stb**$\}$.

An admissible augmentation can be viewed as adding "auxiliary" arguments in order to simplify the combinatorial structure of the given argumentation framework and, at the same time, maintaining all the credulous acceptability conclusions (see Proposition 23). We consider this simplified structure a *normal form* as follows.

Definition 27 An argumentation framework $AF = (A, D)$ is in *normal form* if for each $a \in A$ there are no $b, c \in a^+$ such that $b \neq c$ and $(b, c) \in D$. A set $S \subseteq A$ with the property, that $(a, b) \notin D$ for $a, b \in S$ and $a \neq b$, is referred as *d-conflict-free*.

Some properties of an argumentation framework in normal form are given in the next proposition. Note that the part ii) of this proposition shows that an argumentation framework is in normal form if and only if it can be constructed by adding its arguments one after one (the order does not matter), such that each new argument cannot attack two attacking arguments already added, and cannot be attacked by a pair of two attacking arguments already added.

Proposition 28

(i) Let $AF = (A, D)$ be an argumentation framework in normal form. Then for each $a \in A$ the set a^- is d-conflict-free. Moreover, in any set of four arguments of AF there are two non-attacking arguments.

(ii) An argumentation framework $AF = (A, D)$ is in normal form if and only if for any ordering $A = \{a_1, a_2, \ldots, a_n\}$, the sets $\vec{a}_i^{\,-} = a_i^- \cap \{a_1, \ldots, a_{i-1}\}$ and $\vec{a}_i^{\,+} = a_i^+ \cap \{a_1, \ldots, a_{i-1}\}$ are d-conflict-free, for all $i \in \{2, \ldots, n\}$.

Proof (i) Suppose that there is $a_0 \in A$ such that a_0^- is not a d-conflict-free set, that is, there are $b, c \in a_0^-$ such that $b \neq c$ and $(b,c) \in D$. But then, $a_0, c \in b^+$ and $(c, a_0) \in D$, that is the set b^+ is not d-conflict free, a contradiction.
If there are four pairwise attacking arguments $\{a,b,c,d\} \subseteq A$, then the underlying undirected graph of AF contains a complete graph K_4 as an induced subgraph, with nodes a,b,c,d and the edge $\{a,b\}$ generated by the attack $(a,b) \in D$ (see Figure 2.2 below). Since a^+ in AF is d-conflict-free, we are forced to have $(c,a) \in D$ and $(d,a) \in D$; but then, a^- contains c and d, and since $(c,d) \in D$ or $(d,c) \in D$, a^- is not d-conflict-free, a contradiction.

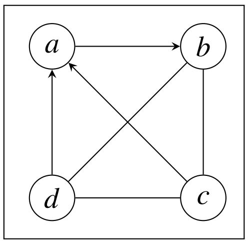


Figure 2.2.: An induced K_4 in AF.

(ii) Clearly, if AF is in normal form, then for any ordering $A = \{a_1, a_2, \dots, a_n\}$, and any $i \in \{2, \dots, n\}$, a_i^+ and a_i^- are d-conflict-free sets, therefore their subsets $\vec{a}_i^+$ and $\vec{a}_i^-$ are d-conflict-free. Conversely, let $AF = (A,D)$ satisfying the property stated. If AF is not in normal form, there are $a,b,c \in A$ such that $(a,b),(a,c),(b,c) \in A$. Any ordering of A with $a_1 := a, a_2 := b, a_3 := c$ has $\vec{a}_3^- = \{a,b\}$ which is not d-conflict-free, a contradiction. □

The next algorithm eliminates an attack between arguments attacked by the same argument in a given argumentation framework.

Input $AF = (A,D)$ *an argumentation framework,* $a,b,c \in A$ *with* $(a,b),(a,c),(b,c) \in D$;
Output $AF' = (A',D')$;
add to A two new arguments a_1, a_2 giving A';
put in D' all attacks in D;
delete from D' the attack (a,b);
add to D' the attacks $(a,a_1),(a_1,a_2),(a_2,b)$;
Return AF'

Algorithm 1: ELIM1$(AF; a,b,c)$

The effect of ELIM1$(AF; a,b,c)$ is depicted in the Figure 2.3. The squiggly arrows signify sets of all attacks, between arguments a,b,c and the sets of arguments in the rectangular boxes.

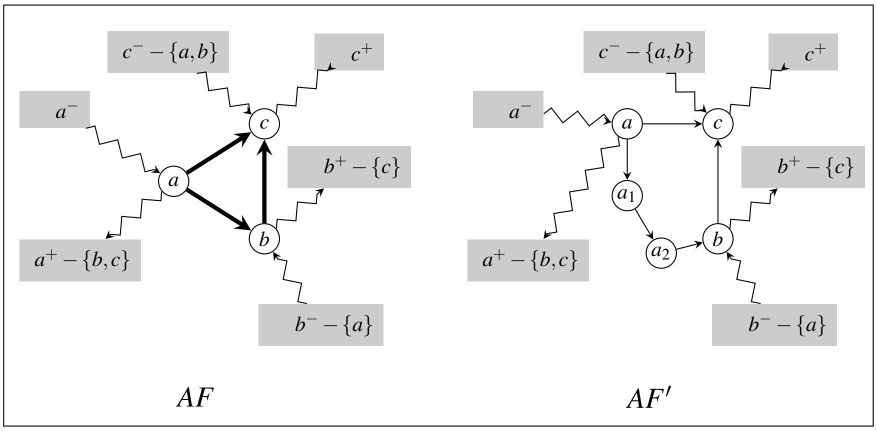


Figure 2.3.: Elimination of a bad triangle.

Proposition 29 *The argumentation framework $AF' = (A', D')$, which is returned by ELIM1$(AF; a, b, c)$, is an admissible augmentation of AF.*

Proof. Let $S \subseteq A$ be an admissible set in AF. We prove that $S' \subseteq A'$ is an admissible set in AF', where:

$$S' = \begin{cases} S \cup \{a_2\} & \text{if } a \in S, \\ S \cup \{a_1\} & \text{if } a \notin S, b \in S, \\ S & \text{if } a \notin S, b \notin S. \end{cases}$$

If $S \subseteq A$ is an admissible set containing a in AF, then $S' = S \cup \{a_2\}$ is a conflict-free set in AF'. Indeed, no attack between the arguments in A is added by the algorithm ELIM1, hence S is conflict free in AF'. The only attacks containing a_2 are (a_1, a_2) and (a_2, b). But $a_1 \notin S$ (because $a_1 \notin A$), and $b \notin S$ (because $a \in S$, $(a, b) \in D$, and S is conflict-free set in AF). It follows that $S \cup \{a_2\}$ is a conflict-free set in AF'. The attack (a_1, a_2) against $S \cup \{a_2\}$ is defeated by (a, a_1), since $a \in S$. Any attack (x, y) with $x \in A - S$ and $y \in S$ is defeated by an attack (z, x) with $z \in S$, since S is admissible set in AF. It follows that $S \cup \{a_2\}$ is a conflict-free set in AF' which defends itself against any attack in AF', that is, $S \cup \{a_2\}$ is an admissible set in AF'.

If S is an admissible set in AF such that $a \notin S$ but $b \in S$, then adding a_1 to S we obtain a conflict-free set in AF' (since $a \notin S$ and $a_2 \notin S$, the only attacks involving a_1 – (a, a_1) and (a_1, a_2) – are not between arguments from $S \cup \{a_1\}$). The attack (a_2, b) on $S \cup \{a_1\}$ is defeated by (a_1, a_2). The attack (a, a_1) must be defeated by some argument $x \in a^- \cap S$, because in AF the attack (a, b) must be defeated. Any attack (x, y) with $x \in A - S$ and $y \in S$ is defeated by an attack (z, x) with $z \in S$, since S is admissible set in AF. It follows that $S \cup \{a_1\}$ is an admissible set in AF'.

If S is an admissible set in AF not containing a and b, then S remains conflict-free since no attacks between arguments in A are added. Also all attacks from an argument in S remain in AF', and no new attack against S is introduced. It follows that S continues to defend itself against any attack in AF', hence S is an admissible set in AF'.

On the other hand, let $S' \subseteq A'$ be an admissible set in AF'. We prove that $S = S' \cap A$ is an admissible set in AF.
If S' is an admissible set containing a_2 in AF', then $a_1, b \notin S'$ (since S' is conflict-free and $(a_1,a_2),(a_2,b) \in D'$). Since $(a_1,a_2) \in D'$ and S' is admissible, it follows that a_1 must be attacked by S' in AF'. The only attack on a_1 in AF' is (a,a_1). Hence $a \in S'$. $S' - \{a_2\}$ is conflict free in AF, because $b \notin S'$. Any attack (x,y) with $x \in A - S$ and $y \in S$ is defeated by an attack (z,x) with $z \in S$, since S' is admissible set in AF'. It follows that $S' - \{a_2\} = S' \cap A$ is an admissible set in AF.

If S' is an admissible set containing a_1 in AF', a similar proof shows that $S' - \{a_1\} = S' \cap A$ is an admissible set in AF.

If S' is an admissible set in AF' such that $a_1, a_2 \notin S'$, we can suppose that $b \notin S'$. Otherwise, if $b \in S'$ then the attack (a_2,b) can not be defeated by S', since the only attack on a_2 in AF' is (a_1,a_2). Since the only additional attack involving at least one argument in S' can be (a,b), it follows that S' is a conflict-free set in AF and also defends itself against any attack in AF (because it was an admissible set in AF'). □

Proposition 30 *The argumentation framework $AF' = (A',D')$ returned by calling ELIM1($AF;a,b,c$) satisfies $AF \sqsubseteq_\sigma AF'$ for $\sigma \in \{$**comp, pref, gr, stb**$\}$.*

Proof. For $\sigma \in \{\textbf{comp}, \textbf{pref}\}$ the proof follows from Proposition 29. Indeed, if $S \in \sigma(AF)$ then S is an admissible set in AF and, by Proposition 29, can be extended to an admissible set in AF'. Since any admissible set can be extended to a complete or preferred extension, it follows that there is $S' \in \sigma(AF')$ such that $S \subseteq S'$. Conversely, if $S' \in \sigma(AF')$ then S' is an admissible set in AF' and, by Proposition 29, $S' \cap A$ can be extended to an admissible set in AF. Since any admissible set can be extended to a complete or prefered extension, it follows that there is $S \in \sigma(AF)$ such that $S' \cap A \subseteq S$.

For $\sigma = \textbf{stb}$ it is not difficult to verify that if $S \in \textbf{stb}(AF)$ then $S' \in \textbf{stb}(AF')$, where

$$S' = \begin{cases} S \cup \{a_2\} & \text{if } a \in S, \\ S \cup \{a_1\} & \text{if } a \notin S \end{cases}$$

and, if $S' \in \textbf{stb}(AF')$ then $S \in \textbf{stb}(AF)$, where

$$S = \begin{cases} S' - \{a_2\} & \text{if } a \in S', \\ S' - \{a_1\} & \text{if } a \notin S'. \end{cases}$$

For $\sigma = \textbf{gr}$, we use Proposition 26 and Observation 7. Clearly, if each $x \in Arg(AF)$ satisfies $x^- \neq \emptyset$, then the same property holds in AF' and $\textbf{gr}(AF) = \textbf{gr}(AF') =$

$\{\emptyset\}$. Suppose that $\mathbf{gr}(AF) = \{S\}$, $S \neq \emptyset$ and let Lab a **gr**-labelling of AF such that $S = Lab^{-1}(I)$. If $a \in S$, then we extend Lab to AF' by taking $Lab(a_1) = O$, $Lab(a_2) = I$, and the linear ordering of $Lab^{-1}(I)$ in AF' is obtained by considering a_2 the successor of a. It is not difficult to see that we obtain a **gr**-labelling of AF'. If $a \notin S$, and $Lab(a) = O$ then a **gr**-labelling of AF' is obtained by taking $Lab(a_1) = I$ and the linear ordering of $Lab^{-1}(I)$ in AF' is obtained by considering a_1 the successor of an attacker of a labeled I. If $a \notin S$, and $Lab(a) = U$ then Lab remains a **gr**-labelling of AF'. A similar analysis can be used to show that the restriction to AF of a **gr**-labelling of AF' gives rise to a **gr**-labelling of AF. □

By iterating the algorithm ELIM1, we obtain:

```
AF' := AF;
foreach a,b,c ∈ Arg(AF) s.t. (a,b),(a,c),(b,c) ∈ Def(AF) do
  | AF' := ELIM1(AF'; a,b,c)
end
Return AF'
```

Algorithm 2: ELIMALL(AF)

Proposition 31 *For any argumentation framework $AF = (A, D)$ there is an admissible augmentation $AF' = (A', D')$ in normal form. Furthermore, AF' can be constructed from AF in $O(|A|^3)$ time.*

Proof. Using Propositions 20 and 29, the above iteration of the algorithm ELIM1 returns an admissible augmentation AF' of the given AF. The **for** condition assures that AF', the returned argumentation framework, is in normal form. It remains to prove that the algorithm finishes.
We call a triangle $\{a, b, c\} \subseteq A'$ with $(a, b), (a, c), (b, c) \in D'$, a *bad triangle*. Clearly, the algorithm finishes when there is no bad triangle in the current argumentation framework.
In each **for**-iteration the total number of bad triangles of the current argumentation framework AF' decreases by 1. Indeed, the algorithm ELIM1$(AF'; a, b, c)$ destroys a bad triangle and creates no new bad triangle, since the two new arguments a_1 and a_2 are not contained in a triangle in the new argumentation framework. Since the number of bad triangles in AF it at most $\binom{|A|}{3}$, and the running time of ELIM1$(AF'; a, b, c)$ is $O(1)$, the final argumentation framework AF' is obtained in $O(|A|^3)$ time. □

Summarizing the results obtained in this section, using Propositions 28ii), 30 and 31, we have the following theorem.

Theorem 32 *Any argumentation framework $AF = (A, D)$ has an admissible augmentation $AF' = (A', D')$ which can be formed by adding the arguments one after* *one such that each argument attacks a d-conflict-free set of its predecessors and is attacked by a d-conflict-free set of its predecessors. Furthermore AF' is also a σ-augmentation of AF for any Dung's classical semantics σ.*

The Figure 2.4 below suggests the way in which the argumentation framework AF' from the above theorem is formed. Any new argument a_{new} added by an agent in a round cannot attack an existing pair of conflicting arguments, that is a_{new} attacks only a coherent set of existing arguments. The agent knows that, if she wants, in a later round can use a surrogate of a_{new} to attack other arguments which in the actual round are in conflict with those selected to be attacked. In the same time, from the set of existing arguments only a coherent set can attack the new argument. The other attacks will be simulated in future rounds by using again special surrogate arguments. In this way, a more logical scene of dispute can be devised, which is however (polynomially) longer as one in which our discipline policy is not followed.

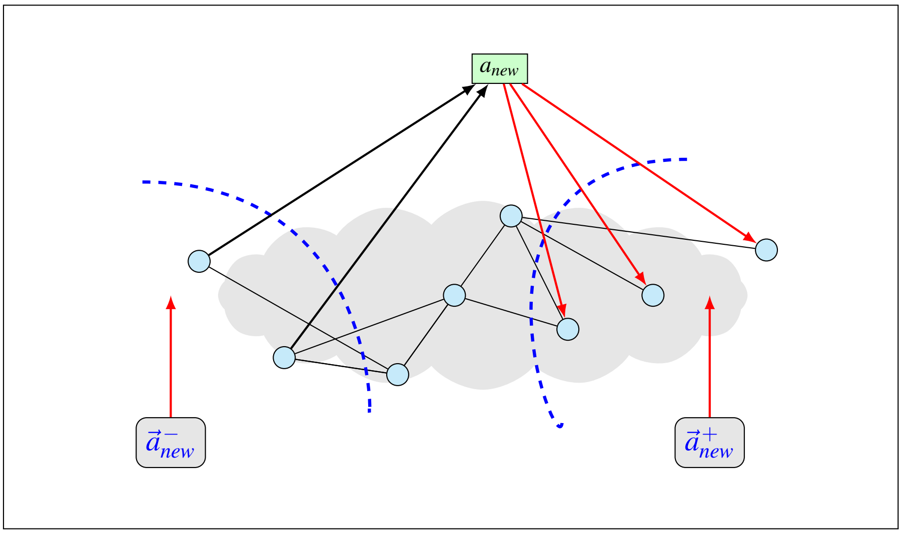


Figure 2.4.: Discipline policy in forming an AF.

3. Polyhedral Labellings for Argumentation Frameworks

Polyhedral labellings associated to an argumentation framework are introduced. A *polyhedral labelling* for an argumentation framework $AF = (A, D)$ is a polytope P_{AF}, that is, a bounded set of solutions $x \in \mathbb{R}^A$ (x_a is the label of the argument $a \in A$), to a system of linear constraints, such that the set of integral vectors in P_{AF} are exactly the incidence vectors of some specific type of Dung's extensions. The linear constraints vary from the obvious $x_a = 1$ for each non attacked argument a, or $x_a + x_b \leq 1$ for each attack $(a, b) \in D$ (in order to assure Dung's conflict-free condition), to more deep inequalities of the form "*the sum of the label of an argument and the labels of all its attackers is at least 1*" or if (b, a) is an attack then "*the label of a is not greater than the sum of the labels of all attackers of b*".

3.1. Introduction

An important intuitive way to express Dung's extension-based semantics is using *argument labellings*, as proposed by Caminada [Cam06a] (originally introduced in Pollock [Pol95]). The idea is to consider symbolic vectors[1] $\lambda \in \{I, O, U\}^A$, such that if $a \in A$ is an argument, then λ_a is the label of a, with the intuitive meaning: $\lambda_a = I$ (i.e. In) if and only if a is accepted, $\lambda_a = O$ (i.e. Out) if and only if a is rejected, and $\lambda_a = U$ (i.e. Undecided) if and only if one abstains from an opinion on whether the argument a is accepted or rejected.

Constraining the labels of arguments with respect to the attack relation D of the argumentation framework, Caminada characterized the subsets of $\{I, O, U\}^A$ corresponding to Dung's semantics (see 0.2).

Another interesting approach to the semantics of argumentation frameworks was introduced by Gabbay in [Gab11, Gab12a, Gab12b] under the name *equational approach*, and independently by Gratie and Florea in [GF11] under the name *fuzzy labellings*. The idea is to consider solutions $x \in [0, 1]^A$ of some non-linear systems of equations associated to the argumentation framework and to relate them to Caminada labellings.

[1] Throughout this chapter, if A is a finite set, we make no distinction between the set B^A of all functions from A to B and the set $B^{|A|}$ of all vectors with components from B and indexed by the elements of A. Supposing a fixed ordering of A, there is an obvious one to one correspondence between them, and we use the notation B^A for both sets.

Using this approach, Gabbay proposes an interesting method to avoid the semantics problems caused by the odd circuits in argumentation frameworks.

In this chapter, I introduce *polyhedral labellings* associated to an argumentation framework. The name suggests the use of ideas from *Polyhedral Combinatorics*, an important topic in *Combinatorial Optimization*, mainly concerned with encoding combinatorial problems by means of systems of linear equations and inequalities. The interest in such representation is that it makes the corresponding combinatorial optimization problems accessible to linear programming techniques (see, for example, Schrijver [Sch03]).

More precisely, if $\chi^S \in \{0,1\}^A$ is the incidence vector of a set $S \subseteq A$ of arguments (that is, $\chi^S_a = 1$ if $a \in S$ and $\chi^S_a = 0$ if $a \notin S$), and $\mathscr{S}$ is a collection of sets of arguments, then the convex hull of their incidence vectors, $\text{conv}\{\chi^S | S \in \mathscr{S}\}$, is a polytope in $\mathbb{R}^A$, therefore there exist a matrix $C \in \mathbb{R}^{m \times |A|}$ and a vector $b \in \mathbb{R}^m$ such that

$$\text{conv}\{\chi^S | S \in \mathscr{S}\} = \{x \in \mathbb{R}^A | Cx \leq b\}.$$

If $w : A \to \mathbb{R}$ is a weight function on A, and we are interested in finding a member S^* of $\mathscr{S}$ of maximum weight (where the weight of S is $w(S) = \sum_{a \in S} w(a)$), since $\mathscr{S}$ is finite and the weight function can be viewed as a linear function on $\mathbb{R}^A$, we could maximize over the convex hull $\text{conv}\{\chi^S | S \in \mathscr{S}\}$, that is, by the above representation, finding

$$\max\{w^T x | x \in \mathbb{R}^A, Cx \leq b\}.$$

This is computationally worthwhile when $\mathscr{S}$ is too large to evaluate the weight of each member S in $\mathscr{S}$, but the description $Cx \leq b$ of the above polytope has polynomial size. Then, we can solve in polynomial time the equivalent linear programming problem obtained (for example, using the ellipsoid method, Khachiyan [Kha79]). An illustration of this approach is discussed in Section 3.2, where we describe an interesting polytope encoding the *non-attacked sets* of arguments in an argumentation framework.

In several cases, the desired system of inequalities, $Cx \leq b$, turns out not to be a complete description, but just gives an approximation of the polytope $\text{conv}\{\chi^S | S \in \mathscr{S}\}$. This can still be useful, since in that case the linear programming problem gives a (hopefully good) upper bound for the combinatorial maximum. These bounds are used in designing branch-and-bound algorithms for the combinatorial problem, which are implemented in state-of-the-art integer programming solvers (e.g. CPLEX, Lowe [Low12]). In fact, exploiting integer programming encodings of the problems is an usual modeling method in areas related to argumentation, such as non monotonic reasoning (Bell et al. [BNS94]), satisfiability (Li, Zhou and Du [LZD04]), or answer set programming (Liu, Janhunen and Niemelä [LJN12]).

Summarizing, a *polyhedral labelling* for an argumentation frameork $AF = (A, D)$ is a polytope P_{AF}, that is, a bounded set of solutions $x \in \mathbb{R}^A$ (x_a is the label of the argument $a \in A$), of a system of linear inequalities (and equations), such that the set of integral vectors in P_{AF} are exactly the incidence vectors of various Dung's admissibility based extensions.

3.2. The non-attacked sets polytope

Definition 33 *Let $AF = (A,D)$ be an argumentation framework. A* ***non-attacked set*** *of arguments is a set $N \subseteq A$ such that $N^- \subseteq N$. Let $\mathcal{N}_{AF} := \{N | N \subseteq A, N^- \subseteq N\}$.*

Trivial non-attacked sets are $\emptyset, A \in \mathcal{N}_{AF}$ for any argumentation framework $AF = (A,D)$. The interest in such sets of arguments is given by the following proposition (see also the "directionality principle" in Baroni and Giacomin [BG07b]).

Proposition 34 *Let $a \in A$ be an argument in the argumentation framework $AF = (A,D)$ and $X \in \mathcal{N}_{AF}$ be a non-attacked set containing a. There is an admissible set S in AF such that $a \in S$ if and only if there is an admissible set S' in AF' such that $a \in S'$, where AF' is the argumentation framework induced by X in AF.*

Proof. Let S an admissible set in AF such that $a \in S$. Then S is conflict-free and $S^- \subseteq S^+$. Then, $S_X = S \cap X$ is a conflict-free set in AF', and $a \in S_X$. If S_X is not an admissible set in AF', then there is $b \in S_X^- \cap X - S_X^+$. Since S is an admissible set in AF, $b \in S^+$. It follows that there is $c \in S - X$ such that $(c,b) \in D$, that is $c \in X^- - X$, contradicting the hypothesis that X is a non-attacking set in AF.

Conversely, if $S' \subseteq X$ is an admissible set in AF' such that $a \in S'$, then it is a conflict-free set in AF. Since in AF we have $S'^- \subseteq X^- \subseteq X$, it follows that in AF we have $S'^- \subseteq S'^+$, that is, S' is an admissible set in AF. □

We show now that $\mathcal{N}_{AF}$ has an interesting polyhedral characterization. For each $X \in \mathcal{N}_{AF}$ we consider its incidence vector $\chi^X \in \{0,1\}^A$, with $\chi_a^X = 1$ if and only if $a \in X$.

Let $\mathbf{N}_{AF} = \{x \in \mathbb{R}^A | x \text{ satisfies } (*)\}$ be the polyhedron defined by

$$(*) \quad \begin{cases} 0 \leq x_a \leq 1 & \forall a \in A, \\ x_a - x_b \geq 0 & \forall (a,b) \in D. \end{cases}$$

Hence, if $x \in \mathbf{N}_{AF}$ then each argument $a \in A$ is labeled with the real number $x_a \in [0,1]$ such that $x_a \geq x_b$ whenever the argument a attacks the argument b. This type of constraints are used to model preferences (see, for example, *"value-based argumentation frameworks"*, Bench-Capon [BC03]).

Theorem 35 *Let $AF = (A,D)$ be an argumentation framework. Then,*

$$\mathbf{N}_{AF} = \text{conv}\{\chi^X | X \in \mathcal{N}_{AF}\}.$$

Proof. For $X \in \mathcal{N}_{AF}$, let $y = \chi^X$. Since $y_a \in \{0,1\}$, the first group of inequalities in $(*)$ is satisfied. Let $(a,b) \in D$. If $|\{a,b\} \cap X| \neq 1$, then $y_a = y_b$ and the second constraint in $(*)$ for (a,b), is satisfied with equality. If $a \in X$ and $b \notin X$ then $1 = y_a > y_b = 0$. Since $X \in \mathcal{N}_{AF}$, we can not have $a \notin X$ and $b \in X$. Hence $\chi^X \in \mathbf{N}_{AF}$ for each $X \in \mathcal{N}_{AF}$. It follows that

$$\mathrm{conv}\{\chi^X | X \in \mathcal{N}_{AF}\} \subseteq \mathbf{N}_{AF}.$$

To prove the converse inclusion, we observe that the integer vectors in $\mathbf{N}_{AF}$ are exactly the incidence vectors of non-attacked sets. Hence it is sufficient to prove that the vertices of $\mathbf{N}_{AF}$ are integral. Let x be a vertex of $\mathbf{N}_{AF}$.
Suppose that $Frac(x) = \{a \in A | 0 < x_a < 1\} \neq \emptyset$, and let $\alpha = \min\{x_a | a \in Frac(x)\}$. Take $\varepsilon > 0$ such that $\alpha - \varepsilon > 0$ and $\alpha + \varepsilon < x_a$ for all $a \in A$ such that $x_a > \alpha$. Then, let $x', x'' \in \mathbb{R}^A$ be such that:

$$x'_a = \begin{cases} \alpha - \varepsilon & \text{if } x_a = \alpha, \\ x_a & \text{if } x_a \neq \alpha \end{cases} \text{ and } x''_a = \begin{cases} \alpha + \varepsilon & \text{if } x_a = \alpha, \\ x_a & \text{if } x_a \neq \alpha. \end{cases}$$

By the choosing of ε, x' and x'' satisfy the first group of inequalities in $(*)$. Since the order of the components in x' and x'' is the same as in x, and since $x \in \mathbf{N}_{AF}$, it follows that the second group of inequalities are satisfied by x' and x''. Hence, $x', x'' \in \mathbf{N}_{AF}$, $x' \neq x''$, and $x = \frac{1}{2}x' + \frac{1}{2}x''$, contradicting the hypothesis that x is a vertex in $\mathbf{N}_{AF}$. $\square$

From Proposition 34, it follows that in order to decide the σ-acceptability of a given argument a in an argumentation framework , for $\sigma \in \{\mathbf{comp}, \mathbf{gr}, \mathbf{pref}\}$, is worthwhile to find a minimum cardinality non-attacked set containing a. This can be obtained with a simple polynomial algorithm (similar to one used for obtaining the grounded extension), but also using a linear programming solver, as a consequence of the Theorem 35.

Indeed, if we solve the linear program $\min\{c^T x | x \in \mathbf{N}_{AF}, x_{a_0} = 1\}$, for $c \in \mathbb{R}^A$ we obtain $\min\{\sum_{a \in X} c_a | X \in \mathcal{N}_{AF}, a_0 \in X\}$. In particular, for $c = \mathbf{1}$ (the vector with all components 1), the minimum value obtained is the minimum cardinality of a non-attacked set of arguments containing a_0. If we solve (in polynomial time) the above linear program and x^0 is the optimal solution, then the set $X = \{a | a \in A, x^0_a > 0\}$ is the minimum cardinality non-attacking set containing a_0.

3.3. Gabbay's Equational Approach

An interesting approach to the semantics of argumentation frameworks was introduced in Gabbay [Gab11, Gab12a, Gab12b] called the **equational approach**. Let $AF = (A, D)$ be an argumentation framework, and let A_0 be the set of arguments not attacked in AF: $A_0 = \{a \in A | a^- = \emptyset\}$. We consider for each argument $a \in A$ a real variable $x_a \in [0,1]$ and we are searching for real solutions of the following system of non-linear equations:

$$Eq_{\max}(AF) \quad \begin{cases} x_a = 1, \text{ if } a \in A_0 \\ x_a = 1 - \max\limits_{b \in a^-} x_b, \text{ otherwise.} \end{cases}$$

The following theorem holds.

Theorem 36 ([Gab12a]) *If $\lambda : A \to \{I, O, U\}$ is a Caminada complete labellings of AF, then taking $x_a = 0$ if $\lambda(a) = O$, $x_a = \frac{1}{2}$ if $\lambda(a) = U$, and $x_a = 1$ if $\lambda(a) = I$, we obtain a solution to the system of equations $Eq_{max}(AF)$. If x is a solution to the system $Eq_{max}(AF)$, then taking $\lambda(a) = O$ if $x_a = 0$, $\lambda(a) = U$ if $0 < x_a < 1$, and $\lambda(a) = I$ if $x_a = 1$ we obtain a Caminada complete labelling $\lambda : A \to \{I, O, U\}$.*

Example. For argumentation framework AF in Figure 3.1 below,

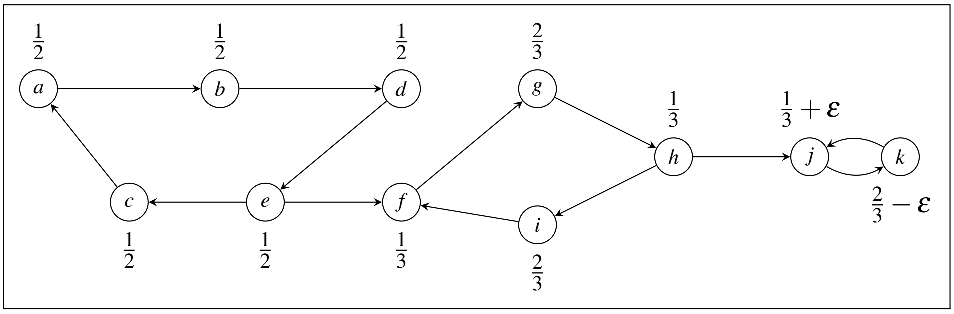


Figure 3.1.: $Eq_{\max}(AF)$ labellings. ε is a small positive number.

the $Eq_{\max}$ system is

$$Eq_{\max}(AF) \quad \begin{cases} x_a = 1 - x_c,\ x_b = 1 - x_a,\ x_c = 1 - x_e,\ x_d = 1 - x_b, \\ x_e = 1 - x_d,\ x_f = 1 - \max(x_e, x_i),\ x_k = 1 - x_j, \\ x_g = 1 - x_f,\ x_h = 1 - x_g,\ x_i = 1 - x_h,\ x_j = 1 - \max(x_h, x_k) \end{cases}$$

and a set of solutions to $Eq_{\max}(AF)$ are suggested.

We can observe that if we translate them to Caminada labellings as in Theorem 36, the result is not so good since, in this case $\mathbf{gr}(AF) = \{\emptyset\}$, and there are odd loops[2]. In order to avoid this, Gabbay [Gab12a] proposes the so called *perturbation method* to solve a system of the form $Eq(AF)$. Essentially, by this method, some variables x_a are forced to be 0 by extending the system $Eq(AF)$ with these new equations. The idea for choosing this forcing is to destroy the loops of AF. This gives rise to interesting semantics (which are, in general, not admissibility based), called **LB-semantics** in Gabbay [Gab12b].

If we make the convention that for $X = \emptyset$ then $\max_{x \in X} = 0$ and $\min_{x \in X} = 1$, then the following proposition holds.

[2] In Gabbay's terminology, a loop is a circuit (see the discussion after Theorem 46).

Proposition 37 *Let $AF = (A,D)$ be an argumentation framework. If x is a solution in $[0,1]$ of the system $Eq_{max}(AF)$ then $x_a = \min_{b\in a^-} \max_{c\in b^-} x_c$, for each $a \in A$.*

Proof. Let x be a solution in $[0,1]$ of the system $Eq_{\max}(AF)$. With our convention, it follows that the second group of equations in $Eq_{\max}(AF)$ is satisfied for each $a \in A$, that is $x_a = 1 - \max_{b\in a^-} x_b$. Then, we have successively,

$$x_a = 1 - \max_{b\in a^-} x_b = 1 - \max_{b\in a^-}(1 - \max_{c\in b^-} x_c) = -\max_{b\in a^-}(-\max_{c\in b^-} x_c) = \min_{b\in a^-} \max_{c\in b^-} x_c. \qquad \square$$

We note that the converse of the above proposition is not true. More precisely, if for a given argumentation framework $AF = (A,D)$, the labelling $x \in [0,1]^A$ satisfies $x_a = \min_{b\in a^-} \max_{c\in b^-} x_c$ for each $a \in A$, then x is not necessary a solution in $[0,1]$ of the system $Eq_{\max}(AF)$. For example, if $AF = (A,D)$ is a circuit, taking $x_a = 1$ for each $a \in A$, then x satisfies $x_a = \min_{b\in a^-} \max_{c\in b^-} x_c$, but it is not a solution of $Eq_{\max}(AF)$.

If, instead of searching the solutions of the system $Eq_{\max}(AF)$, we are searching for real solutions in $[0,1]$ of the following system of non-linear inequalities,

$$Ineq_{\max}(AF) \quad \begin{cases} x_a = 1, \text{ if } a \in A_0 \\ x_a \le 1 - \max_{b\in a^-} x_b, \text{ if } a \in A - A_0, \end{cases}$$

then, is not difficult to prove this set of solutions is convex. Moreover, we can easily translate it as the set of solutions to

$$P_{AF} \quad \begin{cases} 0 \le x_a & \le 1, \forall\, a \in A \\ x_a & = 1, \text{ if } a \in A_0 \\ x_a + x_b & \le 1, \forall (b,a) \in D, \end{cases}$$

that is, this set is a polytope. It is not difficult to show that the integral vectors in this polytope are exactly the incidence vectors of the conflict-free sets in AF containing the set A_0 of non-attacked arguments, therefore loosing the nice semantics property in Theorem 36.

In the next section, using linear constraints suggested by Proposition 37, we restrict the above polytope in order to obtain argumentation significance of the corresponding set of integral vectors.

3.4. Labellings Polytopes

Throughout this section we consider only argumentation frameworks $AF = (A,D)$ without isolated arguments, that is without arguments $a \in A$ such that $a^- \cup a^+ = \emptyset$. Clearly, adding or deleting an isolated argument does not influence the acceptability status of the other arguments.

Definition 38 *Let $AF = (A, D)$ be an argumentation framework. The **admissible sets polytope** of AF is the set $\mathrm{P}_{\mathrm{adm}}(AF)$ of all vectors in $\mathbb{R}^A$ satisfying:*

$$\begin{array}{lrcll} (1) & x_a & \geq & 0 & \forall a \in A, \\ (2) & x_a + x_b & \leq & 1 & \forall (b,a) \in D, \\ (3) & x_a - \sum\limits_{c \in b^-} x_c & \leq & 0 & \forall (b,a) \in D. \end{array}$$

We make the convention that if $b^- = \emptyset$, then the sum $\sum_{c \in b^-} x_c$ in (3) is 0.

Example 1. In Figure 3.2 we consider a simple argumentation framework $AF = (A, D)$ with $A = \{a, b\}$. We have two constraints of type (2), corresponding to the two attacks in $D = \{(a,a),(a,b)\}$. Note that in any argumentation framework if a is a self-attacking argument, then $x_a \in [0, 1/2]$ due to the type (2) constraint $x_a + x_a \leq 1$. The are two type (3) constraints, but the type (3) constraint for a gives $x_a \leq x_a$ which is trivially satisfied. It follows that all constraints giving $\mathrm{P}_{\mathrm{adm}}(AF)$ are those given in the middle of Figure 3.2 and a graphic illustration of the admissible sets polytope is at right. Its vertices are $(0,0), (\frac{1}{2}, 0)$, and $(\frac{1}{2}, \frac{1}{2})$. The only integer point in $\mathrm{P}_{\mathrm{adm}}(AF)$ is $(0,0)$.

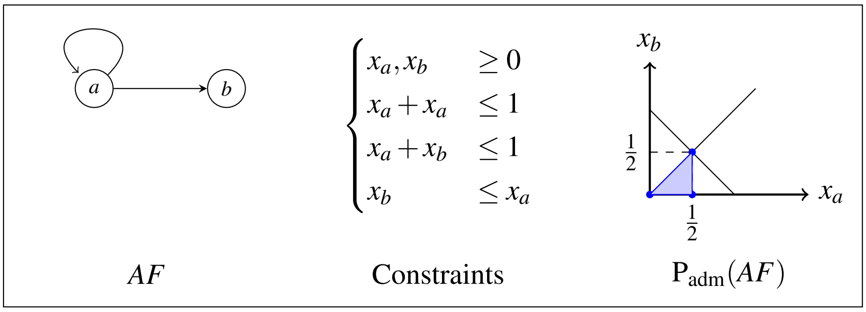


Figure 3.2.: An argumentation framework and its admissible sets polytope.

Example 2. Let $AF = (A, D)$ illustrated in Figure 3.3 with $A = \{a, b, c\}$ and $D = \{(a,b),(b,a),(b,c)\}$. The two mutual attacks (a,b) and (b,a) generate a single type (2) constraint $x_a + x_b \leq 1$. The only non-trivial type (3) constraint is $x_c \leq x_a$. It follows that all constraints giving $\mathrm{P}_{\mathrm{adm}}(AF)$ are those given in the middle of Figure 3.3 and its graphic representation appears at the right. Its vertices are the integral vectors $(0,0,0)^T, (0,1,0)^T, (1,0,0)^T$, and $(1,0,1)^T$ (where, x^T denotes the transpose of the row vector x).

In both above examples the integral vectors in $\mathrm{P}_{\mathrm{adm}}(AF)$ are exactly the incidence vectors of the admissible sets of the argumentation framework. We will prove next that this holds in general and this justifies the name of the considered polytope.

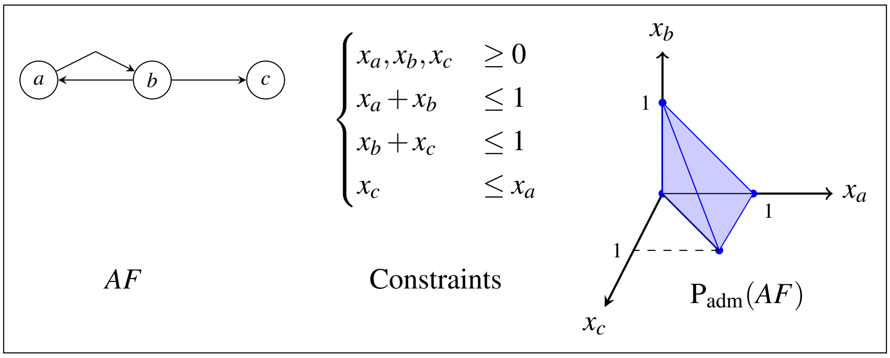


Figure 3.3.: An admissible sets polytope with integer vertices.

Lemma 39 *Let $AF = (A,D)$ be an argumentation framework without isolated arguments. If $x \in \mathbb{Z}^A$ satisfies the constraints* (1) *and* (2) *then x is a* 0-1 *vector.*

Proof. For each $a \in A$ there is $b \in A$ such that $(a,b) \in D$ or $(b,a) \in D$. By constraint (2), $x_a + x_b \leq 1$ and, by constraint (1), $x_b \geq 0$. It follows that $x_a \leq 1$, and, since $x_a \in \mathbb{Z}$, we have $x_a \in \{0,1\}$. □

Lemma 40 *Let $AF = (A,D)$ be an argumentation framework without isolated arguments. An integral vector $x \in \mathbb{Z}^A$ satisfies the constraints* (1) *and* (2) *if and only if x is the incidence vector of a conflict-free set in AF.*

Proof. Let $x \in \mathbb{Z}^A$ be an integral vector satisfying the constraints (1) and (2). By Lemma 39, x is a 0-1 vector, and there is $X \subseteq A$ such that $\chi^X = x$. There is no $(a,b) \in D$ with $a,b \in X$, since then $x_a = x_b = 1$ and the constraint (2) for (a,b) is not satisfied. Hence X is a conflict-free set in AF.

Conversely, if X is a conflict-free set in AF and $x = \chi^X$, then x is a 0-1 vector, hence constraints (1) are trivially satisfied. For any attack $(a,b) \in D$ at most one of a and b are in X, therefore $x_a + x_b \in \{0,1\}$ and the constraints (2) are satisfied. □

Lemma 41 *Let $AF = (A,D)$ be an argumentation framework without isolated arguments. A* 0-1 *vector x satisfies the constraints* (3) *if and only if x is the incidence vector of a set $X \subseteq A$ with the property that $X^- \subseteq X^+$.*

Proof. Let x be a 0-1 vector satisfying the constraints (3) and let $X \subseteq A$ such that $\chi^X = x$. Let $b \in X^-$, that is there is $a \in X$ such that $(b,a) \in D$. By constraint (3), we have $1 = x_a \leq \sum_{c \in b^-} x_c$. It follows that $b^- \cap X \neq \emptyset$ and there is $a' \in A$ such that $a' \in b^-$ and $x_{a'} = 1$. Hence for each $b \in X^-$ there is $a' \in X$ such that $(a',b) \in D$, that is $X^- \subseteq X^+$.

Conversely, let $X \subseteq A$ with the property that $X^- \subseteq X^+$ and let $x = \chi^X$. Let $(b,a) \in D$. If $a \notin X$, then $x_a = 0$, and the constraint (3) trivially holds since x is a 0-1 vector.

If $a \in X$ then $b \in X^-$, and by the hypothesis there is $a' \in X$ such that $(a',b) \in D$. It follows that $\sum_{c \in b^-} x_c \geq x_{a'} = 1 = x_a$, that is the corresponding constraint (3) is satisfied. □

By Lemmas 39, 40 and 41, the following theorem holds.

Theorem 42 *Let $AF = (A,D)$ be an argumentation framework without isolated arguments. The integral vectors of $\mathrm{P}_{\mathrm{adm}}(AF)$ are exactly the incidence vectors of the admissible sets of AF.*

We introduce a set of linear constraints in order to enforce the "directionality principle": the non-attacked arguments must receive label 1, all argument attacked by these must receive label 0, all arguments attacked only by 0 labeled arguments must receive label 1, and so on.

Definition 43 *Let $AF = (A,D)$ be an argumentation framework. The **stable extensions polytope** of AF is the set $\mathrm{P}_{\mathrm{stab}}(AF)$ of all vectors in $\mathbb{R}^A$ satisfying:*

$$\begin{array}{lrcll} (1) & x_a & \geq & 0 & \forall a \in A, \\ (2) & x_a + x_b & \leq & 1 & \forall (b,a) \in D, \\ (4) & x_a + \displaystyle\sum_{b \in a^-} x_b & \geq & 1 & \forall a \in A. \end{array}$$

Example 3. In Figure 3.4 the $\mathrm{P}_{\mathrm{stab}}(AF)$ for the argumentation framework in Example 2 is illustrated. Note that in this particular argumentation framework each argument is attacked by exactly one argument, and constraints (2) and (4) give the equality constraints that appears in the middle of the figure.
The polytope $\mathrm{P}_{\mathrm{stab}}(AF)$ is in this case the line segment

$$\{\lambda x^1 + (1-\lambda)x^2 | 0 \leq \lambda \leq 1\} = \{(1-\lambda, \lambda, 1-\lambda)^T | 0 \leq \lambda \leq 1\},$$

where $x^1 = (0,1,0)^T$ and $x^2 = (1,0,1)^T$ are its vertices.

In order to relate the vectors in $\mathrm{P}_{stab}(AF)$ to complete extensions, we can use Theorem 36 and the following observation.

Proposition 44 *Let $AF = (A,D)$ be an argumentation framework without isolated arguments. The set of solutions in $[0,1]$ to the system of equations $Eq_{max}(AF)$ is contained in $\mathrm{P}_{\mathrm{stab}}(AF)$.*

Proof. Let $x \in [0,1]^A$ be a solution to the system of equations $Eq_{\max}(AF)$. Constraints (1) are clearly satisfied. Since $x_a = 1 - \max_{c \in a^-} x_c$ for all $a \in A$, it follows that if $(b,a) \in D$ we have $x_a + x_b \leq x_a + \max_{c \in a^-} x_c = 1 - \max_{c \in a^-} x_c + \max_{c \in a^-} x_c =$

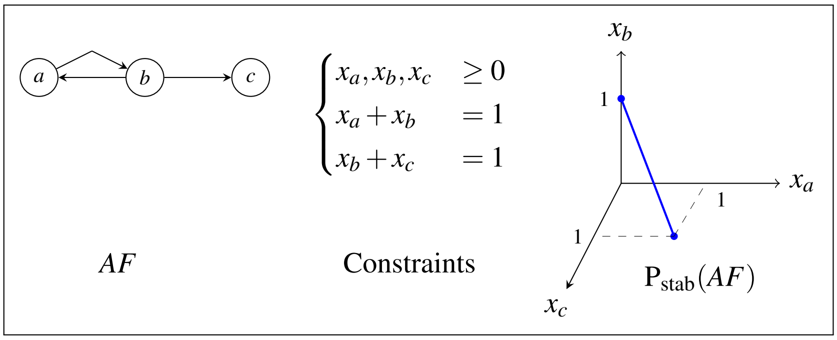


Figure 3.4.: A stable extensions polytope with integer vertices.

1, that is constraints (2) are satisfied by x. Also constraints (4) are satisfied, since for any $a \in A$ we have $x_a + \sum_{b \in a^-} x_b \geq x_a + \max_{c \in a^-} x_c = 1$. □

The integral vectors of $P_{stab}(AF)$ are interesting as the following Lemma shows.

Lemma 45 *Let $AF = (A, D)$ be an argumentation framework without isolated arguments. A 0-1 vector x satisfies the constraints (4) if and only if x is the incidence vector of a set $X \subseteq A$ with the property that $A - X \subseteq X^+$.*

Proof. Let x be a 0-1 vector satisfying the constraints (4) and let $X \subseteq A$ such that $\chi^X = x$. Let $a \in A - X$, that is $x_a = 0$. By constraint (4), we have $x_a + \sum_{b \in a^-} x_b \geq 1$, and since $x_a = 0$, we have $\sum_{b \in a^-} x_b \geq 1$. It follows that there is $b \in a^-$ such that $x_b = 1$, that is there is $b \in X$ such that $(b, a) \in D$. Therefore $A - X \subseteq X^+$.

Conversely, let $X \subseteq A$ with the property that $A - X \subseteq X^+$ and let $x = \chi^X$. We prove that the constraint (4) holds for every $a \in A$. If $a \in X$, then $x_a = 1$ and, since x is a 0-1 vector, the constraint (4) holds trivially for a. If $a \notin X$, then $x_a = 0$ and since $A - X \subseteq X^+$, there is $b_0 \in a^- \cap X$. Since $x_{b_0} = 1$ it follows that $x_a + \sum_{b \in a^-} x_b \geq 0 + x_{b_0} = 1$, that is the constraint (4) holds for a. □

By Lemmas 39, 40 and 45, the following theorem holds.

Theorem 46 *Let $AF = (A, D)$ be an argumentation framework without isolated arguments. The integral vectors of $\mathrm{P}_{\mathrm{stab}}(AF)$ are exactly the incidence vectors of the stable extensions of AF.*

The structure of $\mathrm{P}_{stab}(AF)$ is strongly dependent on the combinatorial structure of AF, more precisely on its family of circuits. We represent here a circuit in $AF = (A, D)$ as a sequence $C = (a_1, a_2, \ldots, a_k)$ of distinct arguments $a_i \in A$ such that $(a_i, a_{i+1}) \in D$, for each $i \in \{1, \ldots, k-1\}$, and $(a_k, a_1) \in D$. C is an even (odd) circuit if k is even (odd).

If AF has no circuits, then it is well known that the grounded extension is also a stable extension, so its incidence vector belongs to $\mathrm{P}_{stab}(AF)$, by Theorem 46.

For $x \in \mathrm{P}_{stab}(AF)$ we denote by

$$Frac(x) = \{a \in A | 0 < x_a < 1\}$$

its set of fractional components. Observe that if $a \in Frac(x)$ then, by constraint (2), we have $x_b < 1$ for all $b \in a^-$. Since, by constraint (4), we have $x_a + \sum_{b \in a^-} x_b \geq 1$, it follows that there is $b \in Frac(x)$ such that $(b,a) \in D$. Similarly, there is $c \in Frac(x)$ such that $(c,b) \in D$. Since $Frac(x)$ is a finite set, continuing the above argument we find a circuit $C = (a_1, a_2, \ldots, a_k)$ with all $a_i \in Frac(x)$. Hence, if AF has no circuits, then all vectors in $\mathrm{P}_{stab}(AF)$ are integral. Since $\mathrm{P}_{stab}(AF)$ is non-empty (containing the incidence vector of the grounded extension), it follows by convexity that it has exactly one point. Hence, the following Corollary of the Theorem 46 holds.

Corollary 47 *Let $AF = (A,D)$ be an argumentation framework without circuits. Then $\mathrm{P}_{\mathrm{stab}}(AF)$ has exactly one point, the incidence vector of grounded extension of AF.*

In order to characterize the incidence vectors of complete extensions of an argumentation framework AF, let us note that, in general, these vectors are not members of $\mathrm{P}_{\mathrm{stab}}(AF)$ as Lemma 45 shows. On the other hand, in the discussion preceding Corollary 47, we have argued that if $x \in \mathrm{P}_{\mathrm{stab}}(AF)$ is such that $Frac(x) \neq \emptyset$ then each argument in $Frac(x)$ has an attacker in the same set. Furthermore, if $x_a = 1$ then $a^+ \cap Frac(x) = \emptyset$ and $a^- \cap Frac(x) = \emptyset$ (by constraints (2)). These facts justify the following definition.

Definition 48 *Let $AF = (A,D)$ be an argumentation framework and $\mathrm{P}_{\mathrm{stab}}(AF)$ be its stable extensions polytope. A vector $x \in \mathrm{P}_{\mathrm{stab}}(AF)$ is called a* ***complete vector*** *if for each $a \in A$ such that $x_a = 0$ there is $b \in A$ such that $(b,a) \in D$ and $x_b = 1$.*

Note that if $x \in \mathrm{P}_{\mathrm{stab}}(AF)$ is such that $Frac(x) = A$, then x is a complete vector. The following characterization of **comp**(AF) holds.

Theorem 49 *Let $AF = (A,D)$ be an argumentation framework without isolated arguments. The integral vectors obtained from complete vectors of $\mathrm{P}_{\mathrm{stab}}(AF)$ by replacing the fractional components with 0 are exactly the incidence vectors of the complete extensions of AF.*

Proof. Let S be a complete extension of AF and $y = \chi^S$. Let $x \in \mathbb{R}^A$ defined by

$$x_a = \begin{cases} y_a & \text{if } y_a = 1, \\ 0 & \text{if } y_a = 0 \text{ and } \exists b \in S \text{ s.t. } (b,a) \in D, \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

Clearly, y is obtained from x by replacing its fractional components with 0. We show that x is a complete vector of $\mathrm{P}_{\mathrm{stab}}(AF)$. x trivially satisfies the constraints

(1), and constraints (2) hold since S is a conflict-free set. Let $a \in A$. If $x_a = 1$ then constraint (4) holds trivially. If $x_a = 0$ then $a \notin S$ and there is $b \in S$ such that $(b,a) \in D$. It follows that $y_b = x_b = 1$ and $\sum_{c \in a^-} x_c \geq x_b = 1$, hence constraint (4) holds. If $x_a = \frac{1}{2}$ it follows that $a \notin S$ and there is no $b \in S$ such that $(b,a) \in D$. Since S is a complete extension, we have $F(S) = S$. Since $a \notin S$, there is $b \in A - S \cup S^+$ such that $(b,a) \in D$. It follows that $x_b = \frac{1}{2}$. Hence $x_a + \sum_{c \in a^-} x_c \geq x_a + x_b = \frac{1}{2} + \frac{1}{2} = 1$, that is, the constraint (4) holds.

Conversely, let $x \in \mathrm{P}_{\mathrm{stab}}(AF)$ be a complete vector, and let $y \in \{0,1\}^A$ be the vector obtained from x by replacing its fractional components with 0. Let $S \subseteq A$ such that $\chi^S = y$. We show that S is a complete extension in AF. Clearly, S is a conflict-free set because x satisfies the constraints (2). Let $a \in S^-$, that is, there is $b \in S$ such that $(a,b) \in D$. Then $a \notin S$, hence $y_a = 0$, and moreover $x_a = 0$, because $x_a + x_b \leq 1$ and $x_b = 1$. Since x is a complete vector, it follows that there is $c \in S$ with $x_c = 1$ such that $(c,a) \in D$. Hence, we proved that $S^- \subseteq S^+$, that is S is an admissible set. Suppose that there is $a \notin S \cup S^+$ such that $a^- - (S \cup S^+) = \emptyset$. Since $a \notin S^+$, we must have $a \in Frac(x)$ and from the constraint (4) for a we have $x_a + \sum_{b \in a^-} x_b = x_a + 0 = x_a \geq 1$, contradiction. Hence we have obtained that S is an admissible set and $F(S) = S$, that is S is a complete extension. $\square$

Part II.

Graphs and Social Choice Theory

On a shared *reasons space*, a *society* expresses a set of possibly shared forms of *subjectivity*, whose deeper interactions enable new consistent collective judgments, creating *social inference relations*. Formally, this is done by considering a novel graph-based model for aggregating dichotomous preferences: *Bipartite Digraphs Debates*. The Chapter 4, based on the papers [Cro13, Cro14a, Cro15a], can be viewed as an attempt to integrate and exploit Dung's argumentation semantics to provide *argumentative aggregation of individual opinions*. The next chapter (based on the paper [CM]) introduces *Bipartite Choice Systems*, – abstracting (many-to-many) two-sided (labour) markets – and presents new properties of *choice functions*.

4. Argumentative Aggregation of Individual Opinions

A novel graph-based model for aggregating dichotomous preferences is introduced. The output *opinion* is viewed as a consensual situation, paving the way of using graph operations to describe properties of the aggregators. The outputs are also dichotomous preferences which could be useful in some applications. New axiomatic characterizations of aggregators corresponding to usual majority or approval & disapproval rule are presented. Integrating and exploiting Dung's Argumentation Frameworks and their semantics into our model is another contribution of the present chapter.

4.1. Introduction

Comparing and assessing different points of view in order to obtain *fair and rational* collective aggregation of them is the main research topic of *Social Choice Theory* (SCT) [Arr63], having major philosophical, economic, and political significance. The most important methodological tool in SCT is the axiomatic method, pioneered by Arrow [Arr50], which consists in formulating normatively desirable properties of aggregation rules as *postulates* or *axioms*, in order to obtain precise characterizations of the aggregation rules that satisfy these properties. The Artificial Intelligence developments, especially in the area of collective decision making in *Multiagent Systems*, have lead to the emergence of a new research area called *Computational Social Choice* (CSC), mainly concerned with the design and analysis of collective decision making mechanisms.

If in classical SCT the objects of aggregation belong to *preferential knowledge* (Arrow, Sen, and Suzumura [ASS02]), recent developments apply the same methodology to other types of information: *beliefs* (Konieczny and Pérez [KP02]), *judgments* (List and Puppe [LP09]), *ontologies* (Porello and Endriss [PE11]), *graphs* (Airiau, Endriss, Grandi, Porello, and Uckelman [AEG$^+$11], Endriss and Grandi [EG12]), and *argumentation frameworks* (Coste-Marquis, Devred, Konieczny, Lagasquie- Schiex, and Marquis [CMDK$^+$07], Dunne, Marquis, and Wooldridge [DMW12]).

Argumentation is a powerful mechanism for automating the decision making process of autonomous agents. Several recent works have studied the problem of accommodating ideas from CSC to Argumentation (Pigozzi [Pig06], Tohme, Bodanza, and Simari [TBS08], Rahwan, Larson, and Tohme [RLT09], Rahwan and Tohme [RT10], Caminada and Pigozzi [CP11], Dunne, Marquis, and Wooldridge [DMW12]). Most of them rely on Dung's *Argumentation Frameworks* and their acceptability semantics [Dun95].

The main objective of this chapter is to borrow ideas from Abstract Argumentation Frameworks to CSC, hence in the *converse direction of the above line of research on this subject*. We introduce and study a new graph-based model for aggregating dichotomous preferences. Although dichotomous preferences over alternatives may lack the expressiveness to capture intensity of the preference, they are natural in many settings, and are studied in different approaches of decision making systems mainly related to approval voting (see Brams and Fishburn [BF78], Laslier and Sanver [LS10], Vorsatz [Vor07], among others) or in connection to randomized mechanisms (Bogomolnaia and Moulin[BM04], Bogomolnaia, Moulin, and Stong [BMS05]).

Let us suggest a new possible application. In peer assessments systems (used in massive open online courses, or in evaluation of grant applications, see Walsh [Wal14], de Alfaro and Shavlovsky [AS14]) the main objective is to get a fair grade for each agent based on the grades proposed by some other agents. Uniform grading is an obvious desideratum for the quality of the outcome. An approach to make uniform the individual grading is to consider each grade given by an agent as good or bad, depending on how this grade compares to the average grade given by this agent.

In our dichotomous setting, we consider two disjoint non-empty finite sets $\mathbb{F}$ and $\mathbb{S}$. $\mathbb{F}$, referred as the set of *facts* (*issues*), is the set of alternatives in a decision making system, e.g. candidates in an election process, normative judgments, goods in purchasing systems, time slots in meeting scheduling systems, etc. Note that all facts are "positive", there are no negative facts. The *individuals* (*persons, agents*) in $\mathbb{S}$, the *society*, expresse their *opinions* (*dichotomous preferences*) that are pairs (L, DL) of disjoint subsets of $\mathbb{F}$: L is the set of facts on which the individual has a positive opinion while DL is the set of facts on which the individual has a negative opinion. The remaining facts in $\mathbb{F} - (L \cup DL)$ are indifferent or unknown to the individual. In applications in which $\mathbb{F}$ and $\mathbb{S}$ are not disjoint, we will consider disjoint copies of them.

Our focus is how to aggregate the individual opinions into a collective one. We are not only interested into positive collective position (as is intensively done in social choice theory) but also into the negative collective position (which can be used also as an explanation of the selected positive facts).

To illustrate our new model let us consider a simple mundane choice situation. Table 4.1 presents an example with available aliments, $F = \{a_1, a_2, a_3, a_4, a_5, a_6\} \subseteq \mathbb{F}$, and the preferences on these aliments of a group of five persons, $S = \{p_1, p_2, p_3, p_4, p_5\} \subseteq \mathbb{S}$, which want to have a common lunch.

	Liked	Disliked
p_1	$\{a_1, a_3\}$	$\{a_5, a_4\}$
p_2	$\{a_3, a_6\}$	$\{a_1, a_4\}$
p_3	$\{a_3, a_5\}$	$\{a_1, a_4\}$
p_4	$\{a_5, a_1, a_2\}$	$\{a_3\}$
p_5	$\{a_5, a_1\}$	$\{a_3, a_2\}$
Majority	$\{a_3, a_1, a_5\}$	$\{a_4\}$

Table 4.1.: Common Lunch Dilemma.

As we can see, p_1 likes (agrees) a_1 and a_3 but dislikes (disagrees) a_5 and a_4. Similarly, we can read the opinions of the other members of S. Note that disliking an aliment may mean that the individual is allergic to it. The table is entitled *Common Lunch Dilemma* since if we consider the *majority opinion* (include each fact in one of the two sets of liked and disliked facts using the majority rule) as output of the debate, then this has the unpleasant property that **each individual is allergic to an aliment in the collective output** $(\{a_3, a_1, a_5\}, \{a_4\})$! Note that this happens despite the majority rule gives (always) a consistent opinion, i.e., a disjoint pair of subsets of F.

The set of individual opinions listed in Table 4.1 (called "profile" in social choice theory) can be represented using a bipartite directed graph, which we call debate, as depicted in Figure 4.1. The two parts of a debate are the set of facts and the set of individuals. Each individual has as out-neighbors its set of liked facts and as in-neighbors the set of disliked facts. Since these two sets are disjoint, we have no symmetric pair of directed edges.

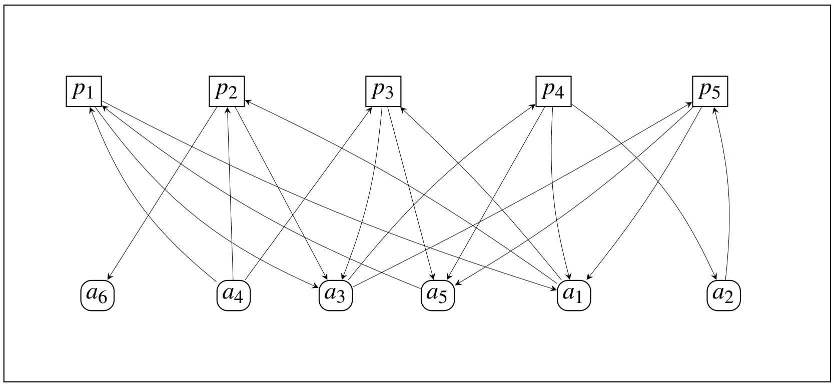


Figure 4.1.: Common Lunch Dilemma – Bipartite Digraph Debate.

The main advantage of this model consists in its visual capacity and the symmetric treatment of the facts and individuals. Also well-established graph theoretical notions (digraph isomorphism, in-degree and out-degree of a vertex, induced subdigraphs, digraph operations, etc.) can be used to describe normative conditions on the aggregating rules.

The aggregate opinion (a disjoint pair of collective agreed and disagreed facts) can also be viewed as a debate (bipartite digraph) in which each individual has the same sets of in-neighbors and out-neighbors (called consensual debate). In this way, axioms on the aggregating rules gain in expressivity (see Section 4.3).

To each digraph without loops we can associate a bipartite digraph, hence a debate. On the other hand, to each debate we can associate two conflict digraphs. These conflict digraphs are isomorphic to the digraph for which the debate is associated. This relationship is exploited in Section 4.5 to describe and analyze, from computational point of view, new (irresolute) aggregators corresponding to the digraphs representing Dung's argumentation frameworks. These gives acceptable solutions to the Common Lunch debate discussed above.

The basic idea of the *argumentative aggregation of individual opinions* is to consider *collective opinions* by merging the opinions of non-conflicting coalitions of individuals. A coalition is conflict-free if the individual's opinions in the coalition does not attack each other. Such a coalition is called an *autarky* if, in addition, has the property that the collective opinion counterattacks any attack of the opinion of an individual not in coalition. This property offers a rational justification for the output opinion.

> Our main rationality hypothesis is that each member of the society believes in the positive value of the collective output O. Therefore if f – a particular proposal (fact) of an individual – is not selected, then the individual understands the reason: f or other fact, that he likes, is disliked by the society's opinion. Moreover, if an individual dislikes a fact agreed by the collective opinion O, the reason is his positive position on a fact disliked by O.

4.2. Graph Based Framework

In this section we introduce our new model for aggregating dichotomous preferences and present graph theoretical concepts and notations used in the next sections. Recall from the Introduction the two disjoint finite non-empty sets $\mathbb{F}$ and $\mathbb{S}$ of *facts* and, respectively, *individuals*.

Definition 50 A *Debate* is a bipartite digraph $D = (F, S; E)$, where $\emptyset \neq F \subseteq \mathbb{F}$ and $\emptyset \neq S \subseteq \mathbb{S}$, and $E \subseteq F \times S \cup S \times F$ contains no symmetric pair of directed edges (i.e., at most one of the pairs (f, s) and (s, f) is a directed edge in E, for every $f \in F$ and $s \in S$).

Let $G = (V,E)$ be a digraph and $v \in V$ a vertex of G. The set of out-neighbors of v is denoted by v_G^+, that is $v_G^+ = \{u \in V | (v,u) \in E\}$. Similarly, the set of in-neighbours of v is $v_G^- = \{u \in V | (u,v) \in E\}$. These notations can be extended to set of vertices by considering, for every $S \subseteq V$, $S_G^+ = \cup_{v \in S} v_G^+$ and $S_G^- = \cup_{v \in S} v_G^-$ (clearly, $\emptyset_G^+ = \emptyset_G^- = \emptyset$).

If $D = (F,S;E)$ is a debate then, for every $s \in S$, s_D^+ is the set of **facts approved by the individual** s and s_D^- is the set of **facts disapproved by the individual** s. The pair (s_D^+, s_D^-) is referred as the **opinion** of individual s on the facts in F. By the above definition of a debate , $s_D^+ \cap s_D^- = \emptyset$.

If $f \in s_D^+$ then s has a "positive" opinion on f, if $f \in s_D^-$ then s has a "negative" opinion on f, and if $f \notin s_D^+ \cup s_D^-$ then s has no opinion on f.

Let $D = (F,S;E)$ be a debate. If $F' \subseteq F$, the **sub-debate** induced by F' is the sub-digraph induced by $F' \cup S$ in D, and is denoted by $D^{F'}$. If $S' \subseteq S$, the **sub-debate** induced by S' is the sub-digraph induced by $F \cup S'$ in D, and is denoted by $D_{S'}$. For $s \in S$, the sub-debate $D_{S-\{s\}}$ is denoted by $D - s$.

If $D_i = (F_i, S_i; E_i)$ $(i = 1,2)$ are debates with $S_1 \cap S_2 = \emptyset$, then their **sum** is $D_1 + D_2 = (F_1 \cup F_2, S_1 \cup S_2; E_1 \cup E_2)$. Clearly, $E_1 \cup E_2$ does not contain symmetric edges, that is, $D_1 + D_2$ is a debate. With this notation, if $D = (F,S;E)$ is a debate such that $|S| \geq 2$ then, for every $s \in S$, we have $D = D_{S-\{s\}} + D_{\{s\}} = (D - s) + D_{\{s\}}$.

Two debates $D = (F,S;E)$ and $D' = (F', S'; E')$ are **isomorphic** if there are bijections $\alpha : F \to F'$ and $\beta : S \to S'$ such that for all $f \in F$ and $s \in S$ $(f,s) \in E$ if and only if $(\alpha(f), \beta(s)) \in E'$, and $(s,f) \in E$ if and only if $(\beta(s), \alpha(f)) \in E'$. Two isomorphic debates are denoted by $D \cong D'$ or $D \cong_{\alpha,\beta} D'$ (when we need to emphasize the isomorphism).

Definition 51 A debate $D = (F,S;E)$ is a *consensual debate* if there are $F_D^{\oplus}, F_D^{\ominus} \subseteq F$ such that $F_D^{\oplus} \cap F_D^{\ominus} = \emptyset$ and, for every $s \in S$, we have $(s_D^+, s_D^-) = (F_D^{\oplus}, F_D^{\ominus})$.
$O_D = (F_D^{\oplus}, F_D^{\ominus})$ is called the *common opinion* of D.

4.3. Aggregators

In this section we define opinions aggregation, describe our versions of the well-known majority and approval&disapproval rules and prove their axiomatic characterization.

Definition 52 Let $\mathscr{D}(\mathbb{F},\mathbb{S})$ be the set of all debates $D = (F,S;E)$ with $F \subseteq \mathbb{F}$ and $S \subseteq \mathbb{S}$. An *aggregator* is a function $\mathbb{A} : \mathscr{D}(\mathbb{F},\mathbb{S}) \to \mathscr{D}(\mathbb{F},\mathbb{S})$ such that

- $\mathbb{A}(D)$ is a consensual debate for every $D \in \mathscr{D}(\mathbb{F},\mathbb{S})$, and
- if $D_1 \cong D_2$ then $\mathbb{A}(D_1) \cong \mathbb{A}(D_2)$.

In words, an aggregator is a functional $\mathbb{A}$ that maps each debate D into a consensual output debate $\mathbb{A}(D)$, in which every individual has the same opinion $O_{\mathbb{A}(D)} = (F^{\oplus}_{\mathbb{A}(D)}, F^{\ominus}_{\mathbb{A}(D)})$. Hence, we can represent $\mathbb{A}(D)$ by specifying this common opinion $(F^{\oplus}_{\mathbb{A}(D)}, F^{\ominus}_{\mathbb{A}(D)})$. Also, $\mathbb{A}$ satisfies the usual social choice theory conditions of *neutrality* and *anonymity* (renaming the facts or the individuals does not change the output modulo this renaming).

We give now two examples of aggregators, corresponding to well-known social choice theory rules. In both examples, the first condition in the above definition is satisfied by construction and the second condition is satisfied since the output consensual debate depends only on the in-degree and out-degree of the fact vertices, and these are invariant under debate isomorphisms. After defining each of these aggregators we consider normative properties of them and prove that these offer novel interesting characterizations.

Majority rule. $\mathbb{A}_M : \mathscr{D}(\mathbb{F},\mathbb{S}) \to \mathscr{D}(\mathbb{F},\mathbb{S})$ such that $\forall D \in \mathscr{D}(\mathbb{F},\mathbb{S})$,

$$O_{\mathbb{A}_M(D)} = (F^{\oplus}_{\mathbb{A}_M(D)}, F^{\ominus}_{\mathbb{A}_M(D)}), \text{ where}$$

$$F^{\oplus}_{\mathbb{A}_M(D)} = \{f \in F \mid |f^-_D| \geq \frac{|S|}{2}\},$$

$$F^{\ominus}_{\mathbb{A}_M(D)} = \{f \in F \mid |f^+_D| > \frac{|S|}{2}\}.$$

Note that, by definition, a fact f is approved (disapproved) by the collective opinion if the number of individuals that like (dislike) f is at least (greater than) half of the total number of individuals. It is not difficult to see that $F^{\oplus}_{\mathbb{A}_M(D)} \cap F^{\ominus}_{\mathbb{A}_M(D)} = \emptyset$.

Also, for every $f \in F$ we have $\mathbb{A}_M(D)^{\{f\}} = \mathbb{A}_M(D^{\{f\}})$, that is, the aggregate opinion on f depends only on the opinions of the individuals in the society on f: to find the aggregate opinion on f, we apply the aggregator on the debate $D^{\{f\}}$ obtained by considering the restriction of D to $\{f\}$ only. This is the usual social choice theory **Independence (I)** condition, that is, the aggregation is done **fact-wise**:

I For every debate $D = (F,S;E)$ and for every $f \in F$ $\mathbb{A}(D)^{\{f\}} = \mathbb{A}(D^{\{f\}})$.

In order to characterize the majority rule, we consider also the following conditions **Unanimity (U)**, **Cancellation (C)**, and **Faithfulness (F)**:

U If D is a consensual debate then $\mathbb{A}(D) = D$.

C For every debate $D = (\{f\},S;E)$ with $|S| \geq 3$, if $s,p \in S$ are such that $f \in s^+_D \cap p^-_D \cup s^-_D \cap p^+_D$, then $\mathbb{A}(D) = \mathbb{A}(D_{S-\{s,p\}})$.

F If $D = (\{f\},\{s,p\};E)$ is a debate such that $f \in s^+_D \cap p^-_D \cup s^-_D \cap p^+_D$, then $O_{\mathbb{A}(D)} = (F^{\oplus}_{\mathbb{A}(D)}, F^{\ominus}_{\mathbb{A}(D)}) = (\{f\},\emptyset)$.

In words, *Cancellation* says that in any debate over a single fact, if there are at least three individuals and two of them have contradictory opinions on this fact, then the output opinion is decided by the remaining individuals. *Faithfulness* says that the output opinion of a debate over a single fact with exactly two individuals with contradictory opinions has a positive position on this fact.

Theorem 53 *The Majority rule, $\mathbb{A}_M$, is the only aggregator $\mathbb{A}$ satisfying conditions* **U, I, C,** *and* **F**.

Proof. Obviously, $\mathbb{A}_M$ satisfies **U, I,** and **F**. To prove that $\mathbb{A}_M$ fulfills **C**, let $D = (\{f\}, S; E)$ be a debate with $|S| \geq 3$, and $s, p \in S$ such that $f \in s_D^+ \cap p_D^-$ (the proof is similar for $f \in s_D^- \cap p_D^+$). Since $f^-_{D_{S-\{p,s\}}} = f_D^- - \{s\}$, it follows that $|f^-_{D_{S-\{p,s\}}}| \geq \frac{|S|-2}{2}$ if and only if $|f_D^-| \geq \frac{|S|}{2}$. Similarly, since $f^+_{D_{S-\{p,s\}}} = f_D^+ - \{p\}$, it follows that $|f^+_{D_{S-\{p,s\}}}| > \frac{|S|-2}{2}$ if and only if $|f_D^+| > \frac{|S|}{2}$. Hence $\mathbb{A}_M(D) = \mathbb{A}_M(D_{S-\{s,p\}})$.

Conversely, let $\mathbb{A}$ be an aggregator satisfying **U, I, C**, and **F**. We prove that $\mathbb{A}(D) = \mathbb{A}_M(D)$ for every debate $D = (F, S; E)$, by induction on $|S|$.

If $|S| = 1$, then D is consensual and by **U** we have $\mathbb{A}(D) = D$ and, since $\mathbb{A}_M$ satisfies **U**, $\mathbb{A}(D) = \mathbb{A}_M(D)$. Also, for $|S| = 2$, for every $f \in S$, $D^{\{f\}}$ is either consensual and $\mathbb{A}(D)^{\{f\}} = \mathbb{A}_M(D^{\{f\}})$ by **U**, or satisfies the hypothesis of **F** and again $\mathbb{A}(D)^{\{f\}} = \mathbb{A}_M(D^{\{f\}})$. By **I**, we have $\mathbb{A}(D) = \mathbb{A}_M(D)$.

In the inductive step, let $D = (F, S; E)$ be a debate with $|S| \geq 3$.
By **I**, in order to prove that $\mathbb{A}(D) = \mathbb{A}_M(D)$ it is sufficiently to prove that $\mathbb{A}(D)^{\{f\}} = \mathbb{A}_M(D^{\{f\}})$ for each $f \in F$. This follows either by **U** or by applying **C** and the induction hypothesis. □

Note that the above new axiomatization of the majority rule (a subject started by May [May52] and followed by several papers, e.g. Maskin [Mas95], Woeginger [Woe03], Miroiu [Mir04], etc.) benefits by the capacity of our framework to express properties of aggregators in terms of simple graph operations.

Approval&Disapproval rule. $\mathbb{A}_{\text{A\&D}} : \mathscr{D}(\mathbb{F}, \mathbb{S}) \to \mathscr{D}(\mathbb{F}, \mathbb{S})$ such that $\forall D \in \mathscr{D}(\mathbb{F}, \mathbb{S})$,

$$O_{\mathbb{A}_{\text{A\&D}}(D)} = (F^{\oplus}_{\text{A\&D}}, F^{\ominus}_{\text{A\&D}}), \text{ where}$$

$$F^{\oplus}_{\text{A\&D}} = \{f \in F |\ |f_D^-| - |f_D^+| \geq |g_D^-| - |g_D^+|, \forall g \in F\},$$
$$F^{\ominus}_{\text{A\&D}} = \{f \in F - F^{\oplus}_{\text{A\&D}} |\ |f_D^+| - |f_D^-| \geq |g_D^+| - |g_D^-|, \forall g \in F - F^{\oplus}_{\text{A\&D}}\}.$$

Let us consider the **score** of $f \in F$ as $score_D(f) = |f_D^-| - |f_D^+|$, that is the difference between the number of individuals, $|f_D^-|$, having a positive position on f, and the number of individuals, $|f_D^+|$, having a negative position on f. Hence, the facts maximizing this score are selected in the positive part of the aggregator's opinion. From the remaining facts, those having the minimum score are included in the negative part of the aggregator's opinion. Clearly, in this case, the aggregation is not fact-wise: despite computing the scores is done fact-wise, the decision of the aggregator on a fact depends on the scores obtained by the other facts.

A characterization of Approval&Disapproval rule can be obtained by considering the above **Unanimity (U)** condition and the following two new conditions: **Summation (S)** and **Additivity (A)**.

S For every consensual debate $D_1 = (F_1, S_1; E_1)$ with $O_{D_1} = (F_1^{\oplus}, F_1^{\ominus})$ and for every debate $D_2 = (F_2, \{s\}; E_2)$ with $s \notin S_1$ and $O_{D_2} = (F_2^{\oplus}, F_2^{\ominus})$, the aggregate debate of their sum, $\mathbb{A}(D_1 + D_2)$, is such that $O_{\mathbb{A}(D_1+D_2)} = (F^{\oplus}, F^{\ominus})$, where

$$F^{\oplus} = \begin{cases} F_1^{\oplus} \cap F_2^{\oplus} & \text{if } F_1^{\oplus} \cap F_2^{\oplus} \neq \emptyset, \\ F_1^{\oplus} & \text{if } F_1^{\oplus} \cap (F_2^{\oplus} \cup F_2^{\ominus}) = \emptyset \text{ and } |S_1| > 1, \\ F_1^{\oplus} \cup F_2^{\oplus} & \text{if } F_1^{\oplus} \cap (F_2^{\oplus} \cup F_2^{\ominus}) = \emptyset \text{ and } |S_1| = 1, \end{cases}$$

and

$$F^{\ominus} = \begin{cases} F_1^{\ominus} \cap F_2^{\ominus} & \text{if } F_1^{\ominus} \cap F_2^{\ominus} \neq \emptyset, \\ F_1^{\ominus} & \text{if } F_1^{\ominus} \cap (F_2^{\oplus} \cup F_2^{\ominus}) = \emptyset \text{ and } |S_1| > 1, \\ F_1^{\ominus} \cup F_2^{\ominus} & \text{if } F_1^{\ominus} \cap (F_2^{\oplus} \cup F_2^{\ominus}) = \emptyset \text{ and } |S_1| = 1. \end{cases}$$

A For every debate $D = (F, S; E)$ with $|S| \geq 2$,

$$\mathbb{A}(D) = \mathbb{A}\big(\mathbb{A}(D - s) + \mathbb{A}(D_{\{s\}})\big), \text{ for every } s \in S.$$

In words, *Additivity* says that in any debate with at least two individuals the output consensual debate is the aggregate debate of the sum of the (consensual) sub-debate induced by any individual and the consensual aggregate debate of the debate obtained by deleting this individual. *Summation* shows how to obtain the aggregate debate of the sum between a consensual debate and a debate with a single individual.

Theorem 54 *The Approval&Disapproval rule, $\mathbb{A}_{A\&D}$, is the only aggregator $\mathbb{A}$ satisfying conditions* **U, S**, *and* **A**.

Proof. We show first that $\mathbb{A}_{\mathrm{A\&D}}$ satisfies **U, S**, and **A**.

U If $D = (F, S; E)$ is a consensual debate with $O_D = (F^{\oplus}, F^{\ominus})$, then

$$score_D(f) = \begin{cases} |S| & \text{if } f \in F^{\oplus}, \\ -|S| & \text{if } f \in F^{\ominus}, \\ 0 & \text{if } f \in F - (F^{\oplus} \cup F^{\ominus}). \end{cases}$$

Since $|S| \geq 1$, it follows that $F^{\oplus}_{\mathrm{A\&D}} = F^{\oplus}$ and $F^{\ominus}_{\mathrm{A\&D}} = F^{\ominus}$, that is, $\mathbb{A}_{\mathrm{A\&D}}(D) = D$.

S Let $D_1 = (F_1, S_1; E_1)$ be a consensual debate with $O_{D_1} = (F_1^{\oplus}, F_1^{\ominus})$, and $D_2 = (F_2, \{s\}; E_2)$ be a debate with $s \notin S_1$ and $O_{D_2} = (F_2^{\oplus}, F_2^{\ominus})$. Then, in the debate $D = D_1 + D_2$, the $score_D(f)$, for $f \in F_1 \cup F_2$, is:

$$score_D(f) = \begin{cases} |S_1| + 1 & \text{if } f \in F_1^{\oplus} \cap F_2^{\oplus} \\ |S_1| - 1 & \text{if } f \in F_1^{\oplus} \cap F_2^{\ominus} \\ |S_1| & \text{if } f \in F_1^{\oplus} - (F_2^{\oplus} \cup F_2^{\ominus}) \\ -|S_1| + 1 & \text{if } f \in F_1^{\ominus} \cap F_2^{\oplus} \\ -|S_1| - 1 & \text{if } f \in F_1^{\ominus} \cap F_2^{\ominus} \\ -|S_1| & \text{if } f \in F_1^{\ominus} - (F_2^{\oplus} \cup F_2^{\ominus}) \\ 1 & \text{if } f \in (F_1 - (F_1^{\oplus} \cup F_1^{\ominus})) \cap F_2^{\oplus} \\ -1 & \text{if } f \in (F_1 - (F_1^{\oplus} \cup F_1^{\ominus})) \cap F_2^{\ominus} \\ 0 & \text{if } f \notin F_1^{\oplus} \cup F_1^{\ominus} \cup F_2^{\oplus} \cup F_2^{\ominus}. \end{cases}$$

Now, it is easy to see that if $O_{\mathbb{A}_{\text{A\&D}}(D)} = (F^{\oplus}_{\text{A\&D}}, F^{\ominus}_{\text{A\&D}})$, then $F^{\oplus}_{\text{A\&D}} = F^{\oplus}$ and $F^{\ominus}_{\text{A\&D}} = F^{\ominus}$, where $F^{\oplus}$ and $F^{\ominus}$ are defined in condition **S**.

A Let $D = (F, S; E)$ be a debate with $|S| \geq 2$.
For every $s \in S$ we have $D = (D - s) + D_{\{s\}}$. Using **U** and **S** it is not difficult to verify that $\mathbb{A}_{\text{A\&D}}(D) = \mathbb{A}_{\text{A\&D}}\big(\mathbb{A}_{\text{A\&D}}(D - s) + \mathbb{A}_{\text{A\&D}}(D_{\{s\}})\big)$.

Conversely, let $\mathbb{A}$ be an aggregator satisfying **U, S**, and **A**. We prove that $\mathbb{A}(D) = \mathbb{A}_{\text{A\&D}}(D)$ for every debate $D = (F, S; E)$, by induction on $|S|$.

If $|S| = 1$, then D is consensual and by **U** we have $\mathbb{A}(D) = D$ and, since $\mathbb{A}_{\text{A\&D}}$ satisfies **U**, $\mathbb{A}(D) = \mathbb{A}_{\text{A\&D}}(D)$. In the inductive step, by **A**, $\mathbb{A}(D) = \mathbb{A}\big(\mathbb{A}(D - s) + \mathbb{A}(D_{\{s\}})\big)$ for $s \in S$. By the induction hypothesis and since $D_{\{s\}}$ is consensual we have $\mathbb{A}\big(\mathbb{A}(D - s) + \mathbb{A}(D_{\{s\}})\big) = \mathbb{A}\big(\mathbb{A}_{\text{A\&D}}(D - s) + \mathbb{A}_{\text{A\&D}}(D_{\{s\}})\big)$. Since $\mathbb{A}_{\text{A\&D}}$ satisfies **S**, $\mathbb{A}_{\text{A\&D}}(D - s) + \mathbb{A}_{\text{A\&D}}(D_{\{s\}}) = \mathbb{A}_{\text{A\&D}}(D)$ and since $\mathbb{A}$ satisfies **U** and $\mathbb{A}_{\text{A\&D}}(D)$ is consensual, we obtain $\mathbb{A}(D) = \mathbb{A}_{\text{A\&D}}(D)$. □

Again, note that the above new axiomatization of the Approval&Disapproval rule (a subject starting with Brams and Fishburn [BF78], followed by several papers, see Xu [Xu10]) is different from that given in Alcantud and Laruelle [AL13] due to the capacity of our framework to express properties of aggregators in terms of simple graph operations.

4.4. Irresolute Aggregation

In order to introduce new principles in doing debate aggregation, we consider *aggregation correspondences*, which map every debate into a set of consensual debates, such that by specifying a rule of selecting a member of this set we obtain an aggregator.

If $F \subseteq \mathbb{F}$, then we denote by $\mathscr{O}(F)$ the set of all F-opinions, that is,

$$\mathscr{O}(F) = \{(F^{\oplus}, F^{\ominus}) | F^{\oplus}, F^{\ominus} \subseteq F \text{ and } F^{\oplus} \cap F^{\ominus} = \emptyset\}.$$

Definition 55 *An **aggregation correspondence** is a function $\mathbb{AC}$ which maps every debate $D \in \mathscr{D}(F,S)$, into a set of F-opinions $\mathbb{AC}(D) \subseteq \mathscr{O}(F)$ such that if $D_1, D_2 \in \mathscr{D}(F,S)$ are isomorphic debates , $D_1 \cong_{\alpha,\beta} D_2$, then $\mathbb{AC}(D_2) = \alpha(\mathbb{AC}(D_1))$.*

Clearly, each $O = (F^{\oplus}, F^{\ominus}) \in \mathbb{AC}(D)$ determines a consensual debate D_O in which each individual opinion is O. Hence, an aggregation correspondence maps every debate into a set of consensual debates, and if we devise a rule to select one from this set, we obtain an aggregator. This qualitative way of aggregation will be exploited in the next section for *argumentative aggregation*. Here we consider two new possible aggregation correspondences which are interesting in themselves.

SAT Aggregation. $\mathbb{AC}_{SAT} : \mathscr{D}(F,S) \to 2^{\mathscr{O}(F)}$, defined by

$$\mathbb{AC}_{SAT}(D) = \{(F^{\oplus}, F^{\ominus}) \in \mathscr{O}(F) \,|\, F^{\oplus} \cap s_D^+ \cup F^{\ominus} \cap s_D^- \neq \emptyset, \forall s \in S\}.$$

In words, an opinion $(F^{\oplus}, F^{\ominus})$ belongs to $\mathbb{AC}_{SAT}(D)$ if it *gets satisfaction* to each person s in the society: $(F^{\oplus}, F^{\ominus})$ agrees with (s_D^+, s_D^-) on at least one fact. This seems a very permissive and simple semantics, but the following example shows that it is not the case. Let $\Phi = C_1 \wedge \ldots \wedge C_m$ a boolean CNF formula, on the set V of variables, where each clause C_i is a disjunction of positive or negative literals. We can associate to Φ a debate $D_\Phi = (F_\Phi, S_\Phi; E_\Phi)$, where for each $v \in V$ we have a fact $f_v \in F_\Phi$, and for each clause $C_i \in \Phi$ we have an individual $s_{C_i} \in S_\Phi$ with $(s_{C_i})^+_{D_\Phi}$ = *the set of facts associated to its positive literals*, and $(s_{C_i})^-_{D_\Phi}$ = *the set of facts associated to its negative literals*. An example of this construction is given in Figure 4.2 below.

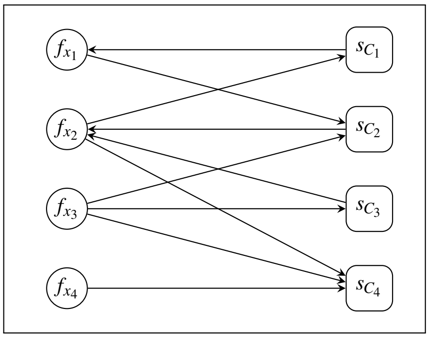


Figure 4.2.: Debate D_Φ associated to $\Phi = C_1 \wedge C_2 \wedge C_3 \wedge C_4$, where $C_1 = x_1 \vee \bar{x}_2$, $C_2 = \bar{x}_1 \vee x_2 \vee \bar{x}_3$, $C_3 = x_2 \vee \bar{x}_3$, $C_4 = \bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4$.

It is not difficult to prove the following proposition.

Proposition 56 *Let Φ be a boolean CNF formula on a set V of variables, and let $D_\Phi = (F_\Phi, S_\Phi; E_\Phi)$ its associated debate. If α is a satisfying assignment for Φ, then $O_\alpha = (F_\alpha^\oplus, F_\alpha^\ominus)$ is an F_Φ-opinion belonging to $\mathbb{AC}_{SAT}(D_\Phi)$, where $F_\alpha^\oplus = \{f_x | \alpha(x) = true\}$ and $F_\alpha^\ominus = \{f_x | \alpha(x) = false\}$. Conversely, if $O = (F^\oplus, F^\ominus) \in \mathbb{AC}_{SAT}(D_\Phi)$, then the assignment α_O, defined by*

$$\alpha_O(x) = \begin{cases} true & if\ f_x \in F^\oplus \\ false & if\ f_x \in F^\ominus \end{cases},$$

is a satisfying assignment for Φ.

Proof. The proof follows from the construction of D_Φ and the definition of the aggregation correspondence $\mathbb{AC}_{SAT}$. □

It follows that the well known SAT problem is (polynomial) equivalent to the problem of deciding if a given abstract debate has at least one **sat**-opinion.

Conceptual Aggregation

The second aggregation correspondence considered in this section is based on the ideas in *Formal Concept Analysis* (FCA), introduced by Wille in [Wil82].

Definition 57 *A **formal context** is an undirected bipartite graph $G = (O, A; E)$. Members $o \in O$ are called **objects** and the elements $a \in A$ are (unary) **attributes**. A **formal concept** for the context $G = (O, A; E)$ is a pair $C = (O_e, A_i)$ where $O_e \subseteq O$ and $A_i \subseteq A$ such that*

(i) *$\forall o \in O_e$ and $\forall a \in A_i$: $\{o, a\} \in E$,*

(ii) *$\forall o \in O - O_e$ $\exists a \in A_i$ such that $\{o, a\} \notin E$, and*

(iii) *$\forall a \in A - A_i$ $\exists o \in O_e$ such that $\{o, a\} \notin E$.*

*O_e is called the **extent** of the concept C and A_i is the **intent** of the concept C.*

In words, the pair $C = (O_e, A_i)$ is a formal concept in the formal context $G = (O, A; E)$ if $O_e \subseteq O$, $A_i \subseteq A$, and $O_e \cup A_i$ induces a complete bipartite subgraph of G that is maximal with this property.

If $D = (F, S; E)$ is a debate, then we can consider the following two formal contexts:

- **The formal context corresponding to the like relation:** $D^\oplus = (F, S; E^\oplus)$, the undirected bipartite graph obtained from D by deleting all arcs from F to S and replacing each arc $(s, f) \in E$ by the undirected edge $\{s, f\}$, and
- **The formal context corresponding to the dislike relation:** $D^\ominus = (F, S; E^\ominus)$, the undirected bipartite graph obtained from D by deleting all arcs from S to F and replacing each arc $(f, s) \in E$ by the undirected edge $\{f, s\}$.

Hence, we replace the attributes in usual *Formal Concept Analysis* by subjective like/dislike opinions of the members of society. Corresponding to the two formal contexts, we obtain **liked/disliked concepts**.

Let us denote by $\mathscr{C}^{\oplus}$ the set of all **liked concepts** for the context $D^{\oplus}$, and by $\mathscr{C}^{\ominus}$ the set of all **disliked concepts** for the context $D^{\ominus}$. The extents of these formal concepts form the sets $\mathscr{F}_e^{\oplus}$, respectively $\mathscr{F}_e^{\ominus}$.

For example, in Figure 4.3 below, a debate D and its corresponding formal contexts $D^{\oplus}$ and $D^{\ominus}$ are depicted.

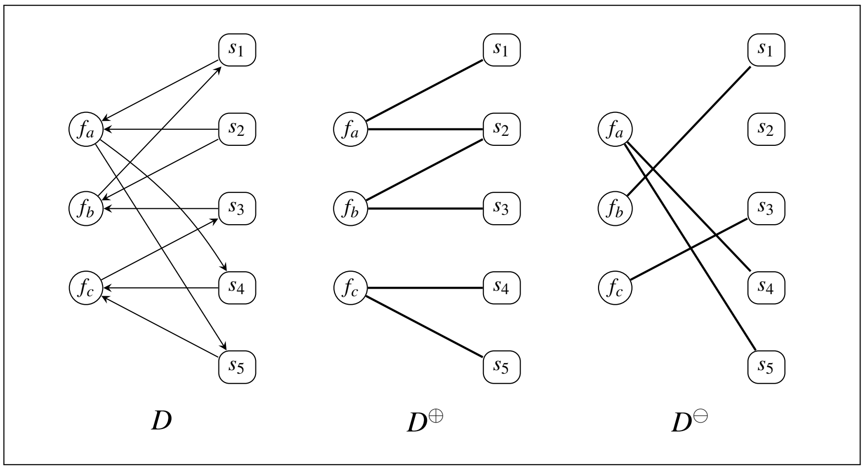


Figure 4.3.: Liked and disliked formal contexts associated to a debate D.

It is not difficult to see that

$$\mathscr{C}^{\oplus} = \big\{(\{f_a\},\{s_1,s_2\}),(\{f_a,f_b\},\{s_2\}),(\{f_b\},\{s_2,s_3\}),(\{f_c\},\{s_4,s_5\})\big\},$$

hence $\mathscr{F}_e^{\oplus} = \big\{\{f_a\},\{f_b\},\{f_c\},\{f_a,f_b\}\big\}$. Similarly,

$$\mathscr{C}^{\ominus} = \big\{(\{f_a\},\{s_4,s_5\}),(\{f_b\},\{s_1\}),(\{f_c\},\{s_3\})\big\}$$

and $\mathscr{F}_e^{\ominus} = \big\{\{f_a\},\{f_b\},\{f_c\}\big\}$.

Definition 58 *The **conceptual aggregation correspondence** is the map* $\mathbb{AC}_{Con} : \mathscr{D}(F,S) \to 2^{\mathscr{O}(F)}$, *defined by*

$$\mathbb{AC}_{Con}(D) = \{(F^{\oplus},F^{\ominus}) \in \mathscr{O}(F) \,|\, F^{\oplus} \in \mathscr{F}_e^{\oplus} - \mathscr{F}_e^{\ominus}, F^{\ominus} \in \mathscr{F}_e^{\ominus} - \mathscr{F}_e^{\oplus}\}.$$

In words, an opinion belongs to the conceptual aggregation correspondence image of a debate if and only if it is a pair of disjoint extents of a liked (which is not disliked) and a disliked (which is not liked) formal concepts. For the above example, this gives $\mathbb{AC}_{Con}(D) = \{(\{f_a,f_b\},\emptyset)\}$. Note, that the conceptual aggregation correspondence could be defined in a more permissive way by using some obvious statistics (similar to those used by the Approval&Disapproval rule).

4.5. Argumentative Aggregation

We can associate to each argumentation framework (without self-attacking arguments) an abstract debate, supporting the idea of collective rationality expressed by the admissibility based extensions.

Definition 59 *Let AF be an argumentation framework such that $(a,a) \notin Def(AF)$, $\forall a \in Arg(AF)$. The* ***debate associated*** *to AF is $D_{AF} = (F_{AF}, S_{AF}; E_{AF})$, where*

- $F_{AF} = \{f_a | a \in Arg(AF)\}$,
- $S_{AF} = \{s_a | a \in Arg(AF)\}$, *and*
- $E_{AF} = \{(s_a, f_a) | a \in Arg(AF)\} \cup \{(f_a, s_b) | (b,a) \in Def(AF)\}$.

D_{AF} is a very particular debate: each fact f_a is liked (approved) by exactly one individual s_a, and each individual's opinions are single minded ($s_a^+ = \{f_a\}$). In words, the individual s_a agrees f_a in D_{AF} and disagrees all f_b for the arguments b attacked by a. An example is illustrated in Figure 4.4 below.

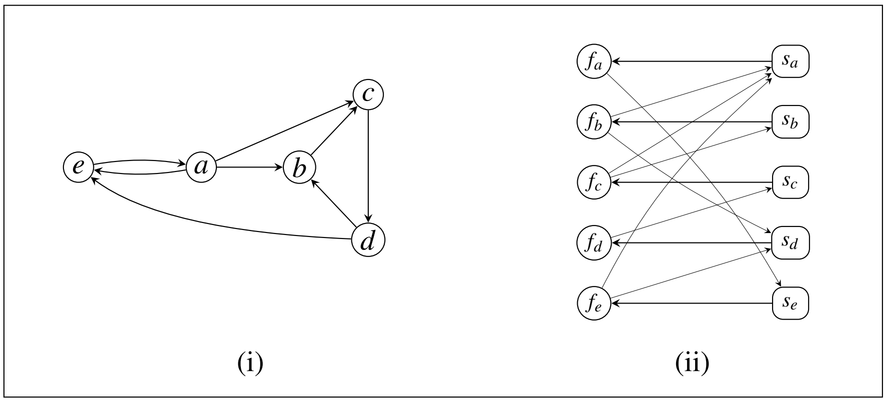


Figure 4.4.: (i) An AF. (ii) The debate D_{AF} associated to AF.

The conflicts between individual opinions in a debate can be naturally viewed as argumentation frameworks in order to use the collective acceptance of Dung's semantics in aggregation.

Definition 60 *If $D = (F,S;E)$ is a debate, then*

the facts argumentation framework associated to D *is* f-$AF_D = (F,C)$, *where* $C \subseteq F \times F$ *and* $(f,g) \in C$ *if and only if* $f_D^- \cap g_D^+ \neq \emptyset$, *and*

the ***opinions argumentation framework associated to*** D *is* o-$AF_D = (S,C')$, *where* $C' \subseteq S \times S$ *and* $(s,t) \in C'$ *if and only if* $s_D^- \cap t_D^+ \neq \emptyset$.

In words: (f,g) is an *attack* in f-$AF_D = (F,C)$ if there is an individual s which approves f and disapproves g; (s,t) is an *attack* in o-$AF_D = (S,C')$ if s disapproves a fact approved by t.

Note that $(f,g) \in F \times F$ is an attack in f-$AF_D = (F,C)$ if and only if the digraph $D + (f,g)$, obtained by adding the edge (f,g) to D, contains at least one $\vec{C}_3$. Similarly, $(s,t) \in S \times S$ is an attack in o-$AF_D = (S,C')$ if and only if the digraph $D+(s,t)$, obtained by adding the edge (s,t) to D, contains at least one $\vec{C}_3$.

These constructions are exemplified in Figure 4.5.

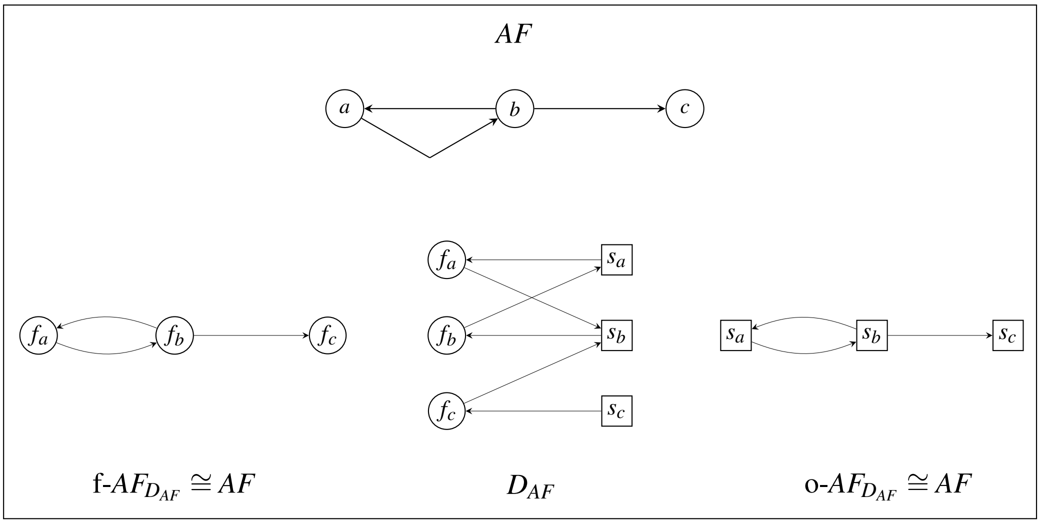


Figure 4.5.: An argumentation framework AF, its associated debate D_{AF}, with both (facts and opinions) associated argumentation frameworks isomorphic to AF.

The above isomorphisms are not incidentally as the following theorem shows.

Theorem 61 *Let AF be an argumentation framework without loops and D_{AF} the debate associated to AF. Then the facts and opinions argumentation frameworks associated to the debate D_{AF} are isomorphic to AF:* f-$AF_{D_{AF}} \cong AF \cong$ o-$AF_{D_{AF}}$.

Proof. Let $A = Arg(AF)$ and $D_{AF} = (F_A, S_A; E_{AF})$ be the debate associated to AF.

To prove that f-$AF_{D_{AF}} \cong AF$, consider bijection $\varphi : F_A \to A$ given by $\varphi(f_a) = a$ for every $a \in A$. Then, $(f_a, f_b) \in Def($f-$AF_{D_{AF}})$ if and only if there is $s \in S_A$ such that (s, f_a, f_b) is an induced $\vec{C}_3$ in $D_{AF} + (f_a, f_b)$. By the definition of the debate D_{AF} it follows that $s = s_a$, and, since (f_b, s_a) is an edge in D_{AF}, it follows that $(a,b) \in Def(AF)$. Conversely, if $(a,b) \in Def(AF)$ then, by the definition of the debate D_{AF}, we have $(s_a, f_a) \in E(D_{AF})$ and $(f_b, s_a) \in E(D_{AF})$. Hence adding (f_a, f_b) to D_{AF} we obtain an induced $\vec{C}_3$, (s_a, f_a, f_b), in $D_{AF} + (f_a, f_b)$. But this means that (f_a, f_b) is an attack in f-$AF_{D_{AF}}$.

To prove that $AF \cong$ o-$AF_{D_{AF}}$, consider bijection $\psi : S_A \to A$ given by $\psi(s_a) = a$ for every $a \in A$. Then, $(s_a, s_b) \in Def(\text{o-}AF_{D_{AF}})$ if and only if there is $f \in F_A$ such that (f, s_a, s_b) is an induced $\vec{C}_3$ in $D_{AF} + (s_a, s_b)$. By the definition of the debate D_{AF} it follows that $f = f_b$, and, since (f_b, s_a) is an edge in $D(AF)$, it follows that $(a, b) \in Def(AF)$. Conversely, if $(a, b) \in Def(AF)$ then, by the definition of the debate D_{AF}, it follows that $(s_b, f_b) \in E(D_{AF})$ and $(f_b, s_a) \in E(D_{AF})$. Hence adding (s_a, s_b) to D_{AF} we obtain an induced $\vec{C}_3$, (f_b, s_a, s_b), in $D_{AF} + (s_a, s_b)$. But this means that (s_a, s_b) is an attack in o-$AF_{D_{AF}}$. □

The opinions argumentation framework associated to a debate D, o-AF_D, can be used to consider particular sets of compatible (opinions of the) individuals.

Definition 62 *Let $D = (F, S; E) \in \mathscr{D}(\mathbb{F}, \mathbb{S})$ be a debate. A **coalition** is any subset $\mathscr{C} \subseteq S$. A coalition $\mathscr{C}$ is **legal** if*

$$O_{\mathscr{C}} = (\mathscr{C}_D^+, \mathscr{C}_D^-) \in \mathscr{O}(F).$$

$O_{\mathscr{C}}$ is the collective opinion of the coalition $\mathscr{C}$.

Note that a coalition $\mathscr{C}$ is legal if and only if $\left(\bigcup_{s\in\mathscr{C}} s_D^+\right) \cap \left(\bigcup_{s\in\mathscr{C}} s_D^-\right) = \emptyset$. Also, the empty coalition and the singletons coalitions are trivially legal coalitions.

For the *Common Lunch Dilemma* debate in Section 4.1, the following figure depicts a coalition ($\{p_1, p_4\}$) which is not legal. The conflicting directed edges entering or leaving their members are also emphasized. We can also see that $\{p_2, p_3\}$ is a legal coalition.

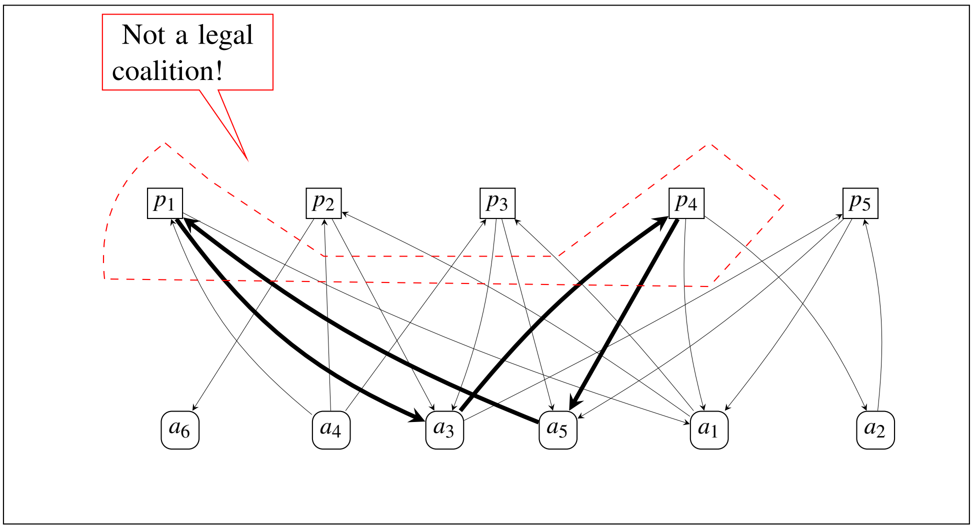


Figure 4.6.: A coalition which is not legal in the *Common Lunch Dilemma* Debate.

Proposition 63 *$\mathscr{C} \subseteq S$ is a legal coalition in the debate $D = (F,S;E)$ if and only if $\mathscr{C}$ is a conflict-free set in o-AF_D.*

Proof. Let $\mathscr{C} \subseteq S$ be a legal coalition in the debate $D = (F,S;E)$. If $\mathscr{C}$ is not a conflict-free set in o-AF_D, then there are $s,t \in \mathscr{C}$ such that $(s,t) \in Def(\text{o-}AF_D)$, that is, $s_D^- \cap t_D^+ \neq \emptyset$. But then, $\left(\bigcup_{s\in\mathscr{C}} s_D^+\right) \bigcap \left(\bigcup_{s\in\mathscr{C}} s_D^-\right) \neq \emptyset$. Hence $\mathscr{C}$ is not a legal coalition, contradicting the hypothesis.

Conversely, let $\mathscr{C} \subseteq S$ be a conflict free set in o-AF_D. Then, for every $s,t \in \mathscr{C}$, we have $s_D^- \cap t_D^+ = \emptyset$. Therefore $\left(\bigcup_{s\in\mathscr{C}} s_D^+\right) \bigcap \left(\bigcup_{s\in\mathscr{C}} s_D^-\right) = \emptyset$, that is $\mathscr{C}$ is a legal coalition in the debate $D = (F,S;E)$. □

Definition 64 *A coalition $\mathscr{C}$ is an **autarky** in the debate $D = (F,S;E)$ if it is an admissible set in the argumentation framework o-AF_D.*

Hence, if $\mathscr{C}$ is an autarky then, for every individual $p \notin \mathscr{C}$, if there is $f \in p_D^- \cap s_D^+$ for some $s \in \mathscr{C}$, then there are $f' \in F$ and $s' \in \mathscr{C}$ such that $f' \in {s'}_D^- \cap p_D^+$. Note that the empty coalition is a trivial autarky.

Definition 65 *A coalition $\mathscr{C}$ is a **strong autarky** in the debate $D = (F,S;E)$ if it is a complete extension in the argumentation framework o-AF_D. A **minimal strong autarky** (**maximal strong autarky**) is a grounded extension (preferred extension) in the argumentation framework o-AF_D.*

In words, $\mathscr{C}$ is a strong autarky in the debate $D = (F,S;E)$ if it is an autarky and, for each $p \notin \mathscr{C}$ such that p is not attacked by $\mathscr{C}$ in o-AF_D, there is $s \notin \mathscr{C}$ such that s attacks p and $\mathscr{C}$ does not attack s in o-AF_D.

Definition 66 *An **oligarchy** in the debate $D = (F,S;E)$ is a coalition $\mathscr{C}$ which is a stable extension in the argumentation framework o-AF_D.*

Clearly, any oligarchy is a maximal strong autarky.

Example. Let D be the debate represented in the Figure 4.7 below.

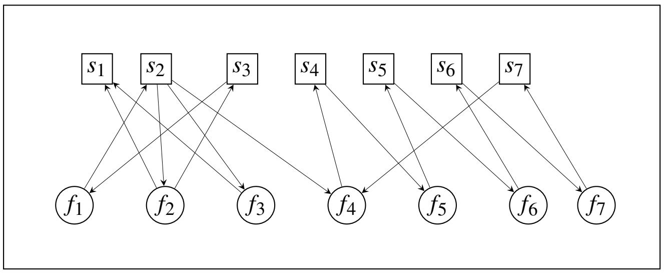


Figure 4.7.: A debate.

$\mathscr{C}_1 = \{s_5, s_7\}$ is an autarky: $O_{\mathscr{C}_1} = (\{f_4, f_6\}, \{f_5, f_7\})$, $O_{s_4} = (\{f_5\}, \{f_4\})$ attacks $O_{\mathscr{C}_1}$ but this counterattacks O_{s_4}; $O_{s_6} = (\{f_7\}, \{f_6\})$ attacks $O_{\mathscr{C}_1}$ but this counterattacks O_{s_6}; no other O_{s_i} attacks $O_{\mathscr{C}_1}$, for $i \in \{1,2,3\}$. $\mathscr{C}_1$ is also a strong autarky (since each opinion of an individual from the set $\{s_1, s_2, s_3\}$ is attacked by the opinion of an individual in the same set, and the opinions of s_4 and s_6 are attacked by $O_{\mathscr{C}_1}$), but it is not a maximal strong autarky since $\mathscr{C}_2 = \{s_1, s_3, s_5, s_7\}$ is also a strong autarky as we can easily verify. Note that $\mathscr{C}_2$ is also a stable coalition.

The above type of coalitions can be equivalently defined directly on the debate in which they are considered as follows.

Definition 67 *Let $D = (F, S; E) \in \mathscr{D}(\mathbb{F}, \mathbb{S})$ be a debate. An individual $s \in S$ is called a **strong eristic** if $s_D^- \cap t_D^+ \neq \emptyset$ for every $t \in S - \{s\}$. An individual $s \in S$ is called a **weak eristic** if $s_D^- \cap t_D^+ \neq \emptyset$ for every $t \in S - \{s\}$ such that $t_D^- \cap s_D^+ \neq \emptyset$.*

In words, an individual s is a strong eristic if it has a negative position on at least one of the facts agreed by every other individual t; s is a weak eristic if it has a negative position on at least one of the facts agreed by every other individual t which has a negative position on a fact agreed by s.

We can verify that p_4 is a strong eristic in the subdebate of the in Figure 4.1 obtained by deleting p_5; also s_b is a strong eristic in the debate $D(AF)$ in Figure 4.5; in this last debate, s_a is a weak eristic.

Let $\mathscr{C} \subseteq S$ be a non-empty coalition in the debate $D = (F, S; E)$. The digraph obtained from D by contracting $\mathscr{C}$ is $D|\mathscr{C} = (F, (S - \mathscr{C}) \cup \{s_{\mathscr{C}}\}; E')$, where $s_{\mathscr{C}}$ is a new "individual" ($s_{\mathscr{C}} \in \mathbb{S} - S$) and $(f, s) \in E'$ ($(s, f) \in E'$) if and only if $(f, s) \in E$ ($(s, f) \in E$) and $s \notin \mathscr{C}$ or $s = s_{\mathscr{C}}$ and there is $t \in \mathscr{C}$ such that $(f, t) \in E$ ($(t, f) \in E$).

Note that $(s_{\mathscr{C}})^+_{D|\mathscr{C}} = \cup_{s \in \mathscr{C}} s_D^+$ and $(s_{\mathscr{C}})^-_{D|\mathscr{C}} = \cup_{s \in \mathscr{C}} s_D^-$. It follows that $D|\mathscr{C}$ has no symmetric pairs of edges if and only if $\mathscr{C}$ is a legal coalition. Hence the following lemma holds.

Lemma 68 *A non-empty coalition $\mathscr{C} \subseteq S$ in the debate $D = (F, S; E)$ is a legal coalition if and only if $D|\mathscr{C}$ is a debate.*

Using this lemma and Definitions 64, 65, and 66, we have the following theorem.

Theorem 69 *Let $\mathscr{C} \subseteq S$ be a non-empty coalition in the debate $D = (F, S; E)$ such that $D|\mathscr{C}$ is a debate. Then,*

i) $\mathscr{C}$ is an autarky if and only if $s_{\mathscr{C}}$ is a weak eristic in $D|\mathscr{C}$;

ii) $\mathscr{C}$ is a strong autarky if and only if $s_{\mathscr{C}}$ is a weak eristic in $D|\mathscr{C}$ and for every $t \in S - \mathscr{C}$ such that $t^-_{D|\mathscr{C}} \cap (s_{\mathscr{C}})^+_{D|\mathscr{C}} = \emptyset$ there is $u \in S - \mathscr{C}$, $u \neq t$ such that $u^-_{D|\mathscr{C}} \cap t^+_{D|\mathscr{C}} \neq \emptyset$ and $u^+_{D|\mathscr{C}} \cap (s_{\mathscr{C}})^-_{D|\mathscr{C}} = \emptyset$;

iii) $\mathscr{C}$ is an oligarchy if and only if $s_{\mathscr{C}}$ is a strong eristic in $D|\mathscr{C}$.

Proof.

(i) If $\mathscr{C}$ is an autarky then it is an admissible set in the argumentation framework o-AF_D. Suppose that $s_{\mathscr{C}}$ is not a weak eristic in $D|\mathscr{C}$. Then, there is $t \in S - \mathscr{C}$ such that $t^-_{D|\mathscr{C}} \cap (s_{\mathscr{C}})^+_{D|\mathscr{C}} \neq \emptyset$ and $(s_{\mathscr{C}})^-_{D|\mathscr{C}} \cap t^+_{D|\mathscr{C}} = \emptyset$. Hence there is $p \in \mathscr{C}$ such that $t^-_D \cap p^+_D \neq \emptyset$. This means that, in o-$AF_D$, t attacks $p \in \mathscr{C}$. Because $(s_{\mathscr{C}})^-_D \cap t^+_D = \emptyset$, it follows that, in o-$AF_D$, $\mathscr{C}$ does not defend the attack on $p \in \mathscr{C}$. This contradicts the hypothesis that $\mathscr{C}$ is an autarky in D.

Conversely, suppose that $s_{\mathscr{C}}$ is a weak eristic in $D|\mathscr{C}$. If $\mathscr{C}$ is not an autarky in D, then it is not an admissible set in o-AF_D. This means that there is $t \in S - \mathscr{C}$ such that t attacks some $p \in \mathscr{C}$ in o-AF_D and $\mathscr{C}$ does not counter-attack p. But then, in $D|\mathscr{C}$, p attacks $s_{\mathscr{C}}$ and this does not counter-attack p, contradicting the hypothesis that $s_{\mathscr{C}}$ is a weak eristic.

(ii) By (i), $\mathscr{C}$ is a strong autarky if and only if $s_{\mathscr{C}}$ is a weak eristic in $D|\mathscr{C}$ (that is, $\mathscr{C}$ is an admissible set in o-AF_D) and for each $t \notin \mathscr{C}$ such that t is not attacked by $\mathscr{C}$ in o-AF_D, there is $u \notin \mathscr{C}$ such that u attacks t and $\mathscr{C}$ does not attack u in o-AF_D. This additional property can be stated in $D|\mathscr{C}$ as enounced in (ii): for every $t \in S - \mathscr{C}$ such that $t^-_{D|\mathscr{C}} \cap (s_{\mathscr{C}})^+_{D|\mathscr{C}} = \emptyset$ there is $u \in S - \mathscr{C}$, $u \neq t$ such that $u^-_{D|\mathscr{C}} \cap t^+_{D|\mathscr{C}} \neq \emptyset$ and $u^+_{D|\mathscr{C}} \cap (s_{\mathscr{C}})^-_{D|\mathscr{C}} = \emptyset$.

(iii) $\mathscr{C}$ is an oligarchy in D if and only if $\mathscr{C}$ is a stable extension in o-AF_D, and since $\mathscr{C}$ is legal, this means that $s_{\mathscr{C}}$ is a strong eristic in $D|\mathscr{C}$.

$\square$

By Theorem 61, for the debate D_{AF} associated to an argumentation framework AF, the above different type of coalitions translate to the corresponding admissible based extensions in AF.

It follows that the decision problems on argumentation frameworks can be polynomially transformed (in fact, in linear time) into instances of the corresponding problems on debates. For example, let us consider the following two such problems (see Subsection 0.3).

CA$_{\mathbf{pref}}$ (Credulous Preferred Acceptance)
Instance : $AF = (A, D)$ and $a \in A$.
Question : Is there $S \in \mathbf{pref}(AF)$ such that $a \in S$?

SA$_{\mathbf{pref}}$ (Skeptical Preferred Acceptance)
Instance : $AF = (A, D)$ and $a \in A$.
Question : Is a a member of each $S \in \mathbf{pref}(AF)$?

The corresponding problems for abstract debates are stated below.

$\mathbf{CA_{msa}}$ (Credulous Maximal Strong Autarky)
Instance : $D = (F,S;E)$ a debate, and a fact $f \in F$.
Question : Is there $\mathscr{C} \subseteq S$ a maximal strong autarky in D such that $f \in \mathscr{O}^+_{\mathscr{C}}$?

$\mathbf{SA_{msa}}$ (Skeptical Maximal Strong Autarky)
Instance : $D = (F,S;E)$ a debate, and a fact $f \in F$.
Question : Is f a member of $\mathscr{O}^+_{\mathscr{C}}$, for each maximal strong autarky $\mathscr{C}$ in D?

Using the time complexity results on the first two decision problems (Dunne and Bench-Capon [DBC02]), we obtain the following corollary of the Theorem 61.

Corollary 70 $\mathbf{CA_{msa}}$ *is an* NP*-complete problem and* $\mathbf{SA_{msa}}$ *is a* Π_2^P*-complete problem.*

Coalitions are very restrictive in some debates. For example, if $D = (F,S;E)$ is a debate with the property that $s^+_D \bigcup s^-_D = F$ for each $s \in S$, then the only legal non-empty coalitions are trivial: singletons and sets of individuals having the same opinion. Indeed, if s and t are distinct individuals with different opinions, then (s,t) or (t,s) is an attack in o-AF_D. Clearly, if $(s^+_D, s^-_D) \neq (t^+_D, t^-_D)$, then since these opinions are distinct partitions of F, there is $f \in t^+_D \cap s^-_D$ (hence s attacks t in o-AF_D) or there is $f \in s^+_D \cap t^-_D$ (hence t attacks s in o-AF_D). If the two individuals strategically desire to be part of a coalition, it is necessary that s renounces at its liked facts in $s^+_D \cap t^-_D$ and t renounces at its liked facts in $t^+_D \cap s^-_D$. This tolerant way of coalition formation, inspired by political practice, is captured in the following definitions.

Definition 71 *Let $D = (F,S;E) \in \mathscr{D}(\mathbb{F},\mathbb{S})$ be a debate and $\mathscr{C} \subseteq S$ a coalition in D. A $\mathscr{C}$-**compromise** is the pair $(\mathscr{C},E')$, where $E' \subset E$ is a set of edges (s,f) with $s \in \mathscr{C}$ such that in the debate $D' = D - E'$ (obtained from D by removing the edges from E') $s^+_{D'} \neq \emptyset$ and $\mathscr{C}$ is a legal coalition.*

In words, some members of a coalition renounce at some liked (approved) facts in order to make the coalition legal. In this way the coalition has the same negative part of its merged opinion (which is used to ”attack” opinions of individuals not in coalition) despite of weakening the positive part.

Definition 72 *A $\mathscr{C}$-compromise $(\mathscr{C},E')$ in the debate $D = (F,S;E)$ is called a **σ-compromise** if the coalition $\mathscr{C}$ is σ in the debate $D' = D - E'$, for $\sigma \in \{$ autarky, strong autarky, minimal strong autarky, maximal strong autarky, oligarchy$\}$. If $(\mathscr{C},E')$ is a σ-compromise in $D = (F,S;E)$, then $D_{(\mathscr{C},E')}$ is the consensual debate in which each individual opinion is the union of opinions of the members of $\mathscr{C}$ in $D - E'$.*

Note that in debates D with $s_D^+ \neq \emptyset, \forall s \in S$, if the coalition $\mathscr{C}$ is σ then $(\mathscr{C}, \emptyset)$ is also a σ-compromise. Also, in the debates D with $|s_D^+| = 1, \forall s \in S$, $(\mathscr{C}, E')$ is a σ-compromise if and only if $E' = \emptyset$ (otherwise $s_{D'}^+ = \emptyset$, for some $s \in \mathscr{C}$), that is if and only if $\mathscr{C}$ is σ; such debates are called *single minded* and hence in single minded debates there is no proper compromises.

We define now our argumentative aggregation correspondences.

Definition 73 *An argumentative aggregation correspondence is a function* $\mathbb{AC}_\sigma$*, which maps every debate* $D \in \mathscr{D}(\mathbb{F}, \mathbb{S})$ *into the following set of consensual debates*

$$\mathbb{AC}_\sigma(D) = \{D_{(\mathscr{C}, E')} | (\mathscr{C}, E') \textit{ is a } \sigma\textit{-compromise}\},$$

for $\sigma \in \{$ *autarky, strong autarky, minimal strong autarky, maximal strong autarky, oligarchy*$\}$.

Example. Let us consider again the debate in Figure 4.6 and $\sigma =$ oligarchy. We observed that $\mathscr{C} = \{p_1, p_4\}$ is not legal. But, $(\mathscr{C}, E')$ with $E' = \{(p_1, a_3), (p_4, a_5)\}$ is an oligarchy-compromise with $F^+_{D_{(\mathscr{C},E')}} = \{a_1, a_2\}$ and $F^-_{D_{(\mathscr{C},E')}} = \{a_5, a_4, a_3\}$, as suggested in figure below.

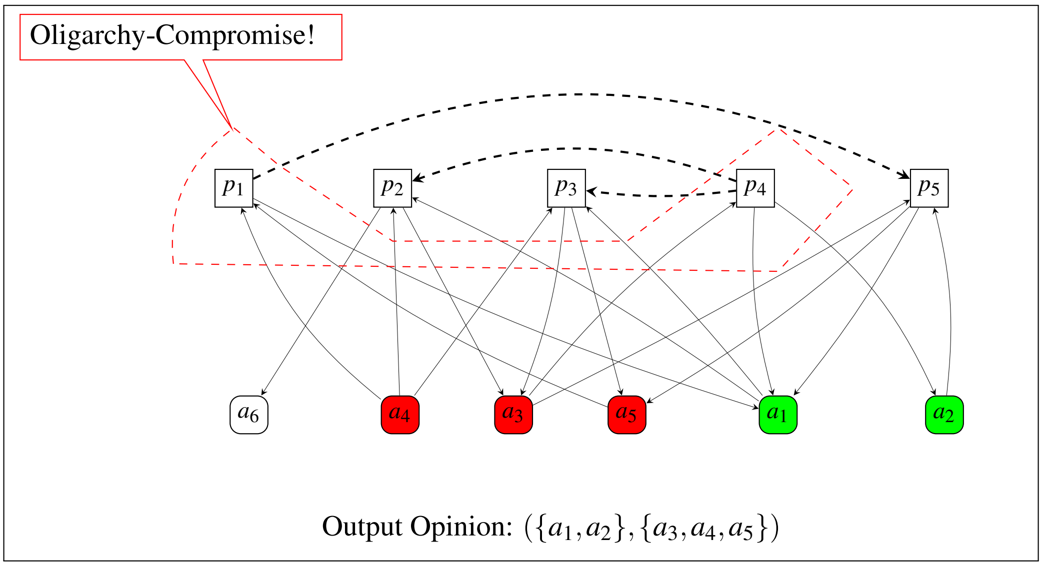


Figure 4.8.: *Common Lunch Dilemma* Debate: Oligarchy-Compromise.

The output opinion, $(\{a_1, a_2\}, \{a_3, a_4, a_5\})$, has the property that each person involved in the *Common Lunch* can eat a_1 or a_2.

Note that if we choose as output the merged opinion of an oligarchy (or oligarchy-compromise), each individual outside this coalition likes a fact which is disliked by some member of the oligarchy. Similar (argumentative) explanations can be made for other choices of σ, depending on the application for which the aggregation is considered.

Each argumentative aggregation correspondence $\mathbb{AC}_\sigma$ gives rise to an argumentative aggregator $\mathbb{A}_\sigma$, by specifying a rule to select one of the consensual debates in $\mathbb{AC}_\sigma(D)$. $\mathbb{A}_\sigma$ satisfies the second condition in Definition 52, by Theorem 61 and the invariance of the admissibility based semantics to the AFs isomorphism.

Since each consensual debate $D_{\mathscr{C};E'} \in \mathbb{AC}_\sigma(D)$ is determined by the collective opinion of a σ-compromise $(\mathscr{C}, E')$, it follows that each liked fact $f \in O^+_{D_{\mathscr{C};E'}}$ has a non-empty set of supporters $S(f) \subseteq \mathscr{C}$. If we define the *support* of $D_{\mathscr{C};E'}$ as $\mathrm{supp}(D_{\mathscr{C};E'}) = \sum_{f \in O^+_{\mathscr{C};E'}} |S(f)|$, then we can keep in $\mathbb{AC}_\sigma(D)$ only the maximum support consensual debates. Other strategy of reducing the set $\mathbb{AC}_\sigma(D)$ is to retain only the debates with the (common) opinion at minimum distance to the entire set of individuals opinion (after defining rational distance functions).

Since the operators $\mathbb{A}_\sigma$ are not "fact wise" and are strongly dependent on the context of the debate to which they are applied, we have the following theorem.

Theorem 74 *Argumentative aggregation operators $\mathbb{A}_\sigma$ do not satisfy independence.*

Proof. Remember that an aggregator $\mathbb{A}$ satisfies independence if, for every debate $D = (F, S; E)$ and for every $f \in F$, we have $\mathbb{A}(D)^{\{f\}} = \mathbb{A}(D^{\{f\}})$.

Consider the debates D and $D^{\{f\}}$ in Figure 4.9 below.

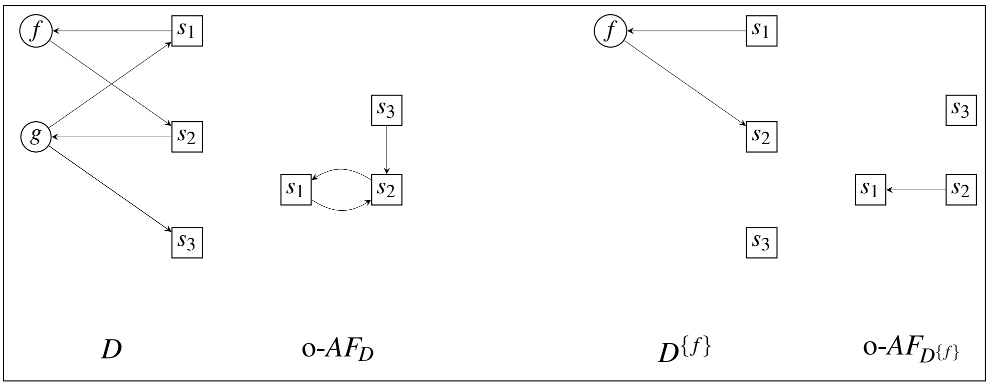


Figure 4.9.: Debates D and $D^{\{f\}}$ and their associated opinion AFs.

Since $|(s_i)^+_D| \leq 1$, a coalition $\mathscr{C}$ in these debates is σ-compromise if and only if it is σ. The only autarky in D is $\{s_1, s_3\}$ hence each individual opinion in $\mathbb{A}_\sigma(D)$ is $(\{f\}, \{g\})$. The only autarky in $D^{\{f\}}$ is $\{s_2, s_3\}$ hence each individual opinion in $\mathbb{A}_\sigma(D^{\{f\}})$ is $(\emptyset, \{f\})$. Therefore $\mathbb{A}_\sigma(D)^{\{f\}} \neq \mathbb{A}(D^{\{f\}})$. $\square$

5. Bipartite Choice Systems

A functional framework abstracting many-to-many two-sided markets is introduced. Sufficient conditions for the existence of stable many-to-many matchings are obtained as particular instances of determining stable common fixed points of two choice functions, by using a generalization of the deferred acceptance Gale-Shapley algorithm, called *Immediate Rejection.* A systematic study of choice functions conditions related to this subject is done, obtaining a new characterization of path independence choice functions. This can be viewed as a formal explanation for the role of the substitutability and independence of irrelevant alternatives in the many-to-many matchings applications.

5.1. Introduction

The seminal paper of Gale and Shapley [GS62] introduced the "two-sided matching model" which turned out to give rise to an entire research area in two-sided (labour) market design. In their original model, the agents are placed in the vertices of the complete bipartite graph $K_{n,n}$, each agent can be matched only with an agent from the other side and it is assumed that each agent has a strict ordering on the set of all its neighbors. A matching is stable if it left no pair of agents on opposite sides of the market who were not matched to each other but would both prefer to be. Gale and Shapley proved that stable matchings always exist by giving a deferred acceptance algorithm to construct it.

In the *Economics* literature, broad generalizations of this model were considered by allowing centralized matching schemes where there could be multiple partners on both sides (many-to-many) of the market (Roth [Rot84], Blair [Bla88], Hatfield and Milgrom [HM05]). Moreover, agent's preferences were given by choice functions on the set of its neighbors that do not necessarily respect an ordering of individuals. More precisely, the choice function of each agent specifies for any set of its neighbors a subset of (most preferred) individuals. New stability concepts for these multi-partner matchings are defined and, in order to obtain a nonempty set of stable matchings, appropriate choice's rule conditions from the *Arrovian Social Choice* literature are considered: *substitutability* and *independence of irrelevant alternatives* (Roth [Rot84], Blair [Bla88], Roth and Sotomayor [RS90]).

Substitutability was introduced, and applied to matching markets, by Kelso and Crawford [KC82]. The importance of independence of irrelevant alternatives condition for the existence of a stable output in these models (Hatfield and Milgrom [HM05]) has been emphasized by Aygün and Sönmez [AS12]. Note that Brandt and Harrenstein [BH11] showed that this condition is equivalent to a very natural *set-rationability* condition of the agent's choice functions.

In this chapter, we introduce a simple framework **Bipartite Choice System (BCS)**, as an abstract functional model for the above many-to-many matchings applications. A BCS is a collection of choice functions indexed by the vertices of a bipartite (multi)graph, such that, for any set A of edges of the bipartite (multi)graph, the choice function in each vertex v selects the most "preferred" subset of edges in A incident to v. Since each edge is incident to exactly two vertices from different sides of the bipartite multigraph, a set of edges chosen by the two sides can be viewed as a **choice matching**.

In many-to-many two-sided markets, the edges are referred as **contracts** and the choice functions C_v are defined on the set of neighbors of the vertex v. We prefer to define the choice functions on the set of edges incident to each vertex, which simplifies the notations and highlights the agreement intuition behind the problems considered. **Stable choice matchings** can be formally defined along the lines introduced by von Neumann and Morgenstern [NM44] as described by Brandt and Harrenstein [BH11].

The existence of stable choice matching in a given BCS is studied by associating to each BCS its **unilateral choice functions** (which are the direct sum of the choice functions of the vertices in each part) and interpreting stable choice matchings as **stable common fixed points** of these two functions. We consider an appropriate generalization of the deferred acceptance Gale-Shapley algorithm – **Immediate Rejection** – and, imposing conditions on the two choice functions, we prove a sufficient condition for the correctness of the algorithm. By relating this condition to well-known existing sufficient conditions in the many-to-many matchings applications, we can view our contribution as an explanation for the role of the substitutability and independence of irrelevant alternatives.

5.2. Stable Choice Matchings

A **choice function** on a finite set U (of alternatives) is a function $f : 2^U \to 2^U$ such that $f(A) \subseteq A$, $\forall A \subseteq U$. As the name suggests, a choice function can be interpreted as a representation of the agents' (strict) preferences over sets of alternatives in U: $f(A)$ is the most preferred subset of A, the alternatives in $f(A)$ are chosen (selected) and those in $A - f(A)$ are rejected.

A **Bipartite Choice System (BCS)** is a collection of choice functions indexed by the vertices of a bipartite (multi)graph such that, for any set A of edges of the bipartite (multi)graph, the choice function in each vertex v selects the most "preferred" subset of edges in A incident to v. The following definition makes clear the notations and terminology.

Definition 75 *A **Bipartite Choice System (BCS)** is a a couple $(G,\mathscr{C})$, where*

- *$G = (S\cup T;E)$ is a bipartite (multi)graph: the set of vertices of G is the union of two disjoint finite non-empty sets S and T and each edge $e \in E$ has associated a 2-set $\{s(e),t(e)\}$, with $s(e) \in S$ and $t(e) \in T$; we say that e is incident to $s(e)$ and $t(e)$. If $A \subseteq E$ and $v \in S\cup T$, then A_v denotes the set of edges in A incident to v.*
- *$\mathscr{C} = \{C_v | v \in S\cup T\}$ specifies, for each vertex v, a choice function $C_v : 2^{E_v} \to 2^{E_v}$.*

Note that if v and v' are two distinct vertices in the same part (S or T) of the bipartite graph then $E_v \cap E_{v'} = \emptyset$.

For a **BCS** $(G = (S,T;E),\mathscr{C})$ we can consider the following **collective** choice functions:

unilateral choice functions $C_S, C_T : 2^E \to 2^E$, defined by

$$C_S(A) = \bigcup_{s\in S} C_s(A_s),\ C_T(A) = \bigcup_{t\in T} C_t(A_t),\ \forall A \subseteq E;$$

bilateral choice function $C : 2^E \to 2^E$, defined by

$$C(A) = C_S(A) \bigcap C_T(A),\ \forall A \subseteq E.$$

In words, the unilateral choice functions $C_S(A)$, $C_T(A)$ select from a set $A \subseteq E$ those edges selected by at least one vertex in a specified part (S or T) of the bipartite graph; the bilateral choice $C(A)$ of a set $A \subseteq E$ is the set of edges in A selected by both parts of the bipartite graph: $e \in A$ belongs to $C(A)$ if and only if $e \in C_{s(e)}(A_{s(e)})$ and $e \in C_{t(e)}(A_{t(e)})$.

Definition 76 *A **choice matching** in the **BCS** $(G = (S,T;E),\mathscr{C})$ is any set $M \subseteq E$ of edges such that*

$$C(M) = M.$$

If the choices functions satisfy $|C_v(A_v)| \leq 1$ for each $A \subseteq E$ and $v \in S\cup T$, then a choice matching is an usual matching in G: since $C(M) = M$ if and only if $\forall v \in S\cup T$, $C_v(M_v) = M_v$, we have $|M_v| = |C_v(M_v)| \leq 1$, for all $v \in V$, that is M is a set of non-adjacent edges in G. Also, if $|C_v(A_v)| \leq 1$ for each $A \subseteq E$ and $v \in S\cup T$, then if M is a choice matching then M' is a choice matching for every $M' \subseteq M$. This property is not fulfilled in general for arbitrary choice functions C_v.

Definition 77 *$M \subseteq E$ is a **stable choice matching** in the **BCS** $(G = (S,T;E),\mathscr{C})$ if*

• $C(M) = M$ ***(internal stability)*** *and*

• $e \notin C(M \cup \{e\}), \forall e \in E - M$ ***(external stability)***.

Let $e = \{s,t\} \in E - M$ be the edge in the external stability condition of M. Since each edge in $C_s(M_s \cup \{e\})$ is incident to s, and each edge in $C_t(M_t \cup \{e\})$ is incident to t, then $e \notin C(M \cup \{e\})$ if and only if $e \notin C_s(M_s \cup \{e\}) \cap C_t(M_t \cup \{e\})$. Hence, if M is a **stable choice matching** and $e \in E - M$, then the external stability condition requires that at least one of the vertices $s(e)$ and $t(e)$ does not choose e if this is added to the set of edges in the choice matching M. It follows that, for a **BCS** $(G = (S,T;E), \mathscr{C})$, were the choice functions satisfy $|C_v(A_v)| \leq 1$ for each $A \subseteq E$ and $v \in S \cup T$, a stable choice matching is a usual stable matching.

Note that the above two conditions in the definition of a **stable choice matching** can be equivalently (see Brandt and Harrenstein [BH11]) stated as

$$M = \{e \in E | e \in C(M \cup \{e\})\}.$$

Examples. Let $G = (S,T;E)$ be the bipartite multigraph in the figure below, where $S = \{s_1, s_2\}$, $T = \{t\}$, $E = \{1,2,3\}$, $E_{s_1} = \{1,2\}$, $E_{s_2} = \{3\}$, and $E_t = \{1,2,3\}$.

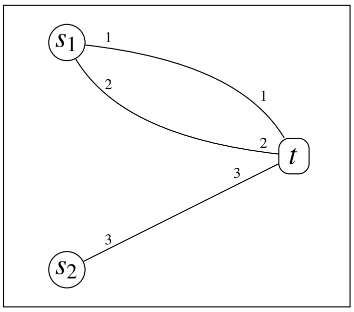


Figure 5.1.: A bipartite multigraph.

1. Let the BCS be $(G, \mathscr{C}^1)$, where

$A(\subseteq E_{s_1})$	$\emptyset$	$\{1\}$	$\{2\}$	$\{1,2\}$
$C^1_{s_1}(A)$	$\emptyset$	$\{1\}$	$\{2\}$	$\{1\}$

$A(\subseteq E_{s_2})$	$\emptyset$	$\{3\}$
$C^1_{s_2}(A)$	$\emptyset$	$\{3\}$

$A(\subseteq E_t)$	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$C^1_t(A)$	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1\}$	$\{2\}$	$\{1\}$

.

The choice matchings in this BCS are $M_1 = \emptyset$, $M_2 = \{1\}$, $M_3 = \{2\}$, $M_4 = \{3\}$. Indeed, it is simply to see that $C^1(M_i) = M_i$, for $i \in \{1,\dots,4\}$); if $A \subseteq \{1,2,3\}$ is such that $|A| \geq 2$, then $|C^1(A)| < 2$ and therefore $C^1(A) \neq A$.
$M_1 = \emptyset$ is not a stable choice matching since $x \in C^1(M_1 \cup \{x\})$, for $x \in \{1,2,3\} - M_1$. $M_3 = \{2\}$ and $M_4 = \{3\}$ are not stable choice matchings: $1 \notin M_3$ but $1 \in C^1(M_3 \cup \{1\})$; $2 \notin M_4$ but $2 \in C^1(M_4 \cup \{2\}$.
$M_2 = \{1\}$ is a stable choice matching: $2 \notin C^1(M_2 \cup \{2\})$, $3 \notin C^1(M_2 \cup \{3\})$.

Hence in the BCS $(G, \mathscr{C}^1)$ **there is exactly one stable choice matching**.

2. Let the BCS be $(G, \mathscr{C}^2)$, where

$A(\subseteq E_{s_1})$	$\emptyset$	$\{1\}$	$\{2\}$	$\{1,2\}$
$C^2_{s_1}(A)$	$\emptyset$	$\{1\}$	$\{2\}$	$\{1\}$

$A(\subseteq E_{s_2})$	$\emptyset$	$\{3\}$
$C^2_{s_2}(A)$	$\emptyset$	$\{3\}$

$A(\subseteq E_t)$	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$C^2_t(A)$	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1\}$	$\{3\}$	$\{2\}$	$\emptyset$

.

The choice matchings in this BCS are $M_1 = \emptyset$, $M_2 = \{1\}$, $M_3 = \{2\}$, $M_4 = \{3\}$. Indeed, since $|C^2_t(A)| \leq 1$ there is no choice matching with at least 2 edges, and all M_i above satisfy $C^2(M_i) = M_i$.
No one is a stable choice matching: $3 \in C^2(M_1 \cup \{3\})$, $3 \in C^2(M_2 \cup \{3\})$, $1 \in C^2(M_3 \cup \{1\})$, $2 \in C^2(M_4 \cup \{2\})$.

Hence in the BCS $(G, \mathscr{C}^2)$ **there is no stable choice matching**.

3. Let the BCS be $(G, \mathscr{C}^3)$, obtained from $(G, \mathscr{C}^1)$ by taking the choice functions $C^3_{s_2} = C^1_{s_2}$, $C^3_t = C^1_t$, and $C^3_{s_1}$ given by

$A(\subseteq E_{s_1})$	$\emptyset$	$\{1\}$	$\{2\}$	$\{1,2\}$
$C^3_{s_1}(A)$	$\emptyset$	$\{1\}$	$\{2\}$	$\{1,2\}$

.

As above, we can check that the choice matchings in the BCS $(G, \mathscr{C}^3)$ are $M_1 = \emptyset$, $M_2 = \{1\}$, $M_3 = \{2\}$, $M_4 = \{3\}$, and $M_5 = \{1,2\}$. From these, the only stable choice matching is M_5. The difference between this example and Example 1 will be discussed in the Section 5.4.

If $(G = (S,T;E), \mathscr{C})$ is a BCS, then we can replace the part S by a super-node v_S, the part T by a super-node v_T and for each edge $e \in E$, we replace $s(e)$ by v_S and $t(e)$ by v_T. The contracted bipartite graph obtained, $G' = (S', T'; E) = (\{v_S\}, \{v_T\}; E)$, has only two nodes connected by the set E of multi-edges. If we consider $\mathscr{C}' = \{C_{v_S}, C_{v_T}\}$, where C_{v_S}, C_{v_T} are the choice functions on E defined by $C_{v_S} = C_S$ and

$C_{v_T} = C_T$, then the contracted BCS $(G', \mathscr{C}')$ has the property that $M \subseteq E$ is a choice matching in $(G = (S,T;E), \mathscr{C})$ if and only if it is a choice matching in $(G' = (S',T';E), \mathscr{C}')$. Also, $M \subseteq E$ is a stable choice matching in $(G = (S,T;E), \mathscr{C})$ if and only if it is a stable choice matching in $(G' = (S',T';E), \mathscr{C}')$.

It follows that we can approach the existence of a stable choice matching in a BCS as the existence of a particular common fixed point of two choice functions (C_S, and C_T) on the edge set of the BCS as described below.

Definition 78 *Let f be a choice function on U.*

- *A **fixed point** of f is any set $P \subseteq U$ such that $f(P) = P$.*
- *A **stable fixed point** of f is is any set $P \subseteq A$ such that*

$$P = \{x \in A \,|\, x \in f(P \cup \{x\})\ \}.$$

A choice function has at least one fixed point, namely the empty set.
If P is a stable fixed point and $x \in P$, then $P \cup \{x\} = P$, and by definition $x \in f(P \cup \{x\}) = f(P)$. It follows $P \subseteq f(P)$, that is (since f is a choice function) P is a fixed point of f. If $x \in U - P$, then $x \notin f(P \cup \{x\})$. It follows that an equivalent definition of a stable fixed point is

$$\begin{cases} P \text{ is a fixed point of } f \quad \textbf{(internal stability),} \\ \text{if } x \in U - P \text{ then } x \notin f(P \cup \{x\}) \quad \textbf{(external stability).} \end{cases}$$

Definition 79 *Let $f, g : 2^U \to 2^U$ be choice functions on U. A **stable common fixed point** of f and g is a stable fixed point of the choice function $f \wedge g : 2^U \to 2^U$ given by $f \wedge g\,(X) = f(X) \cap g(X)$, for all $X \subseteq U$.*

Note that if P is a stable common fixed point of $f \wedge g$, then (by the internal stability condition) we have $f(P) = g(P) = P$ (that is, P is a common fixed point), and (by the external stability condition) $x \notin f(P \cup \{x\})$ or $x \notin g(P \cup \{x\})$, for all $x \in U - P$.

Clearly, a stable choice matching in a BCS $(G = (S,T;E), \mathscr{C})$ is a stable common fixed point of the unilateral choice function C_S and C_T on E. This motivates the study of the existence of stable common fixed points of two choice functions on the same universe U.

5.3. Choice Functions Conditions

Throughout this section, $f : 2^U \to 2^U$ is a fixed choice function. We are discussing conditions to be fulfilled by two choice functions to guarantee the existence of a stable common fixed point. Starting from well known conditions in the matching research area, we consider new conditions motivated by the proof of correctness of the algorithm presented in Section 5.4.

5.3.1. Substitutability

An essential condition for the existence of a stable matching is **Substitutability (SUB)**. It was introduced to the matching literature by Kelso and Crawford [KC82], emphasized by Roth [Rot84], Blair [Bla88], Roth and Sotomayor [RS90] and expressed by Hatfield and Milgrom [HM05] in a form already known in the Arrovian social choice literature as Chernoff (see Moulin [Mou85]) or Sen's condition α ([Sen71]):

$$\textbf{SUB} \quad \text{if } X \subseteq Y \text{ then } f(Y) \cap X \subseteq f(X), \text{ for all } X, Y \subseteq U.$$

Hatfield and Milgrom [HM05] observed that **SUB** is equivalent to the following **Monotony of rejection (Mon)** condition:

$$\textbf{Mon} \quad \text{if } X \subseteq Y \text{ then } X - f(X) \subseteq Y - f(Y), \forall X, Y \subseteq U.$$

The next proposition shows some properties of the fixed points of substitutable choice functions.

Proposition 80 *Let $f : 2^U \to 2^U$ be a choice function that satisfies **SUB**. Then*

i) *If P is a fixed point of f then P' is a fixed point of f, for all $P' \subseteq P$.*

ii) *For every $A \subseteq U$ we have*

$$\{P \,|\, P \subseteq f(A) \text{ and } P \text{ is a fixed point of } f\} = 2^{f(A)}.$$

iii) *Let P be a fixed point of f. If there is $A \subseteq U - P$ such that $A \subseteq f(P \cup A)$ then $a \in f(P \cup \{a\})$ for every $a \in A$. Furthermore A is a fixed point of f.*

Proof. i) Taking $X := P'$ and $Y := P$ in the **SUB** condition we obtain $P' = Y \cap X = f(Y) \cap X \subseteq f(X) = f(P')$.

ii) By i) it suffices to prove that **SUB** implies **idempotency** $f(f(A)) = f(A)$, for every $A \subseteq U$. But this is immediate: taking $X := f(A)$ and $Y := A$ in the **SUB** condition we obtain $f(A) \cap A \subseteq f(f(A))$, that is, $f(A) \subseteq f(f(A))$.

iii) Taking $X := P \cup \{a\}$ and $Y := P \cup A$ in the **SUB** condition, we obtain $f(P \cup A) \cap (P \cup \{a\}) \subseteq f(P \cup \{a\})$. Since $A \subseteq f(P \cup A)$ and $a \in A$, we have $a \in$

$f(P \cup \{a\})$. Moreover, taking $X := A$ and $Y := P \cup A$ in the **SUB** condition, we obtain $f(P \cup A) \cap A \subseteq f(A)$. Since $A \subseteq f(P \cup A)$, we have $A \subseteq f(A)$. □

The property ii) justifies the term **internal stability** in the definition of a stable fixed point for a choice function satisfying **SUB**.

Note that iii) above shows that, in presence of substitutability, a stronger "set-wise" **external stability** (there is no $A \subseteq U - P$, $A \neq \emptyset$, such that $A \subseteq f(P \cup A)$) is equivalently to the "pair-wise" one given in the definition of a stable fixed point.

We close this subsection by presenting another equivalent form of **SUB**, **Conservative Choice (CC)**, which will be used in the next subsection.

$$\textbf{CC} \quad f(A) - B \subseteq f(A - B), \text{ for all } A, B \subseteq U.$$

In words, if the alternatives from the set B become unavailable, then the remaining alternatives chosen by f from A are from those selected by f from $A - B$.

Lemma 81 *A choice function f satisfies **CC** if and only if it satisfies **SUB**.*

Proof. **CC** $\Rightarrow$ **SUB**. Let $X \subseteq Y$. We take in **CC** $A := Y$ and $B := Y - X$. Then $f(A) - B = f(Y) - (Y - X) = f(Y) \cap X$ and $f(A - B) = f(Y - (Y - X)) = f(X)$. By **CC**, we have $f(Y) \cap X = f(A) - B \subseteq f(A - B) = f(X)$.

SUB $\Rightarrow$ **CC**. We take in **SUB** $X := A - B$ and $Y := A$. Then $X \subseteq Y$, $f(Y) \cap X = f(A) \cap (A - B) = f(A) - B$, and $f(X) = f(A - B)$. By **SUB**, $f(A) - B = f(Y) \cap X \subseteq f(X) = f(A - B)$. □

Note that the condition **CC** with set B of cardinality 1 was used by Blair [Bla88]. It is not difficult to prove by induction on the cardinality of B that Blair's weaker condition implies **CC**.

An interesting application of this lemma is given by the following proposition.

Proposition 82 *Let f be a choice function satisfying **SUB** and $X, Y \in 2^U$ such that $f(X) \subseteq Y \subseteq X$. Then $f(X) \subseteq f(Y)$.*

Proof. By Lemma 81 f satisfies **CC**. Let us take $A := X$ and $B := X - Y$ in **CC**. Then $f(A) - B = f(X) - (X - Y) = f(X)$ and $A - B = X - (X - Y) = Y$. By **CC**, we have $f(X) \subseteq f(Y)$.

□

5.3.2. Conservative Rejection

We consider a new condition, called **Conservative Rejection (CR)**, which is stronger than substitutability by making the hypothesis in **SUB** less strict:

$$\textbf{CR} \quad \text{If } f(X) \subseteq Y \text{ then } f(Y) \cap X \subseteq f(X), \text{ for all } X, Y \subseteq U.$$

The reason for its name is given by the following equivalent form

$$\mathbf{CR}^* \quad x \in X - f(X) \Rightarrow x \notin f(Y \cup \{x\}), \forall X, Y \subseteq U, \text{ such that } f(X) \subseteq Y,$$

requiring that if an alternative x is rejected from X, then x is rejected from any superset of $f(X) \cup \{x\}$. We can refer to both **CR** and **CR*** as the Conservative Rejection, as the following lemma confirms.

Lemma 83 *A choice function f satisfies **CR** if and only if it satisfies **CR***.*

Proof. **CR** $\Rightarrow$ **CR***. Let $X, Y \subseteq U$ such that $f(X) \subseteq Y$ and $x \in X - f(X)$. Then $f(X) \subset Y \cup \{x\}$. By **CR** we have $f(Y \cup \{x\}) \cap X \subseteq f(X)$. It follows $x \notin f(Y \cup \{x\})$, that is **CR*** holds.

CR* $\Rightarrow$ **CR**. Let $X, Y \subseteq U$ such that $f(X) \subseteq Y$. Suppose that **CR** does not hold. It follows that there is $x \in f(Y) \cap X$ such that $x \notin f(X)$. Hence $x \in X - f(X)$. Since $x \in f(Y) \subseteq Y$ it follows that $x \in f(Y \cup \{x\})$, that is **CR*** does not hold. □

The following example shows that Conservative Rejection is strictly stronger than substitutability. Let $U = \{1,2,3\}$ and the choice function f on U given by

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$f(A)$	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1\}$	$\{2\}$	$\{1\}$

.

It is easy to check that f satisfies **SUB**. However, for $X := \{1,2,3\}$, $x := 2$, and $Y := f(X)$ we have $2 = x \in X - f(X) = \{1,2,3\} - \{1\} = \{2,3\}$, but $2 = x \in f(f(X) \cup \{x\}) = f(\{1\} \cup \{2\}) = f(\{1,2\}) = \{1,2\}$, that is f does not satisfies **CR***.

A relaxation of the **CR*** condition, **Locally Conservative Rejection (LCR)**,

$$\mathbf{LCR} \qquad \text{if } a \in A - f(A) \text{ then } a \notin f(f(A) \cup \{a\}),\ \forall A \subseteq U,$$

can be used to express its relationship with substitutability.

Proposition 84 *A choice function f satisfies **CR** if and only if f satisfies **SUB** and **LCR**.*

Proof. We use Lemmas 81 and 83.

CR* $\Rightarrow$ **SUB** $\wedge$ **LCR** : By taking $X := A$, $x := a$ and $Y := f(A)$ in **CR***, we obtain **LCR**. Since **SUB** is equivalent to **Mon**, we prove that **CR*** $\Rightarrow$ **Mon** . Let $X \subseteq X'$. We prove that $X - f(X) \subseteq X' - f(X')$. Let $Y := X'$ and $x \in X - f(X)$. Then, since $X \subseteq X'$, we have $x \in X'$ and $Y \cup \{x\} = X'$. By **CR***, we have $x \notin f(X')$, and hence $x \in X' - f(X')$.

SUB $\wedge$ **LCR** $\Rightarrow$ **CR*** : Let $x \in X - f(X)$. Then, using **LCR**, we have $x \notin f(f(X) \cup \{x\})$. Applying **Mon** (which is equivalent to **SUB**), we obtain $x \notin f(Y \cup \{x\})$, for $Y \subseteq U$ with $f(X) \subseteq Y$. □

In the presence of **CR** the fixed points of a choice function have the following important property.

Proposition 85 *If the choice function f satisfies **CR** then $f(X)$ is a maximal (w.r.t. set inclusion) fixed point contained in X, for all $X \subseteq U$.*

Proof. By Proposition 84, f satisfies **SUB** and, by Proposition 80 (ii), $f(X)$ is a fixed point of f contained in X. Let P be a fixed point of f such that $f(X) \subseteq P \subseteq X$. By **CR** we have $f(P) \cap X \subseteq f(X)$, that is, $P \cap X \subseteq f(X)$, giving $P \subseteq f(X)$. □

Let us consider an established condition of choice functions known as **Outcast** (see Aizerman and Aleskerov [AA95])

Outcast if $f(X) \subseteq Y \subseteq X$ then $f(Y) = f(X)$, for all $X, Y \subseteq U$.

The importance of this condition in matching area (used by Blair [Bla88], and Roth [Rot84]) is well-known and it is equivalent to the following **Independence of Rejected Contracts (IRC)** emphasized by Aygün and Sönmez in [AS12]:

IRC if $x \notin f(X \cup \{x\})$ then $f(X \cup \{x\}) = f(X)$, for all $X \subseteq U$, and $x \in U$.

For the sake of completeness we give the proof here:

Lemma 86 *The choice function f satisfies **Outcast** if and only if f satisfies **IRC**.*

Proof. Outcast $\Rightarrow$ IRC . If $x \notin f(X \cup \{x\})$ then $f(X \cup \{x\}) \subseteq X \subseteq X \cup \{x\}$. By **Outcast** we have $f(X \cup \{x\}) = f(X)$, hence **IRC** holds.
IRC $\Rightarrow$ Outcast . We prove **Outcast** by induction on $|X - Y|$.
If $|X - Y| = 1$ then $X = Y \cup \{x\}$; the hypothesis of **Outcast** is $f(Y \cup \{x\}) \subseteq Y \subseteq Y \cup \{x\}$; the first inclusion implies $x \notin f(Y \cup \{x\})$, and by IRC we have $f(Y \cup \{x\}) = f(Y)$, that is $f(Y) = f(X)$.

In the inductive step, assume $Y = (X - Y') - \{x\}$, where $Y' \subseteq f(X) - X$ and $x \in X - Y'$. The hypothesis of **Outcast** is $f(X) \subseteq (X - Y') - \{x\} \subseteq X$. It follows $f(X) \subseteq (X - Y') \subseteq X$ and, by induction hypothesis, $f(X - Y') = f(X)$. Hence $f(X - Y') \subseteq (X - Y') - \{x\} \subseteq X - Y'$, and (by the induction basis) $f(X - Y' - \{x\}) = f(X - Y') = f(X)$. □

The main result of this subsection is the following characterization of choice functions satisfying **CR**.

Theorem 87 *The choice function f satisfies **CR** if and only if f satisfies **SUB** and **Outcast**.*

Proof. CR $\Rightarrow$ SUB $\wedge$ Outcast. By Propositions 84 we have **CR $\Rightarrow$ SUB** . We prove that **CR $\Rightarrow$ Outcast**. Let $f(X) \subseteq Y \subseteq X$. If **CR** holds, then, by Proposition 84 and Lemma 81, **CC** holds, and by Proposition 82 we have $f(X) \subseteq f(Y)$. If $f(Y) = f(X)$ then **Outcast** holds. Otherwise, there is $x \in f(Y) - f(X)$. Since $f(X) \subseteq Y \subseteq X$ we have $x \in X - f(X)$. By **CR***, we obtain $x \notin f(Y \cup \{x\}) = f(Y)$, a contradiction.

SUB $\wedge$ **Outcast** $\Rightarrow$ **CR** . By Proposition 84 it is sufficient to prove **Outcast** $\Rightarrow$ **LCR** . Let $a \in A - f(A)$. Then $f(A) \subseteq f(A) \cup \{a\} \subseteq A$. Using **Outcast** we have further $f(f(A) \cup \{a\}) = f(A)$, therefore $a \notin f(f(A) \cup \{a\})$. Hence **LCR** holds.

□

Remarks.

1. Aizerman and Malishevski [AM81] have shown that the conjunction of **SUB** and **Outcast** is equivalent to an influential and natural consistency condition for choice functions introduced by Plott [Plo73]

   Path Independence $\quad f(A \cup B) = f(f(A) \cup B)$, for all $A, B \subseteq U$.

   In words, **Path Independence** says that in order to evaluate $f(X)$ (for some $X \subseteq U$) we can decompose X in an arbitrary (finite) path of smaller parts X_i and replace them recursively by $f(X_i)$.

   Hence Theorem 87 gives a new characterization of choice functions satisfying **Path Independence**.

2. Another important characterization of the conjunction of **SUB** and **Outcast** was established also by Aizerman and Malishevski [AM81]: A choice function f satisfies **SUB** and **Outcast** if and only if

   $\exists n \in \mathbf{N}$ and functions $f_i : U \to \mathbf{N}, i \in \{1, \ldots, n\}$ such that $\forall X \subseteq U$

   $$f(X) = \bigcup_{i=1}^{n} \{x \in X | f_i(x) \geq f_i(y), \forall y \in X\}$$

A seemingly stronger form of the **CR** condition is the following **Strongly Conservative Rejection (SCR)** (which will be used in the next section):

SCR $\forall x \in U$, if $x \notin f(X \cup \{x\})$ then $x \notin f(Y \cup \{x\}), \forall X, Y \subseteq U$ such that $f(X) \subseteq Y$.

If, in **SCR**, we consider $x \in X$ then the hypothesis $x \notin f(X \cup \{x\})$ means $x \in X - f(X)$ and therefore **SCR** implies **CR***. Interestingly, the converse implication holds.

Proposition 88 *The choice function f satisfies **CR** if and only if f satisfies **SCR**.*

Proof. Using Lemma 83 and the above observation, it suffices to prove that if **CR** holds then

$$x \notin X \cup f(X \cup \{x\}) \Rightarrow x \notin f(Y \cup \{x\}), \forall Y \subseteq U, \text{ such that } f(X) \subseteq Y.$$

Let $x \in U - X$ and $x \notin f(X \cup \{x\})$. Since f satisfies **CR***, by Theorem 87, it follows that f satisfies **Outcast** and its equivalent form **ICR**. Therefore we have $f(X \cup \{x\}) = f(X)$. Let $X' := X \cup \{x\}$. We have $x \in X' - f(X')$. Since f satisfies **CR***, we obtain $x \notin f(Y \cup \{x\}), \forall Y \subseteq U$ such that $f(X') \subseteq Y$. Since $f(X') = f(X)$, **SCR** holds. □

5.4. Stable common fixed points of two choice functions

Let $f, g : 2^U \to 2^U$ be two choice functions. In this section we show that if the two functions satisfy **CR** condition, then a stable common fixed point of them can be found using the following **Immediate Rejection** algorithm, a generalization of the well-known Gale-Shapley **Deferred Acceptance** [GS62] algorithm.

Immediate Rejection Algorithm

```
Input : f,g : 2^U → 2^U choice functions
Output : M ⊆ U a stable common fixed point of f and g

    N ← U ;
    while g(f(N)) ≠ f(N) do
            N ← N − (f(N) − g(f(N))) ;
    M ← f(N) ;
    return M
```

In words: in each while-iteration, the elements rejected by g from those selected by f are removed from a current set N. When there are no rejected elements anymore, the set $M = f(N)$ is a stable common fixed point of f and g.

Let us apply this algorithm for the examples of Bipartite Choice Systems $(G = (S,T;E),\mathscr{C}^i)$ $(i = 1,2,3)$ considered in Section 5.2. Hence, $U := E = \{1,2,3\}$, $f := C_S^i$, and $g := C_T^i$. We also denote by $N^0, N^1, \dots$ the successive values of N during the algorithm.

1. $N^0 = \{1,2,3\}$, $f(N^0) = \{1,3\}$; $g(f(N^0)) = g(\{1,3\}) = \{1\}$; $N^1 = N^0 - (f(N^0) - g(f(N^0))) = \{1,2,3\} - (\{1,3\} - \{1\}) = \{1,2\}$; $f(N^1) = \{1\}$; $g(f(N^1)) = g(\{1\}) = \{1\}$; $M = f(N^1) = f(\{1,2\}) = \{1\}$. M is a stable common fixed point of f and g (in the terminology of Section 5.2, M is a stable choice matching in $(G = (S,T;E),\mathscr{C}^1)$, and we already verified this). Therefore, the **Immediate Rejection Algorithm** returns a stable common fixed point of f and g. Note that $g := C_T^1$ does not satisfies **CR** condition: $2 \notin g(\{1,2,3\} \cup \{2\}) = \{1\}$, but $2 \in g(g(\{1,2,3\}) \cup \{2\}) = g(\{1\} \cup \{2\}) = g(\{1,2\}) = \{1,2\}$.

2. In this case, we observed that $f := C_S^2$, and $g := C_T^2$ have no stable common fixed points. However, the **Immediate Rejection Algorithm** proceeds as follows: $N^0 = \{1,2,3\}$, $f(N^0) = \{1,3\}$; $g(f(N^0)) = g(\{1,3\}) = \{3\}$; $N^1 = N^0 - (f(N^0) - g(f(N^0))) = \{1,2,3\} - (\{1,3\} - \{3\}) = \{2,3\}$; $f(N^1) = \{2,3\}$; $g(f(N^1)) = g(\{2,3\}) = \{2\}$; $N^2 = N^1 - (f(N^1) - g(f(N^1))) = \{2,3\} - (\{2,3\} - \{2\}) = \{2\}$; $f(N^2) = \{2\}$; $g(f(N^2)) = g(\{2\}) = \{2\}$; $M = f(N^2) = f(\{2\}) = \{2\}$. Hence the **Immediate Rejection Algorithm** fails. Note that, again $g := C_T^1$ does not satisfies **CR** condition: $3 \notin g(\{1,2\} \cup \{3\}) = \emptyset$, but $3 \in g(g(\{1,2\}) \cup \{3\}) = C_t(\{1\} \cup \{3\}) = \{3\}$.

3. In this case, we observed that $f := C_S^3$, and $g := C_T^3$ have exactly one stable common fixed point, namely $\{1,2\}$. However, the *Immediate Rejection* algorithm proceeds as follows: $N^0 = \{1,2,3\}$, $f(N^0) = \{1,2,3\}$; $g(f(N^0)) = g(\{1,2,3\}) = \{1\}$; $N^1 = N^0 - (f(N^0) - g(f(N^0)) = \{1,2,3\} - (\{1,2,3\} - \{1\}) = \{1\}$; $f(N^1) = \{1\}$; $g(f(N^1)) = g(\{1\}) = \{1\}$; $M = f(N^1) = f(\{1\}) = \{1\}$. Hence the **Immediate Rejection Algorithm** fails (M is only a common fixed point that is not stable). Note that g is the same as in Example 1, therefore does not satisfy **CR** condition.

The next theorem shows the correctness of the **Immediate Rejection Algorithm** when the two choice functions f and g satisfy the **CR** condition.

Theorem 89 *If f and g satisfy **CR**, then the **Immediate Rejection Algorithm** returns a stable common fixed point of f and g.*

Proof. Since N shrinks in every iteration, the algorithm terminates.

Let k be the number of iterations and let N^0 to N^k be the values of N during the algorithm. Then $N^0 = U$, $N^j = N^{j-1} - B^j$, where $B^j = f(N^{j-1}) - g(f(N^{j-1}))$, for $j = 1, \ldots, k$, and $M = f(N^k)$.

We first show that M is a common fixed point of f and g.
By the condition of the while-loop, we have $g(f(N^k)) = f(N^k)$, therefore $g(M) = M$, since $M = f(N^k)$. Also, $f(M) = f(f(N^k)) = f(N^k) = M$ since f is idempotent, by Propositions 80 and 84.

We next show that M is a stable common fixed point of the two functions, i.e., $a \notin f(M \cup \{a\}) \cap g(M \cup \{a\})$ for every $a \in U - M$.
The elements not in M are the elements removed from N during the while-loop and the elements in N^k rejected by f, i.e.,

$$U - M = (U - N^k) \cup (N^k - M).$$

Let $a \in U - M$ be arbitrary. We distinguish two cases according to which of the above (disjoint) two sets a belongs.

Case $\mathbf{a \in U - N^k}$. Let i be such that $a \in B^i = N^{i-1} - N^i$. Then $a \in f(N^{i-1})$ and $a \notin g(f(N^{i-1}))$. Since $f(N^{i-1}) = f(N^{i-1}) \cup \{a\}$ we can write $a \notin g(f(N^{i-1}) \cup \{a\})$.

We show $a \notin g(f(N^j) \cup \{a\})$, for $j = i, \ldots, k$, by induction on j.
Since $N^j = N^{j-1} - B^j$ and f satisfies **CC**, we have

$$f(N^{j-1}) - B^j \subseteq f(N^{j-1} - B^j) = f(N^j).$$

Since $g(f(N^{j-1})) = f(N^{j-1}) - B^j$, this implies

$$g(f(N^{j-1})) \subseteq f(N^j).$$

By induction hypothesis $a \notin g(f(N^{j-1}) \cup \{a\})$. By Proposition 88, g satisfies **SCR** (with $X = f(N^{j-1})$ and $Y = f(N^j)$), therefore $a \notin g(f(N^j) \cup \{a\})$, and the induction step is complete.

We have established $a \notin g(f(N^k) \cup \{a\}) = g(M \cup \{a\})$. Thus $a \notin f(M \cup \{a\}) \cap g(M \cup \{a\})$.

Case $\mathbf{a} \in \mathbf{N^k} - \mathbf{M}$. Then $a \notin f(N^k)$. By Proposition 84, f satisfies **LCR**, therefore $a \notin f(f(N^k) \cup \{a\}) = f(M \cup \{a\})$. Thus $a \notin f(M \cup \{a\}) \cap g(M \cup \{a\})$.

□

By Proposition 84, the **CR** condition is the conjunction of **SUB** and **LCR**. While substitutability is known to be a necessary condition for the correctitude of the generalizations of Gale and Shapley algorithms (see Hatfield and Kominers [HK16]), the **LCR** condition (expressing the external stability of $f(X)$ in any $X \subseteq U$) is also necessary as the following proposition shows. Let us denote by **Immediate Rejection**(f, g, U') the subset of $U' \subseteq U$ returned by the above algorithm.

Proposition 90 *Let f be an idempotent choice function on U such that **Immediate Rejection**(f, g, U') is a stable common fixed point of f and g contained in U', for any choice function g on U satisfying **CR**. Then f satisfies **LCR**.*

Proof. Suppose that f does not satisfy **LCR**. Then, there is $X \subseteq U$ and $x \in X - f(X)$ such that $x \in f(f(X) \cup \{x\})$. Take $U' := X$ and g the identity choice function on U: $g(A) = A$, for every $A \subseteq U$. Clearly, g satisfies **CR**. Then, **Immediate Rejection**$(f, g, U') = f(X)$ since $f(X)$ is accepted by g and $f(f(X)) = f(X)$, by idempotency of f. However, $f(X)$ is not a stable common fixed point contained in X since $x \in X - f(X)$ and $x \in f(f(X) \cup \{x\}) \cap g(f(X) \cup \{x\})$. This contradicts the hypothesis that the **Immediate Rejection** works properly, and the proposition is proved. □

Part III.

Opposition Frameworks

In this part (based on the paper [CM16]) we introduce opposition frameworks, a generalization of Dung's argumentation frameworks. While keeping the attack relation as the sole type of interaction between nodes and the abstract level and simplicity of argumentation frameworks, opposition networks add more flexibility, reducing the gap between structured and abstract argumentation. A *guarded attack calculus* is developed in order to obtain proper generalizations of Dung's admissibility-based semantics. The high modeling capabilities of our new setting offer an alternative instantiation solution (of other existing argumentation frameworks) for arguments evaluation.

6. A New Generalization of Argumentation Frameworks

6.1. Introduction

In this chapter we introduce **Opposition Frameworks** (`OF`s for short) that generalize **Argumentation Frameworks** (`AF`s for short) without self-attacks, by considering more fine-grained notions of conflict-freeness and admissibility.

An `OF` is a labeled directed multigraph whose directed edges represent attacks between its nodes. A node is no longer an atomic argument as in `AF`s, but a composed object, interpreted as the *position* of an agent in a debate. The position of a node v is a finite set $g(v)$ of *facts granted* by v. Depending on the real world problem modeled, the facts can be statements, claims, pieces of evidence, locutions, issues, etc. The set $g(v)$ of facts granted by a node v has no mathematical structure associated; it is, simply, a list of facts approved by the node v, based on which v develops its attacks. Each attack a has a *source* node $s(a)$, a target node $t(a)$, and is labeled by a pair $(\gamma(a), \delta(a))$ of (disjoint) sets of facts. Here $\gamma(a) \subseteq g(s(a))$ is the *guard* of the attack a, and $\delta(a)$ is a nonempty subset of $g(t(a)) - g(s(a))$, representing the facts (granted by the node $t(a)$) that are *denied* by the source node, $s(a)$, of the attack a. So, if a node v attacks a node w via the attack a, that is $s(a) = v$ and $t(a) = w$, then the guard $\gamma(a)$ specifies the set of facts – granted by v – based on which v does not admit the facts in $\delta(a)$ – granted by w.

It follows that, in `OF`s, arguments are seen as ensembles formed by the facts granted by a node together with the attacks issuing from this node. To illustrate how this can arise in real world situations, let us consider the following possible political debate.

Example 1 (adapted from Wang and Luo [WL10]). Let us construct an `OF` by assigning a node for each of the 5 positions in the following debate on the set $\{f_1, \ldots, f_9\}$ of facts:

v_1 : ”$\overline{\text{Reducing emissions of greenhouse gases}}^{f_1}$ is crucially for the $\overline{\text{protection of our health}}^{f_2}$ and, clearly, it is more important than $\overline{\text{developing economy}}^{f_3}$.”

v_2 : ”$\overline{\text{Developing economy}}^{f_3}$ will ensure $\overline{\text{creating job positions}}^{f_4}$ and it is obviously more significant than simply $\overline{\text{protecting the environment}}^{f_8}$.”

v_3 : ”We do not have to focus on $\overline{\text{developing economy}}^{f_3}$ but instead urgently should take measures in order to $\overline{\text{protect environment}}^{f_8}$, e.g., $\overline{\text{reduce emissions of greenhouse gases}}^{f_1}$, and $\overline{\text{save water}}^{f_5}$.”

v_4 : ”Currently it is more important to $\overline{\text{create job positions}}^{f_4}$ for $\overline{\text{increasing number of graduates}}^{f_6}$, than being concerned with $\overline{\text{reducing emissions of greenhouse gases}}^{f_1}$. It is obvious that by $\overline{\text{hiring new people}}^{f_9}$ it is not necessary to be concerned with $\overline{\text{protecting human health}}^{f_2}$.”

v_5 : ”Instead of $\overline{\text{creating several job positions}}^{f_4}$ in order to $\overline{\text{increase the number of graduates}}^{f_6}$, we should concentrate on $\overline{\text{protecting the environment}}^{f_8}$ to guarantee $\overline{\text{earth security}}^{f_7}$.”

Graphically, in Figure 6.1 below, each node is decorated with its granted set and each attack a, with source $s(a) = v_i$ and target $t(a) = v_j$, is labeled with the pair $(\gamma(a), \delta(a))$.

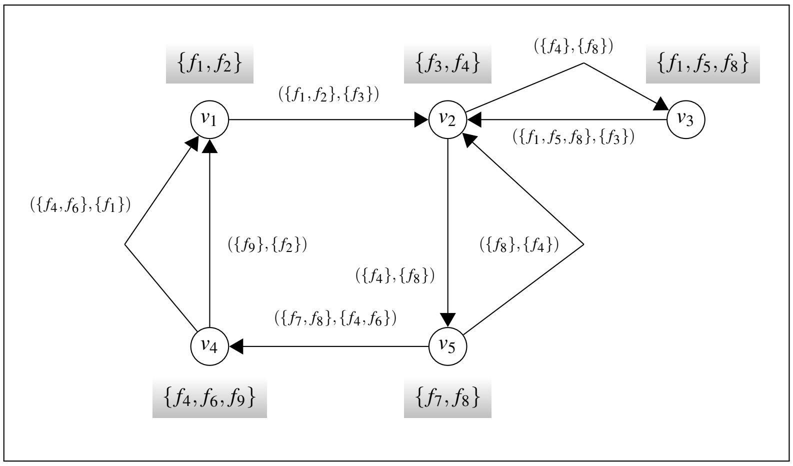


Figure 6.1.: OF modeling the political debate in Example 1.

Note that we have two attacks from v_4 to v_1 which differ by their labels, and this kind of multiple attacks does not exist in Dung's argumentation frameworks.

This example is considered only for illustration purpose. A software system for automatically modeling such a debate, that is to construct positions, facts and attack's labels, must use appropriate natural language processing tools, e.g., Finegan-Dollak and Radev [FDR15], and/or specialized debating websites, e.g., Debatepedia[1], (see also Rahwan et al. [RZR07], or Leite and Martins [LM11]), but it is beyond the scope of this work.

The above example shows that an OF can model which precise part of an ”argument” is in conflict with which part of another ”argument”, without requiring a logical language and an inference relation. Our new formalism is more abstract than the existing structured AFs and it is well suited to represent complex non-logical information.

Intuitively, a *guarded attack* shows the reason why a node attacks another one. Let a be an attack on the node w (that is, $t(a) = w$). We say that a is *harmful to* w if w can not counter-attack a: there is no attack a' from w to the source of a

[1] http://dbp.idebate.org

denying at least one fact in the guard of a ($\forall a'$ with $s(a') = w$ and $t(a') = s(a)$ we have $\delta(a') \cap \gamma(a) = \emptyset$). We develop a *guarded attack calculus* in order to extend the basic notions which underly the classical Dung's semantics, so preserving the diversity of reasoning schemes for AFs. Our new formalism is based on the graph operation of *contraction* and can be described as follows. To decide if a given node v can be *accepted*, we look at the attacks on v; if there is no attack harmful to v, then it is accepted; otherwise, we search a node w which is not in "conflict" with v (no fact denied by an attack from one is granted by the other) and denies at least one fact of the guard of an attack harmful to v; if w does not exist, then v can *not be accepted*; if w exists, we consider their *coalition* $\{v, w\}$ as a new super-node $v_{\{v,w\}}$ with $g(v_{\{v,w\}}) = g(v) \cup g(w)$, delete the nodes v and w, and replace them by the super-node $v_{\{v,w\}}$ as the source or the target of each attack from or to v and w; the decision process is continued using the super-node in the new `OF`. Adapting the notions of *conflict-freeness* and *defense*, we show that there is a sequence of coalition choices for which the above outlined process ends with an accepted (super)node if and only if there is an *admissible set* of nodes containing v. Since `AF`s are particular `OF`s (see the end of Subsection 6.2.1), we obtain a more intuitive and algorithmic way of handling classical admissibility argumentation semantics.

Returning to the Example 1, if we want to see the status of v_2 in this `OF` using the above outlined process, it is obvious that it needs to make a coalition with v_4 in order to deny a fact in the guard of the attack from v_1. The super-node $v_{\{v_2,v_4\}}$ has no harmful attack in the contracted `OF`, therefore v_2 is accepted.

If we delete the set-labels, replace the multiple attack by a single directed edge and call the nodes arguments in the `OF` in Figure 6.1, we obtain (the digraph of) an `AF`; in this `AF` the set of arguments $\{v_2, v_4\}$ is a *preferred* and *stable extension* (see Section 0.2). Hence the outputs of the two frameworks agree. On the other hand, the set $\{v_1, v_3, v_5\}$ is another *preferred* and *stable extension* in the `AF`. However, the set of nodes $\{v_1, v_3, v_5\}$ can not be considered as a "solution" in the `OF` in Figure 6.1, since it does not (collectively) defend the node v_1 against the attack with source v_4 and labeled $(\{f_9\}, \{f_2\})$. It follows that the use of guards, on which `OF`s are developed, provides more accurate outputs than the dichotomy between the existence and lack of attacks, on which Dung's frameworks are based (further differences are highlighted in Subsection 6.2.1 in the comments after Figure 6.2).

Note the difference between our `OF`s and structured/deductive argumentation frameworks (e.g., Caminada and Amgoud [CA07], Prakken [Pra10], Hunter and Gorogiannis [HG11], Amgoud [Amg14]): while in these logic based frameworks the internal structure of the arguments generates and explains the (inferential) nature of the attacks expressed as uniform (i.e., the same for all nodes) rules, the users of `OF`'s are free to choose between uniform or non-uniform rules to construct the attack's labels. These labels can be automatically constructed if the content of a node (that is, its granted set of facts) is equipped with a mathematical (logical, combinatorial, algebraic, etc.) structure. The gain over the AF's instantiation approach is that our model is more general and the use of attack's labels reduces the number of "arguments" to be considered (see also the discussion after the Proposition 101 in the end of Section 6.2.5).

6.2. Opposition Frameworks

6.2.1. Defining the new framework

In this subsection, we define OFs, discuss their compatibility with Dung's structures and specify how to see AFs as OFs.

Definition 91 (**Opposition Framework** (OF)) *An opposition framework is a tuple $OF = (N,F,g,A,s,t,\gamma,\delta)$, where:*

- *N is a finite set of **nodes**; F is a finite set of **facts**; $g : N \to 2^F$ is a function that associates to each node $v \in N$, its **granted** set $g(v)$ of facts in F,*
- *A is a finite set of **attacks**; $s,t : A \to N$ are functions that associate to each attack $a \in A$ its **source** node $s(a)$, and its **target** node $t(a)$,*
- *$\gamma, \delta : A \to 2^F$ are functions that associate to each attack $a \in A$ its **guard** $\gamma(a) \subseteq g(s(a))$, and its **denied set of facts** $\delta(a)$, with $\delta(a) \subseteq g(t(a)) - g(s(a))$.*

In words, an OF is a labeled multi-digraph in which each directed edge (v,w) corresponds to an attack a from the node $s(a) = v$ to node $t(a) = w$, that is based on a set of facts $\gamma(a) \subseteq g(v)$ granted by v, and denies the set $\delta(a) \subseteq g(w)$ of facts granted by w. Throughout this thesis we assume that the sets $g(v)$, $\gamma(a)$, $\delta(a)$ are non-empty.

The condition $\delta(a) \subseteq g(t(a)) - g(s(a))$ forbids the attack a to deny the facts granted by its source. In particular, there is no attack in A such that $s(a) = t(a)$ (there are no self-attacks). However, we can have parallel attacks: a set of attacks $A_0 \subseteq A$ with $|A_0| \geq 2$ and $s(a) = s(a')$ and $t(a) = t(a')$, for every $a, a' \in A_0$.

Graphically, each node is decorated with its granted set and each attack a with source $s(a) = v$ and target $t(a) = w$ is labeled with the pair $(\gamma(a), \delta(a))$, as depicted in Figure 6.1. Let us observe that the granted sets of nodes v_1 and v_4 are $g(v_1) = \{f_1, f_2\}$ and $g(v_4) = \{f_4, f_6, f_9\}$. The two attacks ($a_1$ and a_2) from v_4 to v_1 form a multiple attack of v_4 against v_1. They differ by their labels: $(\gamma(a_1), \delta(a_1)) = (\{f_4, f_6\}, \{f_1\})$ and $(\gamma(a_2), \delta(a_2)) = (\{f_9\}, \{f_2\})$. The meaning is that the attack a_1, based on f_4 and f_6 (granted by the source node of a_1), denies f_1, one of the facts granted by the target of a_1, while the attack a_2, based on f_9, denies f_2. We will assume in the following that multiple attacks have different labels.

The granted set $g(v)$ of a node v can be interpreted as a node interface, exhibiting its pieces of evidence which can be accepted or attacked by the other nodes.

The set $g(v)$ can not be replaced in the digraph representing the `OF` by a set of non-conflicting nodes (viewing the items in $g(v)$ as sub-arguments) due to the rule-based way the attacks are conceived. For example, in Figure 6.2 (i), we consider an `OF` having only two nodes v_1 and v_2, with $g(v_1) = \{f_1, f_2, f_3\}$ and $g(v_2) = \{f_2, f_4\}$, and a symmetric pair of attacks a_1 and a_2. In Figure 6.2 (ii), we transform this `OF` into a digraph (`AF`) with vertices set $\{f_1, f_2, f_3, f_4\}$ and attacks generated by the rules associated to a_2 $((\{f_2\}, \{f_1\}))$ gives the attack (f_2, f_1)) and a_1 $((\{f_3\}, \{f_4\}))$ gives the attack (f_3, f_4)). But then, the set $\{f_1, f_2, f_3\}$ is not conflict free, which contradicts the intuition that in the `OF` the set $g(v)$ granted by a node v is not conflicting.

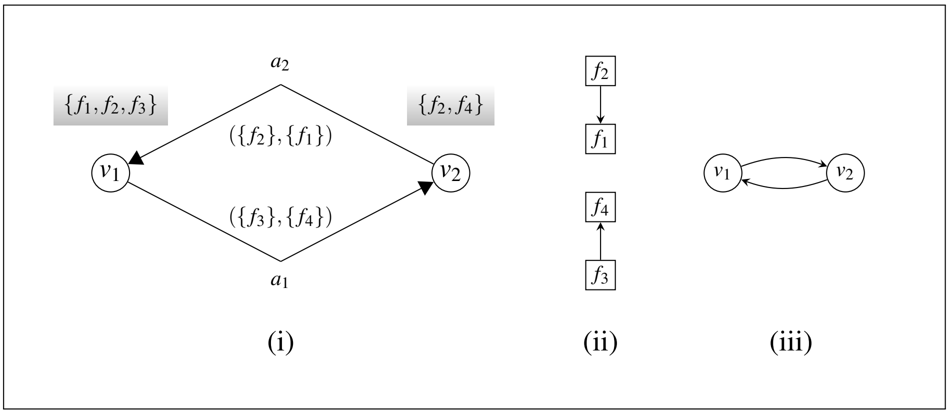


Figure 6.2.: Trying to model an `OF` as an `AF`.

Hence `OF`s offer a more general approach of modeling collective attacks as the one proposed in Nielson and Parsons [NP06], where sets of arguments rather than single arguments may be needed to attack another argument.

On the other hand, if we view the `OF` in (i) simply as the `AF` with two arguments v_1 and v_2 attacking each other (as in Figure 6.2 (iii)), then each argument is credulously stable accepted (see Coste-Marquis *et al.* [CMDM05b]), which is not what the `OF` suggests (the attack a_2 against v_1 can not be defended, because of f_2, which is not denied, and the attack a_1 against v_2 can not be defended, because of f_3). It follows that the semantics of the two models (`AF` and `OF`) differ.

It follows that the acceptability based semantics for `OF`s must be defined in an appropriate way in order to capture the intended intuition and, therefore, to have stronger versions of the Dung's semantics imposed by the fined-grained environment considered in `OF`s.

We close this subsection by noting that any `AF` without self-attacks can be seen as a (trivial) `OF`, specified (for further use) in the following definition, and illustrated by a simple example in Figure 6.3.

Definition 92 (**`OF` associated to an `AF`**) *Let $AF = (Arg(AF), Def(AF))$ be an argumentation framework without self-attacks. The* ***opposition framework associated to*** *AF is $OF_{AF} = (N, F, g, A, s, t, \gamma, \delta)$, where:*
- *$N = F = Arg(AF)$; $g(a) = \{a\}$, $\forall a \in Arg(AF)$; $A = Def(AF)$;*
- *$\forall d = (a, b) \in Def(AF)$, $s(d) = a$ and $t(d) = b$, $\gamma(d) = \{a\}$ and $\delta(d) = \{b\}$.*

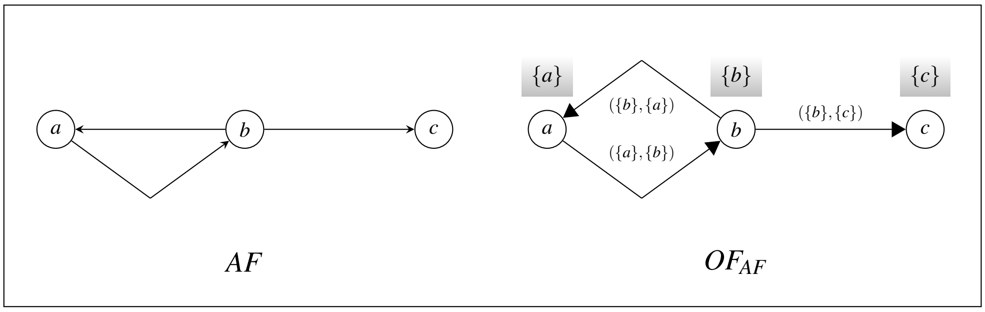


Figure 6.3.: `OF` associated to an `AF`.

6.2.2. Conflict-freeness

Let $OF = (N, F, g, A, s, t, \gamma, \delta)$ be an opposition framework. For $S \subseteq N$, we denote by $\alpha^-(S)$ ($\alpha^+(S)$) the set of attacks having the target (source) in S:

$\alpha^-(S) = \{a \in A | t(a) \in S\}$ and $\alpha^+(S) = \{a \in A | s(a) \in S\}$.

We write $\alpha^+(v)$ ($\alpha^-(v)$) instead of $\alpha^+(\{v\})$ ($\alpha^-(\{v\})$).

A **weak conflict-free (wcf)** set of nodes is any set $S \subseteq N$ such that there is no attack a with $s(a), t(a) \in S$. Clearly, singletons $\{v\}$, for $v \in N$, are wcf sets.

In `AF`s, wcf sets are simple called conflict-free sets, and can be conceived as collective (super)arguments. Formally, if S is a conflict-free set of arguments in an `AF`, we can replace it by a (super)node a_S and each attack from (to) an argument in S is replaced by an attack from (to) v_S (multiple attacks are replaced by a single attack). This graph operation (called contraction) creates no self-attacks, since S is a conflict-free set. The contraction operation of a wcf set S in an `OF` assigns the union of the granted sets of the members of S as the granted set of v_S, the sources (or targets) of attacks from (to) a member of S are replaced by v_S (multiple attacks are accepted).

The problem that can arise is that for some new attack a with $s(a) = v_S$ we can have $\delta(a) \cap g(v_S) \neq \emptyset$. This can be avoided as follows.

A **strong conflict-free (scf)** set of nodes is any set $S \subseteq N$ such that $\forall a \in \alpha^+(S)$, $\delta(a) \cap g(s') = \emptyset$, for every $s' \in S$. In words, no attack with source in S denies a fact granted by a node in S.

The following lemma holds.

Lemma 93 *(i) In an* $\texttt{OF}$ *any scf set is a wcf set. (ii) If in an* $\texttt{OF}$ *the granted sets,* $g(v)$*, are disjoint (i.e.* $g(v) \cap g(w) = \emptyset$*, for all distinct* $v, w \in N$*) then a set of nodes is wcf if and only if it is scf. In particular, if* OF_{AF} *is the* $\texttt{OF}$ *associated to an* $\texttt{AF}$ AF*, then a set* S *of arguments in* AF *is conflict-free if and only if it is scf in* OF_{AF}*.*

Proof. (i) Let $S \subseteq N$ be a scf set. If S is not a wcf set, there is an attack a with $s(a), t(a) \in S$. By the definition of functions δ and γ, we must have $\gamma(a) \subseteq g(s(a))$, and $\delta(a) \subseteq g(t(a)) - g(s(a))$. Let $s = s(a)$ and $s' = t(a)$. Then, $a \in \alpha^+(S)$ and $\delta(a) \cap g(s') \neq \emptyset$. This contradicts the hypothesis that S is a scf set.

(ii) Suppose that $S \subseteq N$ is a wcf set. Let $a \in \alpha^+(S)$. Since S is a wcf set, we have $t(a) \notin S$. Since the granted sets of nodes are disjoint, it follows $g(t(a)) \cap g(s') = \emptyset$, for every $s' \in S$. Hence S is a scf set. □

Scf sets can be safely used in the $\texttt{OF}$'s contraction operation:

Proposition 94 *Let* $OF = (N, F, g, A, s, t, \gamma, \delta)$ *be an opposition framework and* $S \subseteq N$ *a wcf set of nodes. Let* $OF|_S = (N_1, F_1, g_1, A_1, s_1, t_1, \gamma_1, \delta_1)$ *be the tuple, where*
$N_1 = (N - S)\dot{\cup}\{v_S\}$*,* $F_1 = F$*,* $A_1 = A$*,* $\gamma_1 = \gamma$*,* $\delta_1 = \delta$*,*

$$g_1(v) = \begin{cases} g(v) & \text{if } v \in N - S \\ \bigcup_{w \in S} g(w) & \text{if } v = v_S \end{cases}$$

$$s_1(a) = \begin{cases} s(a) & \text{if } a \notin \alpha^+(S) \\ v_S & \text{if } a \in \alpha^+(S) \end{cases} \text{ and } t_1(a) = \begin{cases} t(a) & \text{if } a \notin \alpha^-(S) \\ v_S & \text{if } a \in \alpha^-(S) \end{cases}.$$

Then $OF|_S$ *is an* $\texttt{OF}$ *(obtained* ***by contraction of*** S*) if and only if* S *is a scf set of nodes.*

Proof. Suppose that $OF|_S$ is an opposition framework. To show that S is a scf set in OF, let $a \in A$ with $s(a) \in S$. In $OF|_S$ we have $s_1(a) = v_S$ and therefore $\delta_1(a) \cap g_1(v_S) = \emptyset$. By the definition of functions δ_1 and g_1, it follows that in OF we have $\delta(a) \cap \bigcup_{w \in S} g(w) = \emptyset$, that is $\delta(a) \cap g(w) = \emptyset$, for every $w \in S$. Hence S is a scf set in OF.

Conversely, suppose that S is a scf set in OF. To show that $OF|_S$ is an $\texttt{OF}$, we have to prove that $\delta_1(a) \subseteq g_1(t_1(a)) - g_1(s_1(a)), \forall a \in A_1 = A$. By the definition of $OF|_S$, this holds trivially for every a with $s_1(a) \neq v_S$. If $s_1(a) = v_S$, then $t_1(a) \notin S$ (since S is wcf set) and , therefore, $\delta_1(a) = \delta(a)$, $t_1(a) = t(a)$, and $g_1(t_1(a)) = g(t(a))$. Because $g_1(s_1(a)) = g_1(v_S) = \bigcup_{w \in S} g(w)$ we have to prove that $\delta(a) \subseteq g(t(a)) - \bigcup_{w \in S} g(w)$. But this holds, since $s(a) \in S$ and S is a scf set (i.e., $\delta(a) \cap g(w) = \emptyset$ for every $w \in S$). □

6.2.3. Extending Dung's semantics

Let $OF = (N,F,g,A,s,t,\gamma,\delta)$ be an opposition framework, $S \subseteq N$ and $v \in N$. Let us denote $g(S) = \bigcup_{s \in S} g(s)$. We say that v is **defended by** S if

- $\forall a \in \alpha^+(v)$ we have $\delta(a) \cap g(S) = \emptyset$, and
- $\forall a \in \alpha^-(v)$, $\exists a^c \in \alpha^+(S)$ such that $t(a^c) = s(a)$ and $\delta(a^c) \cap \gamma(a) \neq \emptyset$.

In words, a node v is defended by a set S of nodes if, firstly, no attack with the source v denies a fact granted by a node in S, and, secondly, for any attack a targeting v there is a counter-attack coming from S, targeting the source of a, and denying at least one fact of the guard of a.

In Figure 6.4, the set $\{v_3\}$ defends the node v_1, but the set $\{v_4\}$ does not defend v_1 despite of the attack of $\{v_4\}$ (against the attacker v_2 of v_1), which doesn't deny the fact f_4. Note that there is no attack with source v_1, hence the first condition in the definition of defense holds trivially.

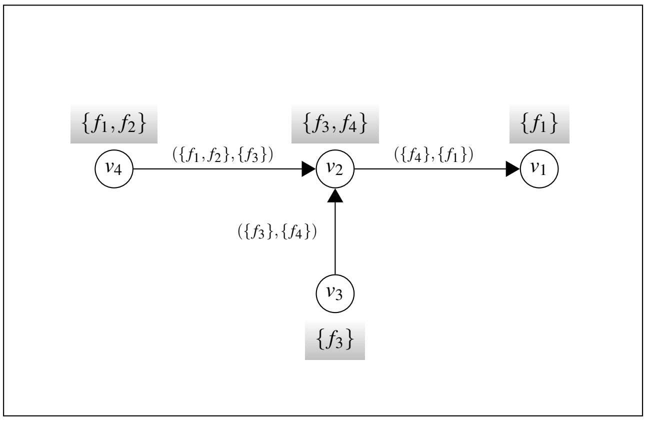


Figure 6.4.: $\{v_3\}$ defends v_1, but $\{v_4\}$ does not defend v_1.

The set of arguments defended by a set S is denoted by $\mathbb{D}(S)$.

Definition 95 *An **admissible set** is any scf set $S \subseteq N$ with the property that for any $a \in \alpha^-(S)$ there is $a^c \in \alpha^+(S)$ such that $t(a^c) = s(a)$ and $\delta(a^c) \cap \gamma(a) \neq \emptyset$.*

In words, a scf set is admissible if the guard of any attack targeting a member of S has at least one fact that is denied by an attack with source in S.

The following proposition gives some basic properties of admissible sets.

Proposition 96 *(i) In an* OF *a set S of nodes is admissible if and only if $S \subseteq \mathbb{D}(S)$.*
(ii) (Dung's Fundamental Lemma for OFs) *Let $OF = (N,F,g,A,s,t,\gamma,\delta)$ be an* OF*, $S \subseteq N$ an admissible set, and $u, v \in \mathbb{D}(S)$. Then,*

1. *$S' = S \cup \{u\}$ is an admissible set, and*

2. *if $\{u,v\}$ is a scf set then $v \in \mathbb{D}(S')$.*

(iii) A set S of arguments in an argumentation framework AF is admissible in AF if and only if it is admissible in OF_{AF}.

Proof. (i) If $S \subseteq \mathbb{D}(S)$ then, by the first condition in the definition of the defense of a node by a set of nodes, it follows that S is a scf set. The second condition in the same definition shows that S is an admissible set. Conversely, if S is an admissible set and $v \in S$, since S is a scf set it follows that $\forall a \in \alpha^+(v)$ we have $\delta(a) \cap g(s') = \emptyset$, for every $s' \in S$. By the definition of an admissible set, it follows that any attack targeting v has at least one fact that is denied by an attack with source in S. Hence $v \in \mathbb{D}(S)$.

(ii) Since S is admissible, it is a scf set. From $u \in \mathbb{D}(S)$ it follows that $S' = S \cup \{u\}$ is a scf set. Any attack targeting a member of S' has at least one fact that is denied by an attack with source in S, since S is admissible and $u \in \mathbb{D}(S)$. Hence S' is an admissible set. To prove the second statement, observe that it is sufficient to prove that $S' \cup \{v\}$ is a scf set. This follows since $S \cup \{u\}$ and $S \cup \{v\}$ are scf sets and the hypothesis that $\{u,v\}$ is a scf set (note that this hypothesis is not necessary for `AF`s).

(iii) If S is an admissible set of arguments in AF then it is conflict-free in AF and therefore is a scf set in OF_{AF}, by Lemma 93 (ii). Furthermore in AF we have $S^- \subseteq S^+$ and, by Definition 92, any attack targeting a member of S has at least one fact that is denied by an attack with source in S. Hence, S is an admissible set in OF_{AF}. The converse implication can be proved in a similar way. □

By Proposition 96, Dung's admissibility based extensions for `AF`s can be extended to `OF`s as follows: in an `OF`

- a **complete extension** is an admissible set S satisfying $\mathbb{D}(S) = S$,
- a **preferred extension** is a maximal (w.r.t. set inclusion) complete extension,
- a **grounded extension** is a minimal (w.r.t. set inclusion) complete extension, and
- a **stable extension** is an admissible set S with the property that each node $v \notin S$ is the target of an attack in $\alpha^+(S)$.

To keep things simple, we will consider here a simple form of acceptance of a node in an `OF`, which corresponds to credulously preferred acceptance in `AF`s.

Definition 97 *A node v is **accepted** in an `OF` if there is an admissible set S in `OF` containing v; otherwise it is **rejected**. The node v is **inceptively accepted** if $\{v\}$ is an admissible set.*

Using the definition of an admissible set (Definition 95), Propositions 94 and 96, we obtain the following result, which is very useful from the algorithmic point of view.

Proposition 98 *A node v is accepted in an opposition framework OF if and only if there is a scf set S such that $v \in S$ and v_S is inceptively accepted in $OF|_S$.*

Since AFs are very special cases of OFs, using Proposition 96 (iii), we obtain the following result on the complexity of deciding if a node can be accepted.

Proposition 99 *Deciding if a node is accepted in an* OF *is an* NP*-complete problem.*

Proof. The hardness follows by adapting a known polynomial reduction from the satisfiability problem given by Dimopoulos and Torres [DT96] or Dunne and Bench-Capon [DBC02] for AFs (see Section 0.3). Obviously, verifying if a guessed set contains the given node and is admissible can be done in polynomial time (in the "size" of the OF), so the decision problem is in NP. □

6.2.4. A DPLL type Acceptance Algorithm

In this subsection, we give a Davis-Putnam-Logemann-Loveland (DPLL) type algorithm (Davis *et al.* [DLL62]) for deciding the acceptance of a node in an OF, that improves a backtrack search exhaustive algorithm by the eager use of a "unit rule" at each step in the construction of a scf set S as in Proposition 98.

Let $OF = (N, F, g, A, s, t, \gamma, \delta)$ be an OF and $v \in N$. If we denote by $\Delta(v) = \bigcup_{a \in \alpha^+(v)} \delta(a)$ the set of facts denied by the attacks out of v, then the set of *harmful attacks* to v is

$$\alpha^-_{harm}(v) = \{a \in \alpha^-(v) \,|\, \gamma(a) \cap \Delta(v) = \emptyset \text{ or } s(a) \notin \bigcup_{a' \in \alpha^+(v)} t(a')\}.$$

If $\alpha^-_{harm}(v) = \emptyset$ then v is inceptively accepted, otherwise we are looking for a node w to make a coalition with v in order to deny as many attacks as possible from $\alpha^-_{harm}(v)$. The coalition $\{v, w\}$ must be a scf set, therefore w must belong to the *backing* set of nodes associated to v:

$$bck(v) = \{w \in N \,|\, \{v, w\} \text{ is scf, } \Delta(w) \cap \big(\bigcup_{a \in \alpha^-_{harm}(v)} \gamma(a)\big) \neq \emptyset\}.$$

For each attack $a \in \alpha^-_{harm}(v)$, we denote by $nem(a)$ the set of *nemeses nodes* in $bck(v)$ which have at least one attack that denies at least one fact in the guard of a:

$$nem(a) = \{w \in bck(v) \,|\, \exists a' \in \alpha^+(w) \cap \alpha^-(s(a)) \text{ s.t. } \delta(a') \cap \gamma(a) \neq \emptyset\}.$$

If there is $a \in \alpha^-_{harm}(v)$ with $nem(a) = \emptyset$, then v can make no coalition in order to counterattack a, so v can not be accepted. If $nem(a)$ is a singleton, $nem(a) = \{w_0\}$,

then v is *forced* to make a coalition with w_0 for denying at least one fact in $\gamma(a)$ (there is no scf set S such that $v \in S$, $w_0 \notin S$ and S counterattacks the attack a on v). This is the "unit rule" which will be followed every time when a candidate for coalition is searched.
If $|nem(a)| \geq 2$, then v must try to make coalitions with each node in $nem(a)$ to see if it can extend to a self-defending scf set.

The resulting algorithm can be described as follows:

```
Accept(OF, v)
------------------------------------------------------------
  Input:  OF = (N,F,g,A,s,t,γ,δ)  an OF,  v ∈ N.
  Output: YES  if  v  is  accepted,  NO  otherwise.

(ACCEPT)     if α⁻_harm(v) = ∅ then return YES.
(REJECT)     if ∃a ∈ α⁻_harm(v) s.t. nem(a) = ∅ then return NO.
(UNIT RULE)  if ∃a ∈ α⁻_harm(v) s.t. |nem(a)| = 1 then
                          let w ∈ N s.t. nem(a) = {w};
               return  Accept(OF|_{v,w}, v_{v,w}).
(BRANCH)     Candidates ← ⋃_{a∈α⁻_harm(a)} nem(a)
             while Candidates ≠ ∅ do
                 w ←  a node in Candidates
                 if  Accept(OF|_{v,w}, v_{v,w})
                    then return YES
                    else Candidates ← Candidates − {w}.
```

Proposition 100 Accept(OF, v) *returns YES if and only if there is an admissible set S in the* OF *OF such that $v \in S$.*

Proof. The proof follows from Proposition 98 and the discussion before the description of the algorithm. □

To see the advantages of this algorithm over chronological backtracking schemes, we consider the OF associated to the AF in Figure 6.5 (showing also its favorable position over similar algorithms for AFs, e.g. Nofal *et al.* [NAD14]).

Applying Accept(OF, v), the attack a with $s(a) = z$ and $t(a) = v$ is from $\alpha^-(v)$, and since $nem(a) = \{w\}$, the "unit rule" Accept$(OF|_{\{v,w\}}, v_{\{v,w\}})$ is called, which returns YES, that is v is accepted since $\{v, w\}$ is an admissible set. On the other hand, a chronological backtrack search could try any of the 2^n scf sets $\{v\} \cup A$, for $A \subseteq \{u_1, \ldots, u_n\}$, (which are not admissible sets) before the solution $\{v, w\}$ is discovered.

Note that the algorithm described can be easily modified to return the "explanation set" S in case of acceptance.

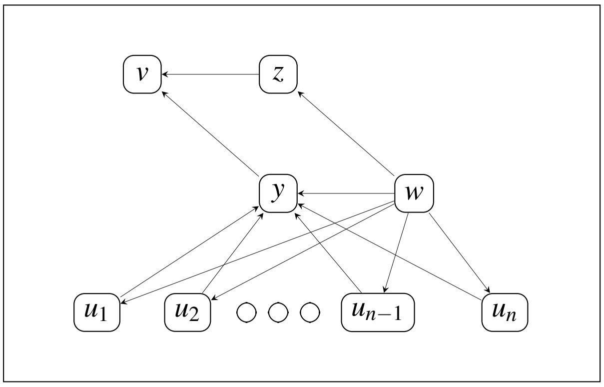


Figure 6.5.: No scf set $\{v\} \cup A$ is admissible, for $A \subseteq \{u_1, \ldots, u_n\}$.

6.2.5. Logical Semantics

In this subsection, we characterize admissible sets in an OF by the models of a formula expressed in propositional logic (for AFs this is done by Besnard and Doutre [BD04]).

More precisely, if $OF = (N, F, g, A, s, t, \gamma, \delta)$ is an OF, then we consider a propositional variable x_v for each $v \in N$. We want to construct a formula Φ over variables $\{x_v | v \in N\}$ such that $S \subseteq N$ is an admissible set in OF if and only if there is a model m of Φ with $S = \{v \in N \,|\, m(x_v) = true\}$.

To characterize the scf sets, let us consider the formula $\text{AtMostOne}(x, y) = \neg(x \wedge y)$. Two nodes v and w belong to the same scf set if and only if there are no facts granted by one and denied by the other: $g(v) \cap \Delta(w) = \emptyset$ and $g(w) \cap \Delta(v) = \emptyset$. Therefore the formula

$$\Phi_1 = \bigwedge_{\substack{v,w \in N, v \neq w \\ g(v) \cap \Delta(w) \neq \emptyset}} \text{AtMostOne}(x_v, x_w)$$

has the property that if m is a model of Φ_1, then $S = \{v \in V | m(x_v) = true\}$ is a scf set, and if S is a scf set then taking $m(x_v) := true$ for $v \in S$ and $m(x_v) := false$ for $v \in N - S$, we obtain a model of Φ_1.

To characterize the admissible sets, let us consider

$$\Phi_2 = \bigwedge_{v \in V} \big(x_v \rightarrow \bigwedge_{a \in \alpha^-(v)} \big(\bigvee_{\substack{u:\, \exists a' \in \alpha^+(u) \text{ s.t.} \\ t(a') = s(a) \,\&\, \delta(a') \cap \gamma(a) \neq \emptyset}} x_u \big)\big).$$

If m is a model of Φ_2 and $S = \{v \in V | m(x_v) = true\}$, then each vertex v in S is either not attacked ($\alpha^-(v) = \emptyset$ and $\bigwedge_{a \in \alpha^-(v)} \big(\ldots\big)$ is $true$), or for each attack a on v at least one fact in the guard of a is denied by an attack from a vertex u in S (that is, $m(x_u) = true$ and $\exists a' \in \alpha^+(u)$ such that $t(a') = s(a)$ and $\delta(a') \cap \gamma(a) \neq \emptyset$). It follows that if m is also a model for Φ_1 then S is a scf set counterattacking each attack against it, that is S is an admissible set. Conversely, if S is an admissible set then it is easy to see that taking $m(x_v) := true$ for $v \in S$ and $m(x_v) := false$ for $v \in N - S$, we obtain a model of $\Phi_1 \wedge \Phi_2$. Hence we have the following proposition.

Proposition 101 *If m is a model of $\Phi = \Phi_1 \wedge \Phi_2$ then $S = \{v \in N | m(x_v) = true\}$ is an admissible set in OF. Conversely, if S is an admissible set in OF then m, given by $m(x_v) := true$ for $v \in S$ and $m(x_v) := false$ for $v \in N - S$, is a model of Φ.*

Proof. Let $S = \{v \in N | m(x_v) = true\}$, for some model m of $\Phi = \Phi_1 \wedge \Phi_2$. If S is not a scf set, then there are $v, w \in S$ such that $\Delta(w) \cap g(v) \neq \emptyset$. Hence $m(x_v) = m(x_w) = true$ and m does not satisfy AtMostOne(x_v, x_w). Since $\Delta(w) \cap g(v) \neq \emptyset$, it follows that AtMostOne(x_v, x_w) occurs in Φ_1. Therefore m does not satisfy Φ_1 and hence m is not a model of Φ, a contradiction. To prove that S is an admissible set, suppose that there is an attack $a \in \alpha^-(S)$ which is not counterattacked by S. If $t(a) = v$ and $s(a) = w$ then $m(x_v) = true$ and $m(x_w) = false$, since S is a scf set. Since a is not counterattacked by S, it follows that for each $u \in N$ such that there is $a' \in \alpha^+(u)$ with $t(a') = w$ and $\delta(a') \cap \gamma(a) \neq \emptyset$, we have $u \notin S$, that is, $m(x_u) = false$. We have obtained that the conjunction in Φ_2 corresponding to x_v evaluates to false under m and hence m is not a model of Φ_2, a contradiction.

Conversely, let S be an admissible set in OF, and m the assignment given by $m(x_v) = true$ if and only if $v \in S$. Then, m is a model of Φ_1 since S is a scf set. Also, m is a model of Φ_2 since for $v \notin S$ the implication $x_v \rightarrow \dots$ evaluates to true and for $v \in S$ the same implication evaluates to true since S counterattacks any attack on v. $\square$

Example. For the OF in Figure 6.4 the above Φ_1, Φ_2 are:
$\Phi_1 = \text{AtMostOne}(x_{v_1}, x_{v_2}) \wedge \text{AtMostOne}(x_{v_2}, x_{v_3}) \wedge \text{AtMostOne}(x_{v_2}, x_{v_4}) \wedge \text{AtMostOne}(x_{v_3}, x_{v_4})$.
Let us write $\Phi_2 = \Phi_{21} \bigwedge \Phi_{22} \bigwedge \Phi_{23} \bigwedge \Phi_{24}$, where $\Phi_{2i} = \left(x_{v_i} \rightarrow \dots\right)$, for $i \in \{1,2,3,4\}$, are the following implications: $\Phi_{21} = x_{v_1} \rightarrow x_{v_3}$, $\Phi_{22} = x_{v_2} \rightarrow false = \neg x_{v_2}$, $\Phi_{23} = x_{v_3} \rightarrow true$, $\Phi_{24} = x_{v_4} \rightarrow true$. Hence, $\Phi_2 = (x_{v_1} \rightarrow x_{v_3}) \wedge \neg x_{v_2}$. It is not difficult to see that the only models of $\Phi_1 \wedge \Phi_2$ are those obtained by setting x_v to true, for $v \in S$, where S is an admissible set of the OF: $\emptyset$, $\{v_3\}$, $\{v_4\}$, $\{v_1, v_3\}$.

The above proposition shows also that it is not possible to simulate an OF with the set of facts F, by considering an argumentation framework $AF = (A, E)$ with $A \subseteq 2^F$. Indeed, as $\Phi = \Phi_1 \wedge \Phi_2$ shows, we have to consider additional constraints to bind the subsets of F (nodes in AF) corresponding to the guards of the attacks issuing from each node in OF. In structured (logical) argumentation this is done by considering an argument together with all its sub-arguments, but this is not practical for most of OFs.

6.3. Related Work

In this chapter we introduced a new generalization of Dung's argumentation framework which is conceptually different from other generalized abstract argumentation frameworks, see Brewka *et al.* [BPW14], or Modgil [Mod13].

The main idea is to keep the abstract level of the original structures, its simplicity and intuitive approach but, at the same time, to increase their modeling capabilities. The "nodes" of our `OF`s have a minimal **content** expressed as finite non-empty sets of facts (the node's position), which are used to relate the "attacks" between two nodes to their positions. This gives a new perspective on the "consistent sets of nodes" which goes beyond the usual conflict-freeness (which is responsible for some rationality violations observed in the instantiation-based argumentation, Caminada and Amgoud [CA07], Amgoud [Amg14]). More precisely, in our strong conflict-free sets we forbid not only the attacks between their members but also require that the outside attacks are not in contradiction with their members positions. Unfortunately, this does not prevent that (when our `OF`s are used as instantiation destination of logical structured argumentation frameworks) nodes with mutually consistent positions to be globally contradictory. This happens because the attack relation is binary.

We introduced a simple recursive definition of acceptance: a node (the position expressed by a node) is accepted in an `OF` if either it can counterattack all attacks targeting it or there is another "compatible" node such that in the `OF` obtained by "contracting" these two nodes in a single "supernode" this supernode is accepted. Note that this type of acceptance is different from that considered in *abstract dialectical frameworks* Brewka and Woltran [BW10], or *GRaph-based Argument Processing with Patterns of Acceptance* [Brewka and Woltran [BW14], where the acceptance of a node is a function defined on the set of its parents (that is the nodes having a directed edge to it). Also our approach is conceptually different from proof procedures, see Modgil and Caminada [MC09]. Technically, using "guarded attacks" and a suitable graph operation of contraction, we proved that this type of acceptance is compatible in the particular case of `AF`s with Dung's admissibility-based semantics, showing that it is actually a proper generalization of the Dung semantics. Hence, if we use `OF`s instead of `AF`s as a target system to evaluate arguments in structured argumentation frameworks, then we obtain the same results if the OF_{AF} (see Definition 92) is considered for reusability reasons. However, a more fine-grained generation of the target `OF` – by explicitly devising rules of attacks (via their guards), which are non-uniform (depend on the source/target node) – may be used to obtain improved modeling. We note also that in our guarded attack calculus, the attacks on the attacks Villata *et al.* [VBvdT11]) are implicitly considered. The use of the set of attacks ($\alpha^{-}(v)$ in `OF`s) instead of the set of parents (v^{-} in `AF`s or `ADF`s) in the study of the acceptability of a node v simplifies the description of acceptability algorithms. A novel DPLL type backtracking acceptance algorithm is described.

Some improvements can be further obtained by using *abstract* DPLL *with learning*, Nieuwenhuis *et al.* [NOT06] (see also Brochenin *et al.* [BLWW15]). The idea is to add dummy facts to nodes in order to learn that some (set of) nodes are not useful in finding a successful coalition for the acceptance of a given node v.

The characterization of admissible sets in an OF by the models of a formula expressed in propositional logic shows a lazy way to map an OF to an AF (via naive transformations of this formula) in a manner preserving semantic properties of the first one. The study of the efficiency of such a mapping (similar to that initiated by Brewka *et al.* [BW10] for ADFs) is an interesting future research direction.
AFs have been generalized (see, e.g., Bench-Capon [BC03], Bourguet *et al.* [BAT10], Dunne *et al.* [DHM$^+$11]) by adding weights to arguments or attacks in order to increase their modeling capacity. We can consider a similar extension for OFs, by providing the facts and attacks with weights, called *vitality* for facts, and *strength* for attacks. The facts effectively denied by an attack a are those facts in $\delta(a)$ having a vitality smaller than the strength of the attack a. In this way, an weighted OF could be used to represent families of OFs.

Concluding Discussion

A summary of my specific contributions rendered in this thesis is given below.

- Properties of the equivalent classes of *attacked-related* and *attacking-related* relations expressed in Proposition 9.
- The two-way scanning algorithm $QK(D,\pi)$ devised in Section 0.4 used to construct three quasi-kernels in a digraph without kernels (see Figure 0.7).
- Proving that deciding if there exists a quasi-kernel in digraph containing a specified vertex v is a NP-complete problem, in Section 0.4 (see Figure 0.8).
- Proving in a formal way that a discipline policy can be adopted in forming of an argumentation framework, without changing the semantic properties, in Section 2.3 (Theorem 32 and Figure 2.4).
- The set of integral vectors of the polyhedron $P_{adm}(AF)$ associated to the argumentation framework AF (Definition 38) is exactly the set of characteristic vectors of the admissible sets of AF, in Section 3.4 (Theorem 42).
- The set of integral vectors of the polyhedron $P_{stab}(AF)$ associated to the argumentation framework AF (Definition 43) is exactly the set of characteristic vectors of the stable extensions of AF, in Section 3.4 (Theorem 46).
- Using the polyhedron $P_{stab}(AF)$ to obtain the set of characteristic vectors of the complete extensions of AF, in Section 3.4 (Theorem 49).
- Introducing bipartite *debates* (Definition 50), *consensual debates* (Definition 51) and debate operations, in Section 4.2, implying an unusual definition of *aggregators* (Definition 52), in Section 4.3.
- Characterization of the *majority rule* (Theorem 53), in Section 4.3.
- Characterization of the *approval & disapproval rule* (Theorem 54), in Section 4.3.
- Introducing the *conceptual aggregation correspondence* (Definition 58), in Section 4.4.

- Introducing the argumentation frameworks associated to a debate (Definition 60), the debate corresponding to an argumentation framework (Definition 59) and proving that the two debates associated to the debate corresponding to an argumentation framework are isomorphic (Theorem 61), in Section 4.5.

- Introducing different type of coalitions in a debate: *legal*, *autarky*, *strong autarky*, *oligarchy*, corresponding to conflict-free sets, admissible sets, complete extensions, stable extensions in the (opinions)-argumentation framework associated to the debate (Definitions 64, 65, and 66) and characterizing them using the operation of contraction (Theorem 69), in Section 4.5.

- Introducing *compromises* (Definition 71) and the *argumentative aggregation correspondence* (Definition 73), in Section 4.5.

- Proving that argumentative aggregation operators does not satisfy *Independence* (Theorem 74), in Section 4.5.

- Reducing the study of the existence of a *Stable Choice Matching* (Definitions 76, 77) in a *Bipartite Choice System* (Definition 75) to the study of the existence of **Stable Common Fixed Points** (Definition 78) of two choice functions, in Section 5.2. Note that this approach is in principle different from the existing fixed point approaches (using Tarski's fixed point theorem, e.g., [Fle03] or [HM05]).

- Proving a new characterization of *Path Independence* choice functions by showing that a substitutable choice function satisfies the *Path Independence* condition if and only if it satisfies **LCR** condition (Theorem 87), expressing the external stability of $f(X)$ in any $X \subseteq U$, in Section 5.3.

- Devising a generalized Gale-Shapley algorithm – called *Immediate Rejection* – which returns a stable fixed point of two functions satisfying the *Conservative Rejection* condition (Theorem 89), in Section 5.4. The name is motivated by the rule that once an element (corresponding to a contract in many-to-many two-sided markets), is selected by the first function and rejected by the second function, it is not considered further.

- Proving that the **LCR** condition is necessary (in the maximal domain sense) for the correctitude of the *Immediate Rejection* algorithm (Theorem 90), in Section 5.4.

- Introducing *Opposition Frameworks* (Definition 91), in Subsection 6.2.1, as a labeled multi-digraph in which each directed edge (v,w) corresponds to an attack a from the node $s(a) = v$ to node $t(a) = w$, that is based on a set of facts $\gamma(a) \subseteq g(v)$ granted by v, and denies the set $\delta(a) \subseteq g(w)$ of facts granted by w.

- Introducing *strongly conflict-free sets*, in Subsection 6.2.2, which can be used safely in the graph operation of *contraction* (Proposition 94).

- Approaching the (credulously) acceptance using the graph operation of *contraction* (Proposition 98) in Subsection 6.2.3.
- Devising a DPLL type backtracking acceptance algorithm in Subsection 6.2.4, reducing the search time as explained in Figure 6.5.
- Characterization of admissible sets in an opposition framework by the models of a formula expressed in propositional logic, (Proposition 101) in Subsection 6.2.4.

An obvious future work (necessary for the validation of our various proposals) is to implement real world argumentative decision-making systems and to do systematic experimental evaluations. The principal difficulty is that the problems arising in the field of argumentation are computationally hard. Hence, any implementation must face the problem of reducing the number of sophisticated NP-oracles calls such as SAT, ASP or CSP solvers.

Another compelling future direction is to use the (path independent) choice functions in order to provides the node's granted sets in the opposition frameworks with a (combinatorial) structure and to use the intuitive rejection associated with choice functions to construct the guards of the attacks.

Bibliography

[AA95] M. Aizerman and F. Aleskerov. *Theory of Choice*. Elsevier, 1995.

[AEG+11] S. Airiau, U. Endriss, U. Grandi, D. Porello, and J. Uckelman. Aggregating dependency graphs into voting agendas in multi-issue elections. In *Proc. of IJCAI 2011*, pages 16–22, 2011.

[AM81] M. A. Aizerman and A. V. Malishevski. General theory of best variants choice: Some aspects. *IEEE Transactions on Automatic Control*, pages 1030–1040, 1981.

[Amg14] L. Amgoud. Postulates for logic-based argumentation systems. *International Journal of Approximate Reasoning*, pages 2028–2048, 2014.

[Arr50] K. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, pages 328–346, 1950.

[Arr63] K. Arrow. *Social Choice and Individual Values*. Wiley, 1963.

[AS12] O. Aygün and T. Sönmez. Matching with contracts: The critical role of irrelevance of rejected contracts. Technical report, Boston College Working Paper, 2012.

[AS14] L. de. Alfaro and M. Shavlovsky. CrowdGrader: a tool for crowdsourcing the evaluation of homework assignments. In *SIGCSE 2014*, pages 415–420, 2014.

[ASS02] K.J. Arrow, A.K. Sen, and K. Suzumura. *Handbook of Social Choice and Welfare*. Elsevier, 2002.

[BAT10] J.-R. Bourguet, L. Amgoud, and R. Thomopoulos. Towards a unified model of preference-based argumentation. In *Proc. of FoIKS 2010*, pages 326–344, 2010.

[BC03] T. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13:429–448, 2003.

[BD04] P. Besnard and S. Doutre. Checking the acceptability of a set of arguments. In *Proc. of NMR*, pages 59–64, 2004.

[BF78] S. J. Brams and P. C. Fishburn. Approval voting. *American Political Science Review*, 72:831–847, 1978.

[BG03] P. Baroni and M. Giacomin. Solving semantic problems with odd-length cycles in argumentation. In *Proc. of ECSQARU 2003*, volume 2711 of *LNAI*, pages 440–451, 2003.

[BG07a] P. Baroni and M. Giacomin. Comparing Argumentation Semantics with Respect to Skepticism. In *Proc. of ECSQARU 2007*, pages 210–221, 2007.

[BG07b] P. Baroni and M. Giacomin. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence*, 171:675–700, 2007.

[BGG05] P. Baroni, M. Giacomin, and G. Guida. Scc-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence*, 168:165–210, 2005.

[BGW12] M. Bezem, C. Grabmayer, and M. Walicki. Expressive power of digraph solvability. *Ann. Pure Appl. Logic*, 163:200–213, 2012.

[BH11] F. Brandt and P. Harrenstein. Set-retionalizable choice functions and self- stability. *Journal of Economic Theory*, 146:233–273, 2011.

[Bla88] C. Blair. The lattice structure of the set of stable matchings with multiple partners. *Mathematics*, 13:619–628, 1988.

[BLWW15] R. Brochenin, B. Linsbichler, J. P. Wallner, and S. Woltran. Abstract Solvers for Dung's Argumentation Frameworks. In *Proc. of TAFA 2015*, 2015.

[BM04] A. Bogomolnaia and H. Moulin. Random matching under dichotomous preferences. *Econometrica*, 72:257–279, 2004.

[BMS05] A. Bogomolnaia, H. Moulin, and R. Stong. Collective choice under dichotomous preferences. *Journal of Economic Theory*, 122:165–184, 2005.

[BNS94] C. Bell, R. T. Nerode, A. Ng, and V. S. Subrahmanian. Mixed integer programming methods for computing nonmonotonic deductive databases. *Journal of ACM*, 41:1178–1215, 1994.

[Bon03] J.A. Bondy. Short proofs of classical theorems. *Graph Theory*, 44:159–165, 2003.

[BPW14] G. Brewka, S. Polberg, and S. Woltran. Generalizations of Dung Frameworks and Their Role in Formal Argumentation. *IEEE Intelligent Systems*, 2014.

[BW10] G. Brewka and S. Woltran. Abstract Dialectical Frameworks. In *Proc. of KR 2010*, pages 102–111, 2010.

[BW14] G. Brewka and S. Woltran. GRAPPA: A semantical framework for graph-based argument processing. In *Proc. of ECAI 2014*, pages 153–158, 2014.

[CA07] M. Caminada and L. Amgoud. On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171:286–310, 2007.

[Cam06a] M. Caminada. On the issue of reinstatement in argumentation. In *Proc. of Logics in Artificial Intelligence, 10th European Conference, JELIA, LNCS 4160*, pages 112–123. Springer, 2006.

[Cam06b] M. Caminada. Semi-stable semantics. In *Proc. of COMMA 2006*, pages 121–130, 2006.

[CDLSM02] C. Cayrol, S. Doutre, M.-C. Lagasquie-Schiex, and J. Mengin. "Minimal defence": a refinement of the preferred semantics for argumentation frameworks. In *Proc. of NMR 2002*, pages 408–415, 2002.

[Chv73] V. Chvátal. On the computational complexity of finding a kernel. Technical report, Centre de Recherches Mathématiques, Université de Montréal,, 1973.

[CK12] C. Croitoru and T. Kötzing. Deliberative acceptability of arguments. In *Proc. of STAIRS 2012*, pages 71–82, 2012.

[CK13] C. Croitoru and T. Kötzing. A Normal Form for Argumentation Frameworks. In *TAFA 2013*, pages 32–45, 2013.

[CL74] V. Chvátal and L. Lovász. Every directed graph has a semi-kernel. In *Hypergraph Seminar, Lecture Notes in Mathematics*, volume 411, page 175, 1974.

[CM] C. Croitoru and K. Mehlhorn. Bipartite choice systems - to appear.

[CM16] C. Croitoru and K. Mehlhorn. Opposition frameworks. In *JELIA 2016*, pages 190–206, 2016.

[CMDK+07] S. Coste-Marquis, C. Devred, S. Konieczny, M.-C. Lagasquie-Schiex, and P. Marquis. On the merging of Dung's argumentation systems. *Journal of Artificial Intelligence*, 171:730–753, 2007.

[CMDM05a] S. Coste-Marquis, C. Devred, and P. Marquis. Prudent semantics for argumentation frameworks. In *Proc. of ICTAI*, pages 568–572, 2005.

[CMDM05b] S. Coste-Marquis, C. Devred, and P. Marquis. Symmetric argumentation frameworks. In *Proc. of ECSQARU 2005*, pages 317–328, 2005.

[Coh68] J. E. Cohen. Interval graphs and food webs: a finding and a problem. Technical report, RAND Corporation, 1968.

[CP11] M. Caminada and G. Pigozzi. On judgment aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems*, 22:64–102, 2011.

[Cre95] N. Creignou. The class of problems that are linearly equivalent to satisfiability or a uniform method for proving NP-completeness. *Theoretical Computer Science*, 145:111–145, 1995.

[Cro12] C. Croitoru. Algorithmic aspects of abstract argumentation frameworks. Master's thesis, Saarland University, 2012.

[Cro13] C. Croitoru. Abstract debates. In *Proc. of ICTAI 2013*, pages 713–718, 2013.

[Cro14a] C. Croitoru. Argumentative aggregation of individual opinions. In *JELIA 2014*, pages 600–608, 2014.

[Cro14b] C. Croitoru. Polyhedral labellings for argumentation frameworks. In *SUM*, pages 86–99, 2014.

[Cro15a] C. Croitoru. Bipartite digraph debates. In *Mpref 2015*, 2015.

[Cro15b] C. Croitoru. A note on quasi-kernels in digraphs. *Inf. Process. Lett.*, pages 863–865, 2015.

[DBC01] P. E. Dunne and T. Bench-Capon. Coherence in finite argument systems. Technical report, University of Liverpool, 2001.

[DBC02] P. E. Dunne and T. Bench-Capon. Coherence in finite argument systems. *Artificial Intelligence*, 141:187–203, 2002.

[DHM$^+$11] P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175:457–486, 2011.

[DLL62] M. Davis, G. Logemann, and D. V. Loveland. A machine program for theorem-proving. *Commun. ACM*, pages 394–397, 1962.

[DMP97] Y. Dimopoulos, V. Magirou, and C. H. Papadimitriou. On kernels, defaults and even graphs. *Annals of Mathematics and Artificial Intelligence*, 20:1–12, 1997.

[DMT06] P. M. Dung, P. Mancarella, and F. Toni. A dialectic procedure for sceptical assumption-based argumentation. In *Proc. of COMMA 2006*, pages 145–146, 2006.

[DMW12] P. E. Dunne, P. Marquis, and M. Wooldridge. Argument aggregation: Basic axioms and complexity results. In *Proc. of COMMA 2012*, pages 129–140, 2012.

[DT96] Y. Dimopoulos and A. Torres. Graph theoretical structures in logic programs and default theories. *Theoret. Comput. Sci.*, 170:209–244, 1996.

[Duc80] P. Duchet. Graphes noyau-parfaits. *Annals of Discrete Mathematics*, 9:93–101, 1980.

[Dug12] J. Duggan. Uncovered sets. *Social Choice and Wellfare*, pages 489–453, 2012.

[Dun95] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[Dun07] P. E. Dunne. Computational properties of argument systems satisfying graph-theoretic constraints. *Artificial Intelligence*, 171:701–729, 2007.

[EG12] U. Endriss and U. Grandi. Graph aggregation. In *Proc. of COMSOC-2012*, 2012.

[FDR15] C. Finegan-Dollak and D. R. Radev. Sentence simplification, compression, and disaggregation for summarization of sophisticated documents. *Journal of the Association for Information Science and Technology*, pages 2437–2453, 2015.

[Fle03] T. Fleiner. A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research.*, 28:103–126, 2003.

[Fra97] A. Fraenkel. Combinatorial game theory foundations applied to digraph kernels. *The Electronic Journal of Combinatorics*, 4:100–117, 1997.

[Gab11] D. Gabbay. Introducing equational semantics for argumentation networks. In W. Liu, editor, *Proc. of ECSQARU 2011*, number 6717 in LNAI, pages 19–35. Springer-Verlag, 2011.

[Gab12a] D. Gabbay. An equational approach to argumentation networks. *Argumentation and Computation*, 3:87–142, 2012.

[Gab12b] D. Gabbay. The equational approach to cf2 semantics. In *Proc. of COMMA 2012*, pages 141–152, 2012.

[GF11] C. Gratie and M. Florea. Fuzzy labelling for argumentation frameworks. In *Proc. of ArgMas2011*, pages 1–8, 2011.

[GKTY04] G. Gutin, K.M Koh, E.G. Tay, and A. Yeo. On the number of quasi-kernels in digraphs. *Jounal of Graph Theory*, 46:48–56, 2004.

[GS62] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69:9–15, 1962.

[HG11] A. Hunter and N. Gorogiannis. Instantiating abstract argumentation with classical logic arguments: Postulates and properties. *Artificial Intelligence*, 175:1479–1497, 2011.

[HK16] J. W. Hatfield and S. D. Kominers. Contract design and stability in many-to-many matching. *Games and Economic Behavior*, pages 78–97, 2016.

[HM05] J. W. Hatfield and P. Milgrom. Matchings with contracts. *American Economic Review*, pages 913–935, 2005.

[JM96] H. Jacob and H. Meyniel. About quasi-kernels in a digraph. *Discrete Mathematics*, 154:279–280, 1996.

[KC82] A. Kelso and V. Crawford. Job matching, coalition formation, and Gross substitutes. *Econometrica*, 50:1483–1504, 1982.

[Kha79] L. G. Khachiyan. A polynomial algorithm in linear programmming. *Soviet Mathematics Doklady*, 20:191—194, 1979.

[KP02] S. Konieczny and R. P. Pérez. Merging information under constraints: A logical framework. *Journal of Logic and Computation*, 12:773–808, 2002.

[LJN12] G. Liu, T. Janhunen, and I. Niemelä. Answer set programming via mixed integer programming. In *Proc. of KR 2012*, pages 32–42, 2012.

[LM11] J. Leite and J. Martins. Social abstract argumentation. In *Proc. of IJCAI 2011*, pages 2287–2292, 2011.

[Low12] J. Lowe. *CPLEX 10 Solver Manual*. GAMS Development Corporation, 2012.

[LP09] C. List and C. Puppe. *Judgment aggregation: A survey*, chapter 19, pages 158–190. Oxford University Press, 2009.

[LS10] J.-F. Laslier and M. R. Sanver, editors. *Handbook on Approval Voting*. Springer, 2010. Surveys all major developments in Approval Voting since the publication of the seminal book by Brams/Fishburn (1983).

[LW98] X. Liu and D. B. West. Line digraphs and coreflexive vertex sets. *Discrete Math.*, 188:269–277, 1998.

[LZD04] R. Li, D. Zhou, and D. Du. Satisfiability and integer programming as complementary tools. In *Proc. of ASPDAC 2004*, pages 879–882, 2004.

[Mas95] E.S. Maskin. Majority rule, social welfare functions, and game forms. In K. Basu, P.K. Pattanaik, and L. Suzumura, editors, *Choice, Welfare, and Development*, pages 100–109. The Clarendon Press, 1995.

[May52] K.O. May. A set of independent, necessary and sufficient conditions for simple majority decisions. *Econometrica*, 20:680–684, 1952.

[MC09] S. Modgil and M. Caminada. *Proof Theories and Algorithms for Abstract Argumentation Frameworks*, chapter 6, pages 105–129. Springer, 2009.

[Mir04] A. Miroiu. Characterizing majority rule: from profiles to societies. *Economics Letters*, 85:359–363, 2004.

[MM65] J. W. Moon and L. Moser. On cliques in graphs. *Israel Journal of Mathematics*, 3:23–28, 1965.

[Mod13] S. Modgil. Revisiting abstract argumentation frameworks. In *Proc. of TAFA 2013*, pages 1–13, 2013.

[Mou85] P. Moulin. Choice functions over a finite set: a summary. *Social Choice and Welfare*, 2:147–160, 1985.

[NAD14] S. Nofal, K. Atkinson, and P. E. Dunne. Algorithms for decision problems in argument systems under preferred semantics. *Artificial Intelligence*, 207:23–51, 2014.

[NL71] V. Neumann-Lara. Seminúcleos de una digráfica. *Anales del Instituto de Matemáticas*, 2, 1971.

[NM44] von J. Neumann and O. Morgenstern. *Theory of Games and Economic Behaviour*. Princeton University Press, 1944.

[NOT06] R. Nieuwenhuis, A. Oliveras, and C. Tinelli. Solving SAT and SAT Modulo Theories: From an abstract Davis–Putnam–Logemann–Loveland procedure to DPLL(T). *ACM*, pages 937–977, 2006.

[NP06] S. Nielson and S. Parsons. A Generalization of Dung's Abstract Framework for Argumentation: Arguing with Sets of Attacking Arguments. In *Proc. of ArgMAS 2006*, pages 54–73, 2006.

[ON08] N. Oren and T. J. Norman. Semantics for evidence-based argumentation. In *Proc. of COMMA 2008*, pages 276–284, 2008.

[PE11] D. Porello and U. Endriss. Ontology merging as social choice. In *Proc. of CLIMA 2011*, pages 157–170, 2011.

[Pig06] G. Pigozzi. Belief merging and the discursive dilemma: An argument-based account to paradoxes of judgment aggregation. *Synthese*, 152:285–298, 2006.

[Plo73] C.R. Plott. Path independence, rationality, and social choice. *Econometrica*, 41:1075–1091, 1973.

[Pol95] J. L. Pollock. *Cognitive Carpentry. A Blueprint for How to Build a Person*. MIT Press, Cambridge MA, 1995.

[Pra10] H. Prakken. An abstract framework for argumentation with structured arguments. *Argumentation and Computation*, 1:93–124, 2010.

[PS96] H. Prakken and G. Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4:331–368, 1996.

[PV02] H. Prakken and G. Vreeswijk. Logics for defeasible argumentation. In *Handbook of Philosophical Logic, 2nd Edition*, volume 4, pages 218–319. Kluwer, 2002.

[Ric53] M. Richardson. Solutions of irreflexible relations. *Annals of Math.*, 53:573–590, 1953.

[RLT09] I. Rahwan, K. Larson, and F Tohme. A characterisation of strategy-proofness for grounded argumentation semantics. In *Proc. of IJCAI 2009*, pages 251–256, 2009.

[Ros70] D. J. Rose. Triangulated graphs and the elimination process. *Journal of Mathematical Analysis and Applications*, 32:597–609, 1970.

[Rot84] A. Roth. Stability and polarization of interests in job matching. *Econometrica*, 52:47–57, 1984.

[RS90] A. Roth and M. Sotomayor. *Two-sided matching: A study in game-theoretic modelling and analysis*. Cambridge Univ. Press, 1990.

[RT10] I. Rahwan and F. Tohme. Collective argument evaluation as judgement aggregation. In *Proc. of AAMAS 2010*, pages 417–424, 2010.

[RZR07] I. Rahwan, F. Zablith, and C. Reed. Laying the foundations for a world wide argument web. *Artificial Intelligence*, 171:897–921, 2007.

[Sch03] A. Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency*. Springer, 2003.

[Sen71] A. K. Sen. Choice functions and revealed preference. *The Review of Economic Studies*, pages 307–317, 1971.

[SM73] L. Stockmeyer and A. Meyer. Word problems requiring exponential time. In *Proc. of STOC 1973*, pages 1–9, 1973.

[TBS08] F. Tohme, G. A. Bodanza, and G. R. Simari. Aggregation of attack relations: A social-choice theoretical analysis of defeasibility criteria. In *FoIKS*, pages 8–23, 2008.

[TY84] R. E. Tarjan and M. Yannakakis. Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs. *SIAM Journal of Computing*, 13:566–579, 1984.

[VBvdT11] S. Villata, G. Boella, and L. van der Torre. Attack Semantics for Abstract Argumentation. In *Proc. of IJCAI 2011*, pages 406–413, 2011.

[Vor07] M. Vorsatz. Approval voting on dichotomous preferences. *Social Choice and Welfare*, 28:127–141, 2007.

[Wal14] T. Walsh. The PeerRank method for peer assessment. In *ECAI 2014*, pages 909–914, 2014.

[WD12] M. Walicki and S. Dyrkolbotn. Finding kernels or solving SAT. *Journal of Discrete Algorithms*, 10:146–164, 2012.

[Wil82] R. Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. In *I. Rival (Ed.), Ordered sets*, pages 445–470, 1982.

[WL10] B. Wang and G. Luo. Extend argumentation frameworks based on degree of attack. In *Proc. of 9th IEEE International Conference on Cognitive Informatics, ICCI 2010*, pages 771–776, 2010.

[Woe03] G.J. Woeginger. A new characterization of majority rule. *Economics Letters*, 81:89–94, 2003.

[Wra76] C. Wrathall. Complete sets and the polynomial-time hierarchy. *Theoretical Computer Science*, 3:23–33, 1976.

Cosmina Croitoru - CV

Personal Details

Birthdate:	06 March 1988
Birthplace:	Iasi, Romania
Citizenship:	Romanian
Gender:	Female
Email:	cosmina.croitoru@gmail.com

Education

2012–today	*PhD Student in Computer Science* at IMPRS-CS and MPII, Saarbrücken, Supervisor: *Kurt Mehlhorn*, Expected Defense: July 2017
2010–2012	*MsC in Computer Science* at Saarland University, Saarbrücken, *1.5/1*
2009–2010	*Erasmus Exchange Student* at Uni Konstanz, Germany, *ECS: 270/300*
2007–2010	*Bachelor in Computer Science* at Al. I. Cuza University Iasi, Faculty of Computer Science, *9.66/10*
1999–2007	*Highschool*, National College, Iasi, Romania

Master thesis

Title:	*Algorithmic Aspects of Argumentation Frameworks*
Supervisors:	*Prof. Dr. Kurt Mehlhorn, Dr. Timo Kötzing*

Teaching Skills

SS 2017	Teaching Assistant, *Artificial Intelligence* (in English)
SS 2016	Teaching Assistant, *Artificial Intelligence* (in English)
SS 2015	Teaching Assistant, *Artificial Intelligence* (in English)
WS 2014	Teaching Assistant, *Theory of Distributed Systems* (in English)
WS 2011	Teaching Assistant, *Algorithms and Data Structures* (in German)
WS 2011	Teaching Assistant, *Computational Thinking* (in German)

Scholarships

2012–2015	IMPRS scholarship for PhD Studies, MPII, Saarbrücken, Germany
2010–2012	IMPRS scholarship for Master, MPII, Saarbrücken, Germany
2009–2010	Erasmus scholarship, Uni Konstanz, Germany
2007–2010	FII merit scholarship, Faculty of Computer Science Iasi, Romania

Languages

Romanian	Native language
English	Fluent
German	Fluent
French	Fair

Publications

2012	C. Croitoru and T. Kötzing. Deliberative acceptability of arguments, In Proc. of STAIRS 2012, pages 71-82, 2012
2013	C. Croitoru and T. Kötzing. A Normal Form for Argumentation Frameworks. In TAFA 2013, pages 32-45, 2013
2013	C. Croitoru. Abstract debates. In Proc. of ICTAI 2013, pages 713-718, 2013
2014	C. Croitoru. Argumentative aggregation of individual opinions. In JELIA 2014, pages 600-608, 2014
2014	C. Croitoru. Polyhedral labellings for argumentation frameworks. In SUM, pages 86-99, 2014
2015	C. Croitoru. Bipartite digraph debates. In Mpref 2015, 2015
2015	C. Croitoru. A note on quasi-kernels in digraphs. Inf. Process. Lett., 863-865, 2015
2016	C. Croitoru and K. Mehlhorn. Opposition frameworks. In JELIA 2016, pages 190-206, 2016