Hybrid Logics for Arguments, Beliefs, and their Dynamics

MSc Thesis *(Afstudeerscriptie)*

written by

Max Rapp
(born March 28th, 1989 in Bayreuth, Germany)

under the supervision of **Prof Dr Sonja Smets** and **Dr Fernando Velázquez-Quesada**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
September 19th, 2018	Dr Floris Roelofsen
	Dr Ronald de Haan
	Prof Dr Davide Grossi
	Dr Luca Incurvati

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

We study abstract argumentation, argument based belief and their dynamics in the setting of Hybrid logic. For this purpose, we develop Hybrid Argumentation Logic (HAL), a logic optimized to express the concepts of abstract argumentation theory. Combining HAL, a theory of support between arguments and propositions, and a logic for the claims of arguments in a product logic, we study three notions of argument based belief that reflect ways in which agents might form beliefs from arguments: credulous belief, skeptical belief and strong belief. In addition, two update operations for abstract argumentation frameworks are introduced and studied: argumentation framework union and argumentation framework intersection. These serve as semantics for two dynamic modal operators for HAL, yielding the logic $HAL_{\cup}^{\cap}$. It is shown that $HAL_{\cup}^{\cap}$ is at least as expressive as first-order logic and hence undecidable. We then study how argumentation framework union affects credulous, skeptical and strong belief respectively and find that they have different dynamic properties. In addition we provide sufficient conditions for preservation of belief under argumentation framework union.

Acknowledgements

The completion of this Master thesis marks the end of a long path that represents much more to me than the thesis itself: academically, the end of my two years at the ILLC, the end of my time as student; personally, the end of two years of long-distance relationship , the end of the intellectually most inspiring time of my life, the achievement of a goal for which I had been fighting much longer than those two years. For this I am indebted to a great many people of whom I can only mention the most important ones here.

On the academic level, first and foremost I would like to thank my supervisors Fernando Velázquez-Quesada and Sonja Smets for supporting me with their advice, feedback and criticism, and taking the time to read my incomprehensible drafts - if necessary late at night.

Beyond that I would like to thank the amazing teaching staff at the ILLC for making the Master of Logic the most exciting, challenging, thought-inspiring, engaging program I have ever experienced. Likewise, I would like to thank my fellow students who provided the environment and a community needed to survive this program, be it late at night in the MOL-room or be it during countless foosball and coffee breaks.

On the personal side, I wish to thank my parents and siblings who have supported me through all this years as an “ewiger Student”, never losing faith that my journey would go somewhere. I could rely on them during times when I was unable to rely on myself.

Above all, I would like to thank my fiancé Anne who was willing to let me go away to follow a dream; Who made the sacrifice to let me stay away longer than planned; and who stuck with me in spite of the toll the distance put on our relationship.

Contents

Chapter 1

Introduction

This work is about how people change their opinions. An opinion for our purposes is a belief that is backed by arguments. People change their opinions through a process involving the creation, exchange, modification or retraction of arguments. But when does a new argument prompt an agent to change her opinion? When does a modification of an existing argument do the trick? Which kinds of opinions are easy to change, which are "safe"? Which arguments are fundamental to an agent's world view? In the present work these questions will be investigated using tools from epistemic logic and abstract argumentation theory. In epistemic logic, the study of justified belief has recently seen increased interest (van Benthem and Pacuit [2011],Baltag, Renne, and Smets [2014],Baltag, Bezhanishvili, Özgün, and Smets [2016]). In addition there is a strong research tradition on the dynamics of belief change (see Baltag and Renne [2016]). Abstract argumentation theory Dung [1995] deals with modelling the structure of argumentations and provides the tools needed to decide which arguments should be accepted in a debate. Putting these elements together this work develops a formal framework to define and study notions of belief based on argumentation and their behaviour when the underlying argumentation changes.

Example 1. Alice is worried about the threat sea level rise might pose to her home, the island nation of Atlantis. Today she encountered the following on the news:

- a_1: A report from the Atlantian Scientific Society predicting the complete inundation of Atlantis City due to sea level rise.
- a_2: A "bubble" by the Atlantian strategos stating: "The concept of sea level rise was created by and for the Athenians in order to make Atlantian manufacturing non-competitive".
- a_3: A new study shows that bloodletting cures the cold.

Suppose Alice has about the same degree of trust in the Atlantian Scientific society and the strategos. Should she believe sea level rise is real? Should she reserve judgement and seek out further information? And what about the seemingly unrelated stuff about bloodletting?

The goal of this work is to develop formal tools that allow Alice and her fellow Atlantians to form beliefs based on the arguments they are presented with in situations like Example 1 and to investigate how these beliefs change when new arguments are received. We structure our task as follows: in Chapter 2 we will lay down the basic notions of abstract argumentation theory needed for our investigation. In Chapter 3, we introduce two logical systems for describing abstract argumentation frameworks, one based on normal modal logic (K_U), one based on hybrid logic (HAL). Next, Chapter 4 develops a logic for the claims of arguments, a notion of support that connects arguments to their claims and conjoins these elements with HAL to create a logic for argument-based belief. In Chapter 5, we will investigate the dynamics of argumentation frameworks in the setting of our logical framework. Chapter 6 studies how dynamics such as adding and removing arguments affect the notions of belief defined in Chapter 4 and gives some answers to the questions posed at the beginning of this introduction. Finally, Chapter 7 touches on some of the great many questions that this work has to leave open and concludes by revealing the final fate of Atlantis.

Chapter 2

Basics of Abstract Argumentation Theory

2.1 Basic Notions

It is difficult to pin down what exactly constitutes an argument. The logician might be tempted to view an argument as a sort of informal and incomplete description of a proof from certain axioms. But many, probably most, arguments that people use to justify their opinions would fail this test: the arguments that feature in the justification of people's opinion do so not on merit of validity but by being convincing. What constitues a convincing argument is very much in the eye of the beholder: just ask people during an election campaign which candidate won the latest televised debate.

Given this difficulty one might give up hope to understand argumentation and opinion change, but this would be premature. Besides their "internal" properties - such as the mode of inference and the rhetorical strategies they employ, the evidence they appeal to or the assumptions they make - arguments also have "external" properties. Most important among these are the relation in which an argument stands to other arguments and the relation in which it stands to the proposition that is up for debate. The latter will be discussed in Chapter 4 of this work. The first is studied in the field of abstract argumentation theory discussed in this section. We base our discussion on the seminal paper by Dung 1995.

The basic idea of abstract argumentation theory is to investigate what can be said about argumentation based only on how arguments attack each other. Thus arguments are conceived of as atoms that stand in an attack relation. The resulting structure is captured as a directed graph where the nodes are arguments and the edges represent "attacks" between arguments.

Definition 1 (Argumentation framework)**.** An argumentation framework (or relational structure) $\mathcal{A}$ is a tuple $\langle A, \twoheadrightarrow \rangle$ where A is a non-empty set and $\twoheadrightarrow \subseteq A \times A$. For $X \subseteq A$, $a, b \in A$ we write $a \twoheadrightarrow b$ iff $(a, b) \in \twoheadrightarrow$ and $X \twoheadrightarrow b$ if there is

$\mathcal{A}_1$
a_1
a_2
a_3

Figure 2.1: Alice's Argumentation Framework

$a \in X$ with $a \twoheadrightarrow b$. We say that a (X) *attacks* b iff $a \twoheadrightarrow b$, ($X \twoheadrightarrow b$). Mirroring this, a (X) attacks $Y \subseteq A$ iff there is some $b \in Y$ such that a (X) attacks b.

Example 2. How can we represent the scenario from example 1 as an argumentation framework? In general, extracting argumentation frameworks from natural language is more of an art than a science but a possible interpretation might be the following:

$$\mathcal{A}_1 = (A = \{a_1, a_2, a_3\}, \twoheadrightarrow = \{(a_1, a_2), (a_2, a_1)\}\}$$

We will frequently graphically represent such argumentation framework in a table as seen in Figure 2

Throughout this thesis, we will work with *finite* argumentation frameworks unless otherwise specified. This assumption not only makes our life easier in many ways, it can also be justified well in the context of argumentation: for one argumentation is in large part an application driven field and infinite structures are the exception in applications in AI [1] and do not occur "in the wild". In addition, presently we seek to model the belief state of an agent in terms of arguments and at least for boundedly rational human agents, infinite structures as belief makers are not cognitively plausible. Nevertheless, we will sometimes briefly discuss which of our results carry over to infinite argumentation frameworks.

Given such an argumentation framework, a question we frequently ask is "which arguments can be accepted in a rational way?" One way of operationalizing rationality in argumentation theory is the concept of admissibility. Intuitively, a set of arguments is admissible if all its member are acceptable in a certain sense and can be accepted together without incurring a conflict. In what follows, we make our notions of conflict and acceptability precise. Intuitively, it make sense to define conflict in terms of the attack relation: it appears spurious to accept two arguments that attack each other. We therefore define the following "neutrality" function:

Definition 2 (Neutrality-Function, Grossi [2012])**.** Given an argumentation framework $\mathcal{A} = \langle A, \twoheadrightarrow \rangle$ the neutrality function $n_{\mathcal{A}} : 2^A \twoheadrightarrow 2^A$ of $\mathcal{A}$ is defined as

$$n_{\mathcal{A}}(X) = \{x \in A : X \not\twoheadrightarrow x\}$$

[1] though not unheard of, see Belardinelli et al. [2015]

A set of arguments X thus is *conflict free* in an argumentation framework $\mathcal{A}$ iff $X \subseteq n_{\mathcal{A}}(X)$.

As for acceptability, the intuition is that we can accept an argument if we can *defend* it. Intuitively, an attacker of an attacker of a given argument defends that argument against its attacker. An argument a (or set X) (perfectly) defends an argument b iff a (X) attacks all attackers of b. More generally, we define the following defense function:

Definition 3 (Defense Function)**.** Given an argumentation framework $\mathcal{A} = \langle A, \twoheadrightarrow \rangle$ the characteristic or defense function $d_{\mathcal{A}} : 2^A \twoheadrightarrow 2^A$ of $\mathcal{A}$ is defined as

$$d_{\mathcal{A}}(X) = \{x \in A : \forall y \in A \text{ if } y \twoheadrightarrow x \text{ then } X \twoheadrightarrow y\}$$

An argument a is said to be *acceptable* to a set X if X defends it or in other words $a \in d_{\mathcal{A}}(X)$. Finally, a set of arguments X is *self-acceptable* iff $X \subseteq d_{\mathcal{A}}(X)$.

Clearly, it makes sense to regard a set of arguments that can defend itself against all of its attackers as stronger than one that can not. In addition, we can ask for maximal self-acceptable sets, i.e. sets of arguments that contain everything they defend.

We can construct these maximmal sets by iterating the defense function: intuitively, if this iteration reaches a fixpoint we obtain one of the maximal sets we asked for: it defends all of its members and nothing else.

In other words, denote $d_{\mathcal{A}}(d_{\mathcal{A}}(X))$ by $d^2_{\mathcal{A}}(X)$ and so forth. We call the series $\{d^n_{\mathcal{A}}(X)\}_{n \in \mathbb{N}}$ a stream and say it stabilizes with limit or fixpoint $d^m_{\mathcal{A}}(X)$ iff m is a natural number s.t. $d^m_{\mathcal{A}}(X) = d^{m+1}_{\mathcal{A}}(X)$.

It is often useful to consider not only direct attacks and defenses between arguments but also indirect connections: intuitively, an attacker of a defender of a given argument attacks it, although indirectly. To pin down these notions, we define the following extensions of the attack relation

Definition 4 ((Odd, Even) Transitive Closure)**.** Let $\mathcal{A}$ be an argumentation framework with attack relation $\twoheadrightarrow$. We denote by $\twoheadrightarrow^+$, $\twoheadrightarrow_{odd}$, $\twoheadrightarrow_{even}$ the transitive, odd-transitive and even-transitive closure of $\twoheadrightarrow$ defined recursively as

$$a \twoheadrightarrow_1 b \text{ iff } a \twoheadrightarrow b \tag{2.1}$$

$$a \twoheadrightarrow_{n+1} b \text{ iff there is c such that } a \twoheadrightarrow_n c \text{ and } c \twoheadrightarrow b \tag{2.2}$$

Then $\twoheadrightarrow^+ = \bigcup_{n \in \mathbb{N} \setminus 0} \twoheadrightarrow_n$, $\twoheadrightarrow_{odd} = \bigcup_{n \in \mathbb{N}_{odd} \setminus 0} \twoheadrightarrow_n$, $\twoheadrightarrow_{even} = \bigcup_{n \in \mathbb{N}_{even} \setminus 0} \twoheadrightarrow_n$.

Putting the notion of (odd and even) transitive closure to use we can pin down indirect attack and defense.

Definition 5 (Reachability, Indirect Attack, Indirect Defense)**.** Let $\mathcal{A}$ be an argumentation framework with attack relation $\twoheadrightarrow$.

- If $a \twoheadrightarrow^+ b$ we say that b is reachable from a.
- If $a \twoheadrightarrow_{odd} b$ we say that a indirectly attacks b.

- If $a \twoheadrightarrow_{even} b$ we say that a indirectly defends b.
- If for some b, $a \twoheadrightarrow_{even} b$ and $a \twoheadrightarrow_{odd} b$ we say that a is controversial.

Although we now have the notions of acceptability and conflict in place, it is not obvious how high we should set the bar for acceptance. A set of accepted arguments is called an *extension*. Intuitively, the minimal extension we would want to accept should be conflict-free and self-defending. However, it turns out that many such extensions may exist in a given argumentation framework. Is this a problem? It depends! For example when we would like to define a notion of belief in terms of arguments, it may be reasonable to demand consistency: an agent should not believe "P" and "not P" at the same time. But what if there are conflict-free and self-defending sets of arguments supporting either option? Considerations such as this one have led to the proposal of various "semantics" - i.e. ways of choosing the extensions in a given argumentation framework - in the literature. We mention here only a subset of them.

Definition 6 (Extensions and Semantics)**.** Given an argumentation framework $\mathcal{A} = \langle A, \twoheadrightarrow \rangle$ a *semantic* S is a subset of the powerset of A. A member of S is called an *extension* under S. The following are some important semantics:

- The conflict-free semantic is the set of extensions $\mathcal{CF} = \{X : \text{X is conflict free}\}$
- The self-acceptable semantic is the set of extensions $\mathcal{SA} = \{X : \text{X is self-acceptable}\}$
- The admissible semantic is the set of extensions $\mathcal{AD} = \{X : \text{X is conflict free and self-acceptable}\}$
- The complete semantic is the set of extensions $\mathcal{CO} = \{X : \text{ X is admissible and } X = d_{\mathcal{A}}(X)\}$
- The grounded semantic is the (singletion) set of extensions $\mathcal{GR} = \{X : \text{X is the } \subseteq\text{-least complete extension}\}$
- The preferred semantic is the set of extensions $\mathcal{PR} = \{X : \text{X is a } \subseteq\text{-maximal complete extension}\}$

Some of these semantics allow for several extensions to exist. We still need a way to deal with this. Essentially, we have two choices: accept the argument if it is in some extension; or accept it if it is in all of the extensions. This leads us to the following definitions:

Definition 7 (Acceptance)**.** Given an argumentation framework $\mathcal{A}$ and a semantic S for $\mathcal{A}$, an argument a is

- skeptically accepted under S iff $a \in X$ for all $X \in S$.
- credulously accepted under S iff $a \in X$ for some $X \in S$.

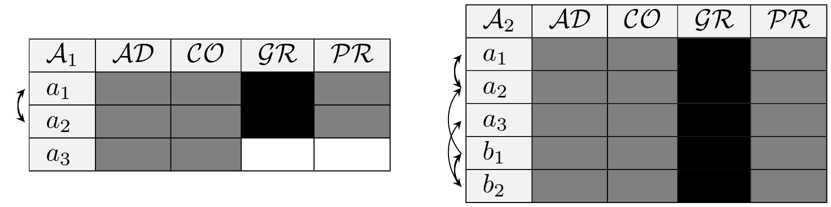

Figure 2.2: Acceptance in Alice's argumentation framework before ($\mathcal{A}_1$) and after ($\mathcal{A}_2$) the conversation with Bob where black marks rejection, gray credulous acceptance and white skeptical acceptance.

- rejected under S iff $a \notin X$ for all $X \in S$.

Example 3. Later that day, Alice has the following conversation wih her friend Bob.

- Alice: "Do you believe sea level rise is really happening?
- Bob: b_1:"We ought to believe what our best science supports."
- Alice: "Anyway, how are you? You sound a bit under the weather"
- Bob: "Terrible! I have a cold."
- Alice: a_3 :"There is a new study saying bloodletting is effective against the common cold and has few side effects. Will you go for bloodletting?"
- Bob: b_2 :"Science is an inherently fallible enterprise. Therefore we should not prematurely believe our scientists' conclusions."

Figure 2.2 shows how her argumentation framework might look like after this exchange. It also shows her options with respect to which arguments she should accept.

Finally, we mention here a property of semantics that will come in handy in our treatment of argument dynamics:

Definition 8 (Directionality)**.** A semantic S satsifies *directionality* iff for any argumentation framework $\mathcal{A}$ and unattacked set $U \subseteq A$, $S((U, \twoheadrightarrow_{\downarrow U})) = \{X \cap U : X \in S(\mathcal{A})\}$ where $\twoheadrightarrow_{\downarrow U}$ is the restriction of $\twoheadrightarrow$ to U. In other words, S is directional iff the S-extensions of the restriction of $\mathcal{A}$ to any unattacked set U are exactly the intersections of the S-extensions of $\mathcal{A}$ with U.

That is under directional semantics the status of an argument a depends only on the arguments from which a is reachable. Hence the argument statusses of members of unattacked sets should remain the same under the semantic when the rest of the argumentation framework is modified.

2.2 Basic Results

We state here without proof a number of results from abstract argumentation theory with respect to the existence and uniqueness of extensions.

Proposition 1 (Dung [1995])**.** Let $\mathcal{A}$ be an argumentation framework. Then $\mathcal{A}$

- has at least one admissible extension (possibly the empty set);
- thereby has at least one complete extension (possibly empty).
- has a unique grounded extension;
- has at least one preferred extension;

A useful way to express the grounded extension is as the least fixpoint of the defense function:

Proposition 2 (Dung [1995])**.** Given an argumentation framework $\mathcal{A}$ a set of arguments $X \subseteq A$ is the grounded extension iff X is the least fixpoint of the defense function $d_{\mathcal{A}}$ or in other words X is grounded iff for the unique m such that $d^m_{\mathcal{A}}(\varnothing) = d^{m+1}_{\mathcal{A}}(\varnothing)$ we have $d^m_{\mathcal{A}}(\varnothing) = X$.

With respect to directionality we have the following result:

Proposition 3 (Baroni and Giacomin [2007])**.** The admissible, grounded, complete, and preferred semantics satisfy directionality.

Finally, there is a strong connection between well-foundedness of argumentation frameworks and how well the semantics behave. An argumentation framework is well-founded iff it contains no infinite chain of arguments $a_1 \twoheadrightarrow a_2 \twoheadrightarrow a_3 ....$[2]. It is easy to see that in the finite case an argumentation framework is well-founded iff it is acyclic. We define cycles in terms of indirect defense and attack as follows:

Definition 9. Let $\mathcal{A}$ be an argumentation framework. We say that $\mathcal{A}$ is cycle-free, odd-cycle-free, even-cycle-free respectively iff there is no argument $a \in A$ s.t. $a \twoheadrightarrow^+ a$, $a \twoheadrightarrow_{odd} a$, $a \twoheadrightarrow_{even} a$ respectively.

If an argument indirectly attacks itself, it is in a way pathological: intuitively, a good argument should not defend arguments that attack it. Likewise, if an argument defends itself (indirectly), that implies that the agent (or whoever came up with the attack relation) is undecided which of two contradicting arguments is stronger. The following proposition shows that both of these "defects" have consequences on the well-behavedness of the argument semantics.

Proposition 4 (Dung [1995],Dunne and Bench-Capon [2001], Dunne and Bench–Capon [2002])**.** Let $\mathcal{A}$ be an argumentation framework.

[2]Note that in Modal Logic this is usually referred to as *converse well-foundedness*

- If $\mathcal{A}$ is well-founded the complete, preferred and grounded extensions coincide (and are thereby unique).
- If $\mathcal{A}$ is finite and cycle-free the complete, preferred and grounded extensions coincide (and are thereby unique).
- If $\mathcal{A}$ is finite and odd-cycle-free then it has at least one non-empty preferred extension.
- If $\mathcal{A}$ is finite and even-cycle-free then it has a unique preferred extension.

Chapter 3

Logics for Argumentation

3.1 Modal Logic for Argumentation: K_U

It is noteworthy that an abstract argumentation framework is nothing else but a relational structure (with labels) which is precisely the fundamental notion used in the standard Kripke-semantics of modal logics. Since there are soundness and completeness results for many classes of relational structures and modal languages, the setting of modal logic arguably allows for a more parsimonious approach to studying abstract argumentation than employing the full expressive power of first order logic. In addition, modal logicians have developed an immense toolkit to work with relational structures and the modal logical setting allows to make use of these powerful tools.

For example, frame definability, the problem of which properties of relational structures can be defined in terms of modal formulas and if so, by which formula precisely, has been extensively studied in modal logic.

It therefore seems natural to investigate abstract argumentation from a modal logical perspective. This approach has been pursued in depth by Grossi (2010, 2011a, 2011b, 2012, 2013, 2014). We shall summarize here some of the basic results.

Definition 10 (Attack Model)**.** Given a set of propositional letters P an argumentation framework $\mathcal{A} = \langle A, \twoheadrightarrow \rangle$ and a valuation function $V_{prop} : P \longrightarrow 2^{\mathcal{A}}$, an attack model is a tuple $\mathcal{M} = \langle \mathcal{A}, V \rangle$.

Definition 11 (The logic K_U for argumentation)**.** Given an attack model $\mathcal{M}$, define the following language:

$$\mathcal{L}(P) : \phi ::= p \mid \bot \mid \neg\phi \mid \phi \wedge \phi \mid \Diamond\phi \mid \Diamond_U \phi$$

where $p \in P$ is a propositional letter.

Satisfaction is defined as usual for $\bot, \neg, \wedge$ plus

$\mathcal{M}, a \vDash \Diamond\phi$ iff there is $a' \in A : a \twoheadleftarrow a'$ and $\mathcal{M}, a' \vDash \phi$
$\mathcal{M}, a \vDash \Diamond_U \phi$ iff there is $a' \in A : \mathcal{M}, a' \vDash \phi$

The valuation function V_{prop} is extended to a function $V : \mathcal{L} \longrightarrow 2^A$ such that $V(\phi) = V_{prop}(\phi)$ if $\phi \in P$ and $V(\phi) := \{a \in A : \mathcal{M}, a \vDash \phi\}$ else.

A sound and complete axiomatization of K_U is given by the propositional tautologies plus the following axioms and rules:

$$\begin{array}{lr}
\Diamond -Normality: & \Diamond(p \vee q) \leftrightarrow (\Diamond p \vee \Diamond q) \\
\Diamond_U -Normality: & \Diamond_U(p \vee q) \leftrightarrow (\Diamond_U p \vee \Diamond_U q) \\
\Diamond -Dual: & \Diamond p \leftrightarrow \neg \Box \neg p \\
\Diamond_U -Dual: & \Diamond_U p \leftrightarrow \neg \Box_U \neg p \\
\Diamond_U -Reflexivity: & p \to \Diamond_U p \\
\Diamond_U -Symmetry: & p \to \Box_U \Diamond_U \\
\Diamond_U -Transitivity: & \Diamond_U \Diamond_U p \to \Diamond_U p \\
\Diamond_U -Inclusion: & \Diamond p \to \Diamond_U p \\
Necessitation: & \text{if } \vdash \phi \text{ then } \vdash \Box\phi \\
\Box_U -Necessitation: & \text{if } \vdash \phi \text{ then } \vdash \Box_U\phi \\
Modus\ Ponens: & \text{if } \vdash \phi \text{ and } \vdash \phi \to \psi \text{ then } \vdash \psi
\end{array}$$

Recall the semantics defined in definition 17. We can express them in the language of K_U as follows:

Proposition 5 (Grossi [2012])**.** Let $\mathcal{A}, V$ be an attack model, $a \in A$ an arbitrary argument and $\phi \in \mathcal{L}(P)$

- $V(\phi)$ is a conflict-free extension of $\mathcal{A}$ iff $\mathcal{A}, V, a \vDash CF(\phi)$ where $CF(\phi) := \Box_U(\phi \to \neg \Diamond \phi)$
- $V(\phi)$ is a self-acceptable extension of $\mathcal{A}$ iff $\mathcal{A}, V, a \vDash SA(\phi)$ where $SA(\phi) := \Box_U(\phi \to \Box \Diamond \phi)$
- $V(\phi)$ is an admissible extension of $\mathcal{A}$ iff $\mathcal{A}, V, a \vDash AD(\phi)$where $AD(\phi) := \Box_U(\phi \to (\neg \Diamond \phi \wedge \Box \Diamond \phi))$
- $V(\phi)$ is a complete extension of $\mathcal{A}$ iff $\mathcal{A}, V, a \vDash CO(\phi)$ where $CO(\phi) := \Box_U((\phi \to \neg \Diamond \phi) \wedge (\phi \leftrightarrow \Box \Diamond \phi))$

Example 4. Suppose we treat propositional letters as topics, say s for the topic sea level rise and b for the topic bloodletting. Define the attack model $\mathcal{M}_1 = (\mathcal{A}, V)$ where $V(s) = \{a_1, a_2\}$ in $\mathcal{A}_1$ and $V(b) = \{a_3\}$ (see Figure 3.1. It is easy to check the following:

- $\mathcal{M}_1 \vDash \Box_U(b \to (\neg \Diamond b \wedge \Box \Diamond b))$
- $\mathcal{M}_1 \vDash \Box_U((b \to \neg \Diamond b) \wedge (b \leftrightarrow \Box \Diamond b))$

$\mathcal{A}_1$	V	$\mathcal{AD}$	$\mathcal{CO}$	$\mathcal{GR}$	$\mathcal{PR}$
a_1	s				
a_2	s				
a_3	b				

Figure 3.1: Alice's original argumentation framework equipped with a valuation to denote the topics s=sea level rise and b=bloodletting.

- $\mathcal{M}_1 \nvDash \Box_U(s \to (\neg \Diamond s \wedge \Box \Diamond s))$

That is, $V(b)$ is admissible and complete while $V(s)$ is neither. Hence one might say Alice has made up her mind on the topic of bloodletting but has currently no opinion regarding sea level rise.

Expressing the grounded and preferred extensions requires richer languages. One way is to take a dynamic view and make use of fixpoint definitions of these semantics by introducing model transformers and iterating them. Another is to equip the language with names and eventualities. Grossi [2012] takes the first approach.

However, in the finite case the grounded and preferred extensions can be expressed in a different way: suppose that for each argument a we had a propositional letter p_a that names a, i.e. $V(p_a) = \{a\}$. Then we could express any set of arguments X as a disjunction $\bigvee_{a \in X} p_a$. Denote this disjunction by $\check{X}$.

Proposition 6. Let $\mathcal{A}$ be a finite argumentation model and $X \subseteq A$. Then

1. $V(\phi)$ is the grounded extension of $\mathcal{A}$ iff

$$\mathcal{A}, V, a \vDash CO(\phi) \wedge \bigwedge_{X \subseteq A} ((CO(\check{X}) \to \Box_U(\phi \to \check{X}))$$

2. $V(\phi)$ is a preferred extension iff

$$\mathcal{A}, V, a \vDash CO(\phi) \wedge \bigwedge_{X \subseteq A, X \neq V(\phi)} (CO(\check{X}) \to \Diamond_U(\phi \wedge \neg \check{X}))$$

In words, $V(\phi)$ is grounded iff it is complete (6.1, first conjunct and it is a subset of every complete extension (6.1, second conjunct). $V(\phi)$ is preferred iff it is complete and it is not a strict subset of any other complete extension (6.2, second conjunct).

Proof. (1) $\Rightarrow$ Suppose $\mathcal{A}, V, a \vDash \Box_U((\phi \to \neg \Diamond \phi) \wedge (\phi \leftrightarrow \Box \Diamond \phi)) \wedge \bigwedge_{S \subseteq A}(\Box_U((S \to \neg \Diamond S) \wedge (S \leftrightarrow \Box \Diamond S)) \to \Box_U(\phi \to S))$. We will go through the conjuncts one by one. The first conjunct is the formula representation of the complete extension (see Proposition 5), hence $V(\phi)$ needs to be a complete extension. The second conjunct asserts an implication for all subsets of Arg. The antecedent of that implication again states the definition of the complete semantic. The consequent states that at all arguments, ϕ implies $\check{X}$. Hence the second conjunct states that for any $X \subseteq A$, iff it is complete, then ϕ implies $\check{X}$ everywhere. But this is the case iff $V(\phi) \subseteq X$. Hence $V(\phi)$ is complete and a subset of every complete extension. But this is just the definition of the grounded semantic, as desired. $\Leftarrow$ Similar.

(2) The proof is similar to (1), except that now we have that for all $X \subseteq A$, $X \neq V(\phi)$, we have that ϕ is no subset of X. But then since $V(\phi)$ is complete, it must be a maximal complete extension under set inclusion and hence $V(\phi)$ is a preferred extension.

$\square$

Example 5. In $\mathcal{M}_1$ (3.1), neither $V(b)$ nor $V(s)$ are a preferred extension. But the sets $X_1 = \{a_1, a_3\}$ and $X_2 = \{a_2, a_3\}$ are. Hence given names $p_{a_1}, p_{a_2}, p_{a_3}$ with $V(p_{a_1}) = \{a_1\}$, $V(p_{a_2}) = \{a_2\}$, $V(p_{a_3}) = \{a_3\}$ we get:

- $\mathcal{M}_1 \models CO(\check{X}_1) \wedge \bigwedge_{X \subseteq A, X \neq X_1}(CO(\check{X}) \rightarrow \Box_U \neg(\check{X}_1 \rightarrow \check{X}))$
- $\mathcal{M}_1 \models CO(\check{X}_2) \wedge \bigwedge_{X \subseteq A, X \neq X_1}(CO(\check{X}) \rightarrow \Box_U \neg(\check{X}_2 \rightarrow \check{X}))$

Thus adding names to the language would allow us to gain additional expressivity with respect to argument semantics. Therefore, in the next section we will investigate such a language.

3.2 Hybrid Logic for Argumentation

3.2.1 Where Modal Logics and Argumentation Frameworks match and where they don't

We have seen in the last section that modal logic equips us with formal languages that allow us to capture a lot of what we would like to be able to say about argumentation frameworks. However, the correspondence between modal language and argumentation frameworks is not a perfect one.

Firstly, semantically speaking, argumentation frameworks are *frames* not models. That means that modal languages are in a way *too expressive* for standard abstract argumentation theory as they provide the full expressiveness of propositional logic to talk about properties of arguments. For example, we might be interested in all arguments from analogy in a given AF. Or in all arguments that are either from analogy or deductive. Modal argumentation logics go beyond standard abstract argumentation theory in providing a language for such properties of arguments.

While further investigation of this added expressiveness would be very interesting in its own right (especially when taking the step to first-order modal languages), the lack of propositional letters in abstract argumentation theory shows that it is really a theory that operates on the frame level rather than the model level. Thus the minimal logic rich enough to capture argumentation frameworks should be a *frame language*, i.e. a language whose formulas can be evaluated on argumentation frameworks without the need for a valuation function to interpret propositional letters.

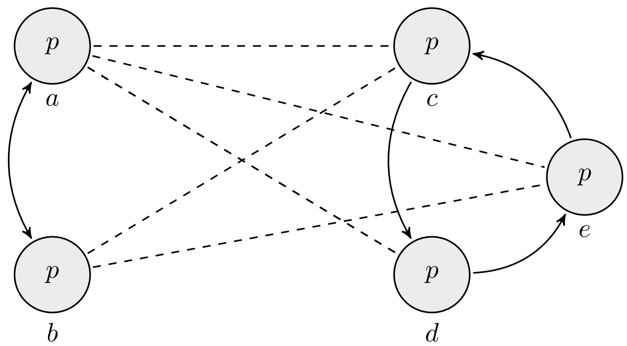

Figure 3.2: Representation of the bisimilar models in the proof for Proposition 7. The dotted lines represent the bisimulation Z.

In addition, we will explore where modal languages are *not expressive enough* to deal with standard abstract argumentation theory. As we have seen in Chapter 2, an important feature of argumentation frameworks are indirect attacks and defenses. Therefore we would like to be able to modally capture the indirect attackers and defenders of an argument.

Say that for this purpose we define the following two modalities:

- $\mathcal{A}, V, a \vDash \Diamond_{odd}\phi$ iff there is an argument b such that $\mathcal{A}, V, b \vDash \phi$ and a is indirectly attacked by b.
- $\mathcal{A}, V, a \vDash \Diamond_{even}\phi$ iff there is an argument b such that $\mathcal{A}, V, b \vDash \phi$ and a is indirectly defended by b.

Is K_U expressive enough to define these operators? We obtain a negative result:

Proposition 7. The operators $\Diamond_{odd}$ and $\Diamond_{even}$ are not definable in K_U.

Proof. Consider the following two models: $\mathcal{A} = (A = \{a,b\}, \twoheadrightarrow = \{a,b\}\times\{a,b\}, V(p) = \{a,b\})$, $\mathcal{B} = (B = \{c,d,e\}, \twoheadrightarrow = \{(c,d),(d,e),(e,c)\}, V(p) = \{c,d,e\})$.
Clearly Z={(a,c),(a,d),(a,e),(b,c),(b,d),(b,e)} is a K_U-bisimulation between $\mathcal{A}$, $\mathcal{B}$ (Compare Figure 3.2. Hence $\mathcal{A}$, $\mathcal{B}$ are K_U-equivalent. But $\mathcal{A} \vDash \Box_{odd}p$ and $\mathcal{A} \nvDash \Box_{even}p$ whereas $\mathcal{B} \vDash \Box_{even}p$ and $\mathcal{B} \nvDash \Box_{odd}p$. Hence $\Box_{odd}, \Box_{even}$ are not definable in K_U.

□

The notion of indirect attacks comes in especially handy when studying cycles in argumentation frameworks. Recall from Chapter 2 that the absence of odd or even cycles is a sufficient condition for uniqueness of extensions under certain semantics for finite argumentation frameworks. Thus we would like to be able to express the presence or absence of cycles in the language. Evaluating locally, a state is in an odd/even cycle iff it can reach itself through the odd/even transitive closure of the attack relation. The odd/even Diamonds only get us half-way towards expressing this: we still need an ingredient to denote individual states.

Again, this can be accomplished by using names. We have already seen in the last section that names add useful expressiveness a modal language for argumentation. Thus one might ask whether it is possible to get them for free by defining them in K_U? This would require that each argument a in a model is identified by a propositional letter p s.t. $\mathcal{A}, a \vDash p$ and there is no $b \in Arg$ s.t. $\mathcal{A}, b \vDash p$. Again, this goes beyond the expressiveness of K_U as the following simple example shows.

Proposition 8. Names are not definable in K_U.

Proof. Consider the following two bisimilar models (a single reflexive point and an infinite, directed chain): $\mathcal{A} = (\{a\}, \{a\} \times \{a\}, V(p) = \{a\})$, $\mathcal{B} = (\mathbb{N}, \{(n, n+1\}_{n \in \mathbb{N}}, V(p) = \mathbb{N})$.

□

Names would also help us in a different way: suppose we use an abstract argumentation framework to represent the arguments an agent is aware of. How can we query the model as to whether the agent knows of a certain argument? Without names we could give a description of the argument in terms of properties it satisfies. Since it would not be practical to model all the properties an argument could possibly satisfy, we would have to restrict the set of properties. Hence there may well be several arguments that satisfy the same properties from among the ones we consider, we cannot be sure we identified the right argument with our query. With names this is not a problem since every name denotes a unique argument. Say we are interested in an argument named i. Our query would thus amount to checking whether the model satisfies the formula $\Box_U i$.

3.2.2 HAL: Hybrid Argumentation Logic

We shall therefore develop a logic that combines the simplicity of a frame language with the expressiveness we desire. It is based on hybrid logic (for detailed treatments of hybrid logic see Areces and ten Cate [2007], Cate [2005] and Blackburn [2001]; Chapter 7.3). In essence, hybrid logic expands normal modal languages by adding a set of constants called nominals to denote individual states. In our case we shall follow the approach established by Hansen [2011] where only a subset of the nominals denote a state. Thus we are equipped with an overabundance of names that allows us to add or remove states when needed. This will come in handy when studying dynamics. In addition, such partially denoting nominals allow some additional frame definability: it is well-known that the class of finite frames is not modally definable. Using a language with a countably infinite number of nominals we can at least provide a necessary condition for a frame to be finite (or a sufficient condition for it to be infinite). We start with the semantics.

Definition 12. Let A be a countable set of arguments, Nom be a countably infinite set of nominals. A *named frame* $\mathcal{N}$ is defined as a tuple $\mathcal{N} = (A, R, \{D_i\}_{i \in Assigned})$ where $Assigned \subseteq Nom$, R is a binary relation over A and every D_i is a subset of A. We call $\mathcal{N}$ *nice* if for every i, D_i is a singleton.

Given such frames we define a logical language as follows.

Definition 13 (Hybrid Argumentation Logic)**.** Let Nom be a set of nominals. Then we define the language $(\mathcal{H}_{odd}, @)$ of HAL as follows:

$$\mathcal{L}(Nom) : \phi ::= i \mid \bot \mid \neg\phi \mid \phi \wedge \phi \mid \Box\phi \mid \Box_{odd}\phi \mid @_i\phi$$

where $i \in Nom$ is a nominal. Given a named frame $\mathcal{N} = (A, \twoheadrightarrow, \{D_i\}_{i \in Assigned})$, satisfaction is defined as follows

$\mathcal{N}, a \models i$ iff $a \in D_i$
$\mathcal{N}, a, \not\models \bot$
$\mathcal{N}, a, \models \neg\phi$ iff $\mathcal{N}, a \not\models \phi$
$\mathcal{N}, a, \models \phi \wedge \psi$ iff $\mathcal{N}, a \models \phi$ and $\mathcal{N}, a, \models \psi$
$\mathcal{N}, a, \models \Box\phi$ iff for all $a' \in A : a \twoheadleftarrow a'$ and $\mathcal{N}, a' \models \phi$
$\mathcal{N}, a \models \Box_{odd}\phi$ iff for all $a' \in A$ s.t. $a \twoheadleftarrow^{odd} a'$ and $\mathcal{M}, a' \models \phi$ where $\twoheadleftarrow_{odd}$ is the odd transitive closure of $\twoheadleftarrow$.
$\mathcal{N}, a, \models @_i\phi$ iff there is a', such that $a' \in D_i$, $\mathcal{N}, a' \models \phi$

We define the following abbreviations: $\top, \vee, \rightarrow, \Diamond$ are as usual, $\mathcal{N}, a \models [@]_a\phi$ iff $\mathcal{N}, a \models \neg @_a \neg\phi$, $\mathcal{N}, a \models \Box_{even}\phi$ iff $\mathcal{N}, a \models \Box_{odd} \Box \phi$ and $\mathcal{N}, a \models \Box^{+}\phi$ iff $\mathcal{N}, a \models \Box_{even}\phi \wedge \Box_{odd} \Box \phi$. In addition, as before, for any set $X \subseteq A$ we use $\check{X}$ to denote the disjunction of nominals $\bigvee_{D_i \subseteq X} i$. Moreover, whenever we work in a finite argumentation framework, we put $\mathcal{N}, a \models \Box_U\phi$ iff $\mathcal{N}, a \models \bigwedge_{D_i \subseteq A} @_i\phi$ and $\mathcal{N}, a \models \Diamond_U\phi$ iff $\mathcal{N}, a \models \bigvee_{D_i \subseteq A} @_i\phi$. Finally, in the finite setting we take $\mathcal{N}, a \models \downarrow x.\phi$ to mean $\mathcal{N}, a \models \bigcap_{D_i \in A}(i \rightarrow \phi)$ where every occurence of x in ϕ is substituted by i.

With these definitions in place we turn to axiomatization. The axiomatization is based on the one provided by Hansen [2011] provides for hybrid logics with partially assigning nominals except that we add the odd and even iteration and induction axioms. They correspond to the usual axioms for transitive closure operators but are adapted for the odd and even case.

(K)	$\Box(\phi \wedge \psi) \leftrightarrow \Box\phi \wedge \Box\psi$
$(K_{[@]})$	$[@]_i(\phi \wedge \psi) \leftrightarrow [@]_i\phi \wedge [@]_i\psi$
$(@-Functionality)$	$@_i\phi \to [@]_i\phi$
$(Introduction)$	$i \to (\phi \leftrightarrow @_i\phi)$
$(Weak\ Reflexivity)$	$[@]_i i$
$(Bridge)$	$(@_i \Diamond j \wedge @_j\phi) \to @_i \Diamond \phi$
$(Weak\ Agree)$	$@_j @_i\phi \to @_i\phi$
$(Back)$	$\Diamond @_i\phi \to @_i\phi$
$(Denote)$	$@_i\phi \to @_i i$
$(Collapse)$	$@_i i \to ([@]_i\phi \to @_i\phi)$
(K_{odd})	$\Box_{odd}(\phi \wedge \psi) \leftrightarrow \Box_{odd}\phi \wedge \Box_{odd}\psi$
$(Odd\ Iteration)$	$\Box_{odd}\phi \leftrightarrow \Box(\phi \wedge \Box_{even}\phi)$
$(Odd\ Induction)$	$\Box\phi \wedge \Box_{odd}(\phi \to \Box\Box\phi) \to \Box_{odd}\phi$
$(Even\ Induction)$	$\Box\Box\phi \wedge \Box_{even}(\phi \to \Box\Box\phi) \to \Box_{even}\phi$
$Necessitation:$	if $\vdash \phi$ then $\vdash \Box\phi$
$Odd-Necessitation:$	if $\vdash \phi$ then $\vdash \Box_{odd}\phi$
$[@]_i-Necessitation:$	if $\vdash \phi$ then $\vdash [@]_i\phi$
$Modus\ Ponens:$	if $\vdash \phi$ and $\vdash \phi \to \psi$ then $\vdash \psi$
$Name$	If $\vdash [@]_i\phi$ and i does not occur in ϕ, then $\vdash \phi$
$Paste$	If $\vdash (@_i \Diamond j \wedge @_j\phi) \to \phi$ and $i \neq j$ and j does not occur in ϕ and ψ, then$@_i \Diamond \phi \to \phi$

Note also, that the following are theorems of HAL (Hansen [2011]):

Proposition 9. The following formulas are HAL-theorems:

1. Weak Self-Dual: : $@_i j \to ([@]_i\phi \leftrightarrow @_i\phi)$
2. Sym: $@_i j \leftrightarrow @_j i$
3. Weak Reverse Agree: $(@_i i \wedge @_j j) \to (@j\phi \leftrightarrow @_i @_j\phi)$
4. Nom: $@ij \to (@_i\phi \leftrightarrow @_j\phi)$
5. @-transitivity: $(@_i j \wedge @_j k) \to @_i k$
6. K_{even}: $\Box_{even}(\phi \wedge \psi) \leftrightarrow \Box_{even}\phi \wedge \Box_{even}\psi$
7. Even Iteration: $\Box_{even}\phi \leftrightarrow \Box\Box(\phi \wedge \Box_{even}\phi)$
8. Even Necessitation: if $\vdash \phi$ then $\vdash \Box_{even}\phi$

Proof. We (sketch-)prove only items 9.6 to 9.8 which do not appear in (Hansen [2011]).

9.6: $\Rightarrow$ We use the definition of $\Box_{even}\phi$ as $\Box\,\Box_{odd}\,\phi$ and then apply first K_{odd} and then K. Then we use the definition of $\Box_{even}\phi$ again to obtain the desired theorem. $\Leftarrow$: similar.

9.7: $\Rightarrow$ We use the the definition of $\Box_{even}\phi$ as $\Box\,\Box_{odd}\,\phi$ and then apply odd-iteration. $\Leftarrow$ We take the contrapositive and then proceed as in the left-to-right direction.

9.8: By applying first Odd Necessitation, then Necessitation, then the definition of $\Box_{even}\phi$ as $\Box\,\Box_{odd}\,\phi$.

□

3.2.3 Soundness and Completeness of HAL?

Before we turn to the questopm of soundness and completeness of HAL, we have to take care of a peculiarity of the frame semantics we are using: completeness for Hybrid logics is usually obtained with respect to the class of all frames using the following standard approach: Take the contrapositive of the (strong) completeness claim, i.e. every consistent set of formulae can be satisfied in some model. Construct a named canonical model on which every Hybrid-logic-consistent set of formulas is satisfiable. However, we have defined names as a frame property, not a model property. Hence we want to prove completeness with respect to the class of nice named frames over Nom, not all frames. We define this class as follows:

$$NAMED = \{\mathcal{N} : (A, \twoheadrightarrow) \text{ is a frame }, Assigned \subseteq Nom\}$$

where $\mathcal{N} = ((A, \twoheadrightarrow, \{D_i\}_{i \in Assigned})$ is nice and named. Hence for completeness we strong completeness of HAL with respect to $NAMED$ we would have to prove that every consistent set of formulae can be satisfied on some frame $\mathcal{N} \in NAMED$. However, we begin with soundness:

Proposition 10. HAL is sound with respect to $NAMED$.

Proof. We check here the validity of the axioms not included by Hansen [2011].

- K_{odd}: $\Rightarrow$: Suppose $\mathcal{N}, a \vDash \Box_{odd}\phi \wedge \Box_{odd}\psi$. Then for any indirect attacker a' of a, $\mathcal{N}, a' \vDash \phi$ and $\mathcal{N}, a' \vDash \psi$. But then $\mathcal{N}, a' \vDash \phi \wedge \psi$ and hence $\mathcal{N}, a \vDash \Box_{odd}(\phi \wedge \psi)$ as desired. $\Leftarrow$ Suppose $\mathcal{N}, a \vDash \Box_{odd}(\phi \wedge \psi)$. Then for any indirect attacker a' of a, $\mathcal{N}, a' \vDash \phi \wedge \psi$. Then a fortiori, $\mathcal{N}, a' \vDash \phi$ and $\mathcal{N}, a' \vDash \psi$. Hence $\mathcal{N}, a \vDash \Box_{odd}\phi \wedge \Box_{odd}\psi$.
- Odd Iteration: $\Rightarrow$: Suppose $\mathcal{N}, a \vDash \Box_{odd}\phi$. Then for any indirect attacker a' of a, $\mathcal{N}, a' \vDash \phi$. Hence a fortiori for any direct attacker a'' $\mathcal{N}, a'' \vDash \phi$. In addition, any indirect attacker of a is an indirect defender of a''. Hence $\mathcal{N}, a'' \vDash \Box_{even}\phi$. But then $\mathcal{N}, a'' \vDash \phi \wedge \Box_{even}\phi$ and therefore $\mathcal{N}, a \vDash \Box(\phi \wedge \Box_{even}\phi$ as desired. $\Leftarrow$: By applying K.

- Odd Induction: Suppose $\mathcal{N}, a \vDash (\phi \rightarrow \Box\phi) \wedge \Box_{odd}(\phi \rightarrow \Box\Box\phi)$ and $\mathcal{N}, a \vDash \phi$. $\mathcal{N}, a \vDash \Box_{odd}\phi$ iff for any indirect attacker a' of a, $\mathcal{N}, a' \vDash \phi$. We prove this by induction. We need to show that for all odd numbers n, if a is reachable from a through a path of length n, $\mathcal{N}, a' \vDash \phi$. For the base case, assume a' is a direct attacker of a. By assumption we have that $\mathcal{N}, a \vDash \phi$ and ($\mathcal{N}, a \vDash \phi \rightarrow \Box\phi$) and thus we immediately get $\mathcal{N}, a \vDash \Box\phi$. Hence $\mathcal{N}, a' \vDash \phi$. Now suppose we have ϕ for all indirect attackers of a such that there is a path from a' to a of length n. Hence $\mathcal{N}, a' \vDash \phi$. Then by assumption, $\mathcal{N}, a' \vDash \Box\Box\phi$ and hence for any a' such that there is a path of length $n + 2$ we have that $\mathcal{N}, a' \vDash \phi$ as desired.

- Even Induction: analogous to odd induction.

- Odd Necessitation: here we only spell out the induction step for odd necessitation that would form part of an induction to prove soundness of all the inference rules. Suppose $\vdash \phi$ and suppose $\vDash \phi$ if $\vdash \phi$. Then for any $a \in \mathcal{N}$, $\mathcal{N}, a \vDash \phi$. But then a fortiori for all indirect attackers a' of a, $\mathcal{N}, a' \vDash \phi$. Hence $\mathcal{N}, a \vDash \Box_{odd}\phi$ as desired.

□

We conjecture that HAL is also complete with respect to $NAMED$ but we do not currently have a proof for this. We sketch how such a proof would look like in what follows leaving the details to future work.

Firstly, our aim has to be weak completeness rather than strong completeness. The reason is that due to the odd transitive closure operator, HAL is not a compact logic. That is, there exist infinite, consistent sets of HAL-formulas that are not satisfiable at a single argument in a frame although every finite set of HAL formulas is satisfiable.

Definition 14 (Consistency)**.** A set of formulas Γ is consistent if there do not exist formulas $\phi_1, ..., \phi_n \in \Gamma$ s.t. $\Gamma \vdash \neg(\phi_1 \wedge ... \wedge \phi_n)$

Consider the following example, which is a variation of a standard example (found e.g. in Blackburn [2001]):

Example 6.

$$\Gamma = \{\Diamond_{odd}\phi, \neg\Diamond\phi, \neg\Diamond\Diamond\Diamond\phi, ...\}$$

Even if we can prove that every finite subset of Γ is satisfiable on a frame, Γ itself is not: for suppose $\mathcal{N}, a \vDash \Box_{odd}\phi$. Then there is some odd number n such that for some argument a' reachable from a through a path of length n, $\mathcal{N}, a' \vDash \phi$. But by assumption $\mathcal{N}, a \vDash \neg\Diamond^n\phi$ and we get a contradiction.

This implies that we cannot prove satisfiability for every consistent set of HAL-formulas and thus strong completeness is out of reach [1]. Hence instead we will aim for weak completeness.

[1]For finitary proof systems that is. Kooi et al. [2006] provide a strongly complete sequent calculus for various Hybrid logics with transitive closure but extending this approach to HAL is beyond the scope of this work.

The structure of the proof would be as follows: first we define a notion of closure that leads to atoms that are finite in the relevant respects. Then we prove a Lindenbaum's Lemma for this setting. Based on this we construct a named and pasted canonical frame. Finally we establish an existence and a truth lemma. Our proof is based on the general approaches from Blackburn [2001]; (Section 4.8 for completeness of PDL; Section 7.3 for completeness of hybrid logic) and combines it with elements from the completeness proof in Hansen [2011].

The notion of closure will be akin to the closure that features in the usual approach to proving weak completeness for PDL. It is defined as follows:

Definition 15 (Closure)**.** Given a formula ϕ and a set of formulas Σ such that $\phi \in \Sigma$ we define the following closure operations:

- Denote by $\sim \phi$ a formula such that $\sim \phi = \psi$ iff ϕ is of the form $\neg\psi$ and $\sim \phi = \neg\phi$ else. Σ is closed under single negations if whenever $\phi \in \Sigma$, $\sim \phi \in \Sigma$.
- Σ is odd Fischer-Ladner closed if it is closed under subformulas and whenever $\Diamond_{odd} \in X$ then $\Diamond\Diamond_{even} \in X$.

Finally, the closure $\neg FL(\Sigma)$ of a set Σ is the smallest superset of Σ that is closed under single negations and odd Fischer-Ladner closed.

The closure of a set of formulas ensures that it contains all relevant formulas needed to admit an existence lemma while at the same time it is finite. Finiteness is important since instead of maximally consistent sets, the canonical model will use maximally consistent *subsets* of the closure as its worlds. These are called atoms. Since atoms only contain finitely many formulas, we can never encounter an infinite set of formulas that is consistent yet not satisfiable like the one in example 6.

Unfortunately, here we have a problem with the nominals. If we want to prove completeness with respect to named frames, we need to make sure that every atom is named, (i.e. it contains a nominal). Moreover, we need to make sure that we have enough named worlds to satisfy all of the modal operators. This is ensured by a property called pastedness: an atom is pasted iff $@_i \Diamond \phi \in B$ implies that for some $j \in Nom$, $@_i \Diamond j \wedge @_j \phi \in B$.

In order to meet these conditions, we have to add to the closure a nominal for every atom as well as satisfaction operators for every nominal and additional formulas of the form $\diamond i$ for the paste condition. The problem is that this process quickly results in an infinite set once again. Therefore we have to prove that the closure is at least finite in the important dimension:

Proposition 11 ($\Diamond$-Finiteness)**.** Let B an atom over (Γ) where Γ is a finite set of formulas. Then the set $\{\phi \in B : \phi \text{ is of the form } \Diamond\psi\}$ is finite.

This result would ensure that we can never encounter an unsatisfiable atom such as in example 6.

Given such a result we would construct a canonical frame. For this we need an extended Lindenbaum Lemma. The lemma is extended in that it ensures we can build not just any canonical frame but a frame of named and pasted atoms.

Proposition 12 (Lindenbaum Lemma)**.** For any consistent HAL-formula ϕ and any set of HAL-formulas Γ s.t. $\phi \in @\neg iFL$ there is an atom B over Γ such that $\phi \in B$ and B is named and pasted.

We can then extend any consistent HAL-formula to a named and pasted atom and we would make use of this to build a named canonical frame. Unfortunately we cannot simply take the collection of all named and pasted atoms as the worlds of the canonical model since there is no guarantuee that names and diamonds would be assigned to worlds "coherently". Instead, we wwould use a single atom to generate all the worlds of the frame. For this purpose the following lemma is crucial:

Proposition 13. Let Σ be an atom over ϕ. For every nominal $i \in Nom$, the set named by i is defined as $\Delta_i = \{\phi : @_i\phi \in \Sigma\}$. Then:

1. For every assigned nominal i Δ_i is an atom over ϕ that contains i iff $\Delta_i \neq \varnothing$.
2. For all nominals i and j, if $i \in \Delta_j$, then $\Delta_j = \Delta_i$
3. For all nominals i and j, $@_i\phi \in \Delta_j$ iff $@_i\phi \in \Sigma$.
4. If $i \in \Sigma$ then $\Sigma = \Delta_i$

This lemma is not hard to prove for HAL. We would then build the canonical frame on the named sets as follows:

Definition 16 (Canonical Frame)**.** Let ϕ be any HAL-formula, $B \in At(\{\phi\})$ and define $ASSIGNED_B = \{i : @_i i \in B\}$. Then the canonical frame over B for ϕ is given by $\mathcal{N}_{B,\phi} = (A, R, \{D_i\}_{i \in ASSIGNED_B})$ where

- $W = \{\Delta_i : i \in ASSIGNED_B\}$
- $\Delta_j \twoheadrightarrow \Delta_i$ iff for all $\phi \in \Delta_j$, $\Diamond\phi \in \Delta_i$.
- $\Delta_j \twoheadrightarrow_{odd} \Delta_i$ iff for all $\phi \in \Delta_j$, $\Diamond_{odd}\phi \in \Delta_i$.
- $D_i = \{\Delta_i\}$

Next we need to show that the canonical frame is getting the job done. For this we need to prove that there are enough worlds with the right properties to satisfy the modalities and satisfaction operators.

Proposition 14 (Existence Lemma)**.** Let $\mathcal{N}_{B,\phi} = (A, R, \{D_i\}_{i \in ASSIGNED_B})$ be the canonical frame yielded by ϕ. Suppose $a \in A$. Then

1. $@_i\phi \in a$ implies that there is a' with $D_i = \{a'\}$ and $\phi \in a'$.

2. $\diamond\phi \in a$ implies that there is a' such that $a' \twoheadrightarrow a$ and $\phi \in a'$.

3. $\diamond_{odd}\phi \in a$ implies that there is a' such that a' indirectly attacks a and $\phi \in a'$

Finally, we need a truth lemma to complete the proof:

Proposition 15 (Truth Lemma)**.** Let ϕ be any HAL-formula and $\mathcal{N}_{B,\phi}$ its canonical frame over B. Then for any $i \in ASSIGNED_B$ and all HAL-formulas $\psi \in @\neg iFL$, $\mathcal{N}_{B,\phi}, \Delta_i \models \psi$ iff $\psi \in \Delta_i$.

3.2.4 Expressiveness of Hybrid Logic for Argumentation

We are now in a position to deliver the expressiveness results promised in the introduction of this chapter. We are interested in such expressiveness on three levels: firstly, which properties of argumentation frameworks can HAL express; secondly, which properties of sets of arguments within an argumentation framework; finally, which properties of individual arguments are expressible.

The properties we are interested in on the argumentation framework level are mainly questions of existence: does the AF contain a cycle? Does an extension under a given semantic exist? On the level of sets of arguments we seek to identify cycles or extensions. Finally, on the argument level we would like to answer the questions of membership in a cycle and rejection, credulous as well as skeptical acceptance under a semantic.

Expressibility of properties of AFs corresponds to the well-studied area of frame definability in modal logic. I.e., is there a formula in the language that characterises a given class of frames? Before we turn to answer this question we have to define what frame definability means in our setting: normally in modal logic, a formula is said to define a class of frames iff it is true in any model that can be built on any frame in the class. However, we already work with frames, not with models. Hence we need the following definition:

Definition 17. Given an argumentation framework $\mathcal{A} = (A, \twoheadrightarrow)$ we call any named argumentation framework $\mathcal{N} = (A', \twoheadrightarrow', \{D_i\}_{i \in ASSIGNED})$ such that $A = A'$, $\twoheadrightarrow = \twoheadrightarrow'$ a named argumentation framework *based on* $\mathcal{A}$. We say that ϕ is valid on $\mathcal{A}$ and write $\mathcal{A} \models \phi$ if for every $\mathcal{N}$ based on $\mathcal{A}$, $\mathcal{N} \models \phi$.

On this level we obtain the following results:

Proposition 16. Let $\mathcal{A}$ be an argumentation framework. $\mathcal{A}$ is

1. well-founded iff $\mathcal{A} \models \Box_U(\overleftarrow{\Box}\phi \to \phi) \to \phi$ where $\overleftarrow{\Box}\phi$ is the reverse box operator defined as $\downarrow x.\, \Box_U\, (\diamond x \to \phi)$;

2. cycle-free iff $\mathcal{A} \models \neg @_i \diamond^+ i$;

3. odd-cycle-free iff $\mathcal{A} \models \neg @_i \diamond_{odd}\, i$;

4. even-cycle-free iff $\mathcal{A} \models \neg @_i \diamond_{even}\, i$;

5. infinite if $\mathcal{A} \models @_i i$. [2]

Proof. 16.1: Recall that the notion of well-foundedness in argumentation theory corresponds to the notion of converse well-foundedness in modal logic. $\Leftarrow$: Suppose $\mathcal{A}$ is well-founded and let $\mathcal{N}$ be based on $\mathcal{A}$, a be any argument such that $\mathcal{N}, a \models \Box_U(\overleftarrow{\Box}\phi \to \phi)$. Then for all $a' \in \mathcal{N}$, $\mathcal{N}, a' \models \overleftarrow{\Box}\phi \to \phi$. Since $\mathcal{N}$ is well-founded, all chains originating from a are finite. That is, for each chain there is a last argument. Pick an arbitrary such argument a''. Since there is no argument a''' s.t. $a'' \twoheadrightarrow a'''$, $\mathcal{N}, a'' \models \overleftarrow{\Box}\phi$. But then by assumption $\mathcal{N}, a'' \models \phi$. Since a'' was chosen arbitrary, this holds for all endpoints of chains originating at a. But then any member of those chains must satisfy ϕ as we prove by induction: for suppose all arguments b reachable from a through a chain of length n satisfy ϕ. Then all arguments b' reachable from a through a chain of length $n-1$ satisfy $\overleftarrow{\Box}\phi$. But then by assumption $\mathcal{N}, b' \models \phi$ as desired. Hence $\mathcal{N}, a \models \phi$ as desired. $\Rightarrow$: Suppose for any $\mathcal{N}$ based on $\mathcal{A}$, $\mathcal{N} \models \Box_U(\overleftarrow{\Box}\phi \to \phi) \to \phi$. Suppose for contradiction there an argument $a \in A$ such that an infinite ascending chain of attacks originates from a. Choose $\mathcal{N}$ such that $\mathcal{N}, a' \models \neg\phi$ for any a' in the chain. Then, since there is a'' such that $a' \twoheadrightarrow a''$ and $\mathcal{N}, a'' \models \neg\phi$, $\mathcal{N}, a' \models \neg\overleftarrow{\Box}\phi$. Hence $\mathcal{N}, a' \models \overleftarrow{\Box}\phi \to \phi$. But by supposition $a' \models \neg\phi$ and we get a contradiction.

16.2: $\Leftarrow$: Suppose $\mathcal{A}$ is cycle-free and let $\mathcal{N}$ be based on $\mathcal{A}$ and $a \in A$. Either $D_i = \{a'\}$ for some $a' \in A$ or $D_i = \{\}$. In the latter case, trivially $\mathcal{N}, a \models \neg @_i \Diamond^+ i$. In the first case, by cycle-freeness, a' does not reach itself. But then $\mathcal{N}, a \models \neg @_i \Diamond^+ i$ as desired. $\Rightarrow$: Suppose for any $\mathcal{N}$ based on $\mathcal{A}$, $\mathcal{N} \models \neg @_i \Diamond_{odd}\, i\}$. Suppose for contradiction $\mathcal{A}$ had a cycle. Pick any member $a \in \mathcal{N}$ of that cycle. Choose $\mathcal{N}$ such that $D_i = a$. Since a is in a cycle, it can reach itself. Hence $\mathcal{N}', a \models \Diamond^+ i$ and we reach a contradiction.

The proofs of 16.3 & 16.4 are analogous to 16.2.

16.5: Suppose for any $\mathcal{N}$ based on $\mathcal{A}$, $\mathcal{N} \models @_i i$. Then i is assigned in any $\mathcal{N}$. Now suppose for contradiction, $\mathcal{A}$ were finite. For any $\mathcal{N}$, every $a \in \mathcal{A}$, we have that $a \in D_j$ for some j since $\mathcal{N}$ is a named frame. Since $\mathcal{A}$ is finite for any $\mathcal{N}$ a finite number of nominals suffices for this. Hence we can find a named frame $\mathcal{N}'$ based on $\mathcal{A}$ such that i is not assigned. Contradiction! Hence $\mathcal{A}$ must be infinite.

$\square$

Expressiveness on the level of sets corresponds to truth conditions on the level of models. This is the level at which we express extensions of specific argumentation frameworks. Here we obtain the following results:

Proposition 17. Let $\mathcal{A}$ be a finite argumentation framework and $X \subseteq A$ be a set of arguments. X is

[2]For a necessary and sufficient condition, additional expressiveness is needed. Namely, we need to be able to quantify over the nominals. This allows for the condition $\mathcal{N}, a \models \exists i.\neg @_i i$. It is easy to see that this condition characterises finite frames in a language with countably many nominals. However, it can be shown that adding the existential quantifier to our language (with predicates) would give it the full expressive power of first-order logic and thus render it undecidable Areces and ten Cate [2007]Blackburn and Seligman [1996].

1. conflict-free iff $\mathcal{N}, a \vDash CF(\check{X})$ where $CF(\check{X}) \coloneqq \Box_U(\check{X} \to \neg \Diamond \check{X})$
2. self-acceptable iff $\mathcal{N}, a \vDash SA(\check{X})$ where $SA(\check{X}) \coloneqq \Box_U(\check{X} \to \Box \Diamond \check{X})$
3. admissible iff $\mathcal{N}, a \vDash AD(\check{X})$ where $AD(\check{X}) \coloneqq \Box_U((\check{X} \to \neg \Diamond \check{X}) \wedge (\check{X} \to \Box \Diamond \check{X}))$
4. complete iff $\mathcal{N}, a \vDash CO(\check{X})$ where $CO(\check{X}) \coloneqq \Box_U((\check{X} \to \neg \Diamond \check{X}) \wedge (\check{X} \leftrightarrow \Box \Diamond \check{X}))$
5. grounded iff $\mathcal{N}, a \vDash GR(\check{X})$ where $GR(\check{X}) \coloneqq \Box_U \check{X} \leftrightarrow \Box_{odd} \Diamond_{odd} \neg \Diamond \top$
6. preferred iff $\mathcal{N}, a \vDash PR(\check{X})$ where $PR(\check{X}) \coloneqq CO(\check{X}) \wedge \bigwedge_{X' \subseteq A, X' \neq X} CO(\check{X}') \to \Box_U \neg(\check{X} \to \check{X}')$
7. cycle-free iff $\mathcal{N}, a \vDash \Box_U \downarrow x.(\check{X} \to \neg \Diamond^+ x)$
8. odd-cycle-free iff $\mathcal{N}, a \vDash \Box_U \downarrow x.(\check{X} \to \neg \Diamond_{odd}\, x)$
9. even-cycle-free iff $\mathcal{N}, a \vDash \Box_U \downarrow x.(\check{X} \to \neg \Diamond_{even}\, x)$

Proof. The expressions for 17.1 - 17.4 carry over from K_U. 17.5: Recall that X is grounded iff it is the least fixpoint of the defense function (proposition 2). $\Rightarrow$ Suppose $\mathcal{N}, a \vDash \Box_U \check{X} \leftrightarrow \Box_{odd} \Diamond_{odd} \neg \Diamond \top$. Then $a \in X$ iff for all indirect attackers of a, there is an unattacked indirect attacker a'. Then a' is defended by the empty set. Since there is an even path from a' to a, there is an even number n such that $a \in d^n_{\mathcal{N}}(\varnothing)$. But then a must be in the least fixpoint of $d_{\mathcal{N}}$ as desired. $\Leftarrow$ Suppose X is the least fixpoint of $d_{\mathcal{N}}$ and $\mathcal{N}, a \vDash \check{X}$. Then $X = \bigcup_{0 \le n \le \omega} d_{\mathcal{N}}(\varnothing)$. But then $a \in X$ iff a is defended by the empty set or indirectly perfectly defended by an argument a' defended by the emptyset. In the first case, trivially $\mathcal{N}, a \vDash \Box_{odd} \Diamond_{odd} \neg \Diamond \top$. In the second case, there exists an unattacked indirect attacker for any indirect attacker of a. But then $\mathcal{N}, a \vDash \Box_{odd} \Diamond_{odd} \neg \Diamond \top$ as desired.

For 17.6, the proof is similar to the proof of proposition 6.2.

17.7 to 17.9 follow immediately from proposition 16.2 to 16.4 by considering the subframe induced by X.

□

Example 7. We have already seen how names work for the attack model $\mathcal{M}_1$. Now consider instead a named argumentation framework $(\mathcal{N}, \{D_i\}_{i \in ASSIGNED}$ where $ASSIGNED = \{i, j, k\}$ and $D_i = \{a_1\}, D_j = \{a_2\}, D_k = \{a_3\}$. The reader can check the following:

- $\mathcal{A} \vDash \neg @_i \Diamond_{odd}\, i$ and is hence odd-cycle-free.
- $\mathcal{A} \nvDash \neg @_i \Diamond_{even}\, i$ and hence contains an even cycle, namely, we have $\mathcal{N} \vDash \neg @_i \Diamond_{odd}\, i$ and is hence odd-cycle-free

3.2.5 Expressiveness of HAL: The Infinite Case

We should note here that the expressions we found in the previous section do not work for infinite argumentation frameworks: disjunctions (conjunctions) over all subsets of the AF would then have infinitely many disjuncts. But such an infinite disjunction (conjunction) is not a formula of HAL language. In fact, infinite disjunctions (conjunctions) correspond to existentially (universally) quantified formulas. Since this quantification occurs both over individual states and subsets of the space, it is clear that most of the previous characterizations correspond to second order formulas in the infinite case. It is well known that on the level of frame definability, modal and first order language expressivity intersect with non-empty differences: some second order frame conditions are expressible in modal language while some first order conditions are not. It is then natural to ask which of the previous second-order formulas can be expressed in hybrid logic and how we need to extend the language of HAL for this purpose.

What we need are equivalents to the disjunctions and conjunctions over all arguments we used in the finite case. We have already hinted at these equivalents through the abbreviations we used in the last section. Namely, the extended language we will investigate is the language HAL enriched with the binder operator $\downarrow x.\phi$ and the global modality $\Box_U\phi$. where $x \in Nom$. Satisfaction for the new formulas is defined as follows:

- $\mathcal{N}, a \vDash \downarrow x.\phi$ iff $\mathcal{N}', a \vDash \phi$ where x is a "fresh" nominal, i.e. $x \notin Assigned$ and $\mathcal{N}' = (A, R, \{D_i\}_{i \in Assigned'})$ where $\{D_i\}_{i \in Assigned'} = \{D_i\}_{i \in Assigned \cup \{S_x\}}$ and $S_x = \{a\}$.
- $\mathcal{N}, a \vDash \Box_U \phi$ iff there is $a' \in A$ s.t. $\mathcal{N}, a' \vDash \phi$.

Employing this additional expressive power we can express some of the semantics covered in the previous section. On the level of sets we have that:

Proposition 18. Let $\mathcal{N}$ be an argumentation framework and $X \subseteq A$ be a set of arguments. X is

1. conflict-free iff $\mathcal{N}, a \vDash CF(\check{X})$ where $CF(\check{X}) := \Box_U(\check{X} \to \neg \Diamond \check{X})$
2. self-acceptable iff $\mathcal{N}, a \vDash SA(\check{X})$ where $SA(\check{X}) := \Box_U(\check{X} \to \Box \Diamond \check{X})$
3. admissible iff $\mathcal{N}, a \vDash AD(\check{X})$ where $AD(\check{X}) := \Box_U((\check{X} \to \neg \Diamond \check{X}) \land (\check{X} \to \Box \Diamond \check{X}))$
4. complete iff $\mathcal{N}, a \vDash CO(\check{X})$ where $CO(\check{X}) := \Box_U((\check{X} \to \neg \Diamond \check{X}) \land (\check{X} \leftrightarrow \Box \Diamond \check{X}))$
5. grounded iff $\mathcal{N}, a \vDash GR(\check{X})$ where $GR(\check{X}) := \Box_U \check{X} \leftrightarrow \Box_{odd} \Diamond_{odd} \neg \Diamond \top$
6. cycle-free iff $\mathcal{N}, a \vDash \Box_U \downarrow x.(\check{X} \to \neg \Diamond^+ x)$
7. odd-cycle-free iff $\mathcal{N}, a \vDash \Box_U \downarrow x.(\check{X} \to \neg \Diamond_{odd}\, x)$

8. even-cycle-free iff $\mathcal{N}, a \models \Box_U \downarrow x.(\check{X} \rightarrow \neg \Diamond_{even}\, x)$

Proof. The propositions carry over from the finite case only that the semantics of the global modality and binder are now adjusted to the infinite case. □

Chapter 4

A Logic for Argument-based Belief

4.1 Motivation

We have now studied various ways of choosing which arguments to accept from an argumentation framework. We are thus a step closer to helping the Atlantians to base all of their beliefs on the outcome of rational argumentations. But one ingredient is missing: so far we have only talked about arguments. But the Atlantians are not interested in believing in arguments, they are interested in beliefs about the *claims* the arguments make. To help them with this we need to provide them with a language for those claims and appropriately connect it to our argumentation logic HAL. This is what we will set out to do in this section.

4.2 Representing Claims

Intuitively, an argument is a device that aims to convince an agent of the truth of a *proposition* by providing support for it. There are many formal languages we could use to represent these propositions. However, for simplicity's sake we will use basic propositional language. In addition, there is certain condition when a persuasion attempt can never succeed: the proposition is false and the agent *knows* this. Knowledge is usually modelled using the modal logic $S5$. Its language is the following:

Definition 18 ($S5$-Language)**.** Let P be a set of atoms. Our language is given by

$\mathcal{L}(P) : \phi ::= p \mid \bot \mid \neg\phi \mid \phi \wedge \phi \mid K\phi$
where $p \in P$.

Semantically, $S5$ is the logic of the class of frames whose relation is an equivalence relation. However, we will go a step further and interpret $S5$ over a global relation. This essentially yields a version of the logic K_U we discussed earlier without a local modality.

Definition 19 ($S5$ Semantics)**.** Let W be a set of worlds and define a function $V_{prop} : P \longrightarrow 2^W$. Then $\mathcal{M} = (W, V)$ is a an $S5$-model where satisfaction is defined as follows:

$\mathcal{M}, w \vDash p$ iff $V_{prop}(p) = w$
$\mathcal{M}, w \nvDash \bot$
$\mathcal{M}, w \vDash \neg\phi$ iff $\mathcal{M}, w \nvDash \phi$
$\mathcal{M}, w \vDash \phi \wedge \psi$ iff $\mathcal{M}, w \vDash \phi$ and $\mathcal{M}, w \vDash \psi$
$\mathcal{M}, w \vDash \Box\phi$ iff $\mathcal{M}, w' \vDash \phi$ for all $w' \in W$.

Finally, we extend V_{prop} to a function $V : \mathcal{L} \longrightarrow 2^A$ such that $V(\phi) = V_{prop}(\phi)$ if $\phi \in P$ and $V(\phi) := \{a \in A : \mathcal{M}, a \vDash \phi\}$ else.

These semantics allow us to view propositions as sets of worlds. The advantage of this is that frequently in epistemic logic, belief is modelled in terms of "spheres". I.e. each agent is associated with a set of worlds they deem most plausible. The K operator is interpreted as representing knowledge. Hence, a proposition is known iff it holds in all worlds the agent considers epistemically possible - that is in all worlds in W. A proposition is believed iff it is true in all worlds in that sphere, that is if it includes it. However, we still have to specify this sphere in terms of arguments. We will turn to this issue in the next section.

4.2.1 The Support Relation

Intuitively, an agent should believe in the propositions which are supported by arguments she accepts. But what does it mean for an argument to support a proposition P? Should the argument be a logical deduction with P as its conclusion? Clearly this requirement is too strong: most arguments are not valid deductions. Instead an argument could support P in many ways, for example by giving evidence in favor of it, by adding credibility to a source that asserts P or merely through rhethoric. In addition, assessing what makes for a *good* argument is non-trivial and arguably subjective. On the other hand, if we place no constraints on the support relation at all, then arguments do not constrain what the agent believes at all. For example the agent might have an accepted argument that supports both P and $\neg P$ and thus inconsistent beliefs. The approach commonly taken in abstract argumentation theory (especially in Logic Programming) is to model arguments as defeasible implications and support as the negation of defeaters (e.g. Dung [1995]). Then argument a attacks b if a supports a defeater for what b supports. In a modal setting with propositions interpreted as sets of worlds, this approach can be mimicked in the following way: we assume that each argument can be exhaustively described by a (consistent) set of propositions. Semantically then, the argument is the set of worlds where all of these propositions obtain (in what follows we shall blur the distinction between an argument and its semantics and refer to both simply by "argument"). Any proposition whose valuation includes the argument is regarded as supported. This resembles the view of arguments as defeasible implications in that the argument is an implication that holds at a subset of the possible worlds in the universe. Information that eliminates all those worlds can be viewed as a defeater for it. The question then is, how should an arguments semantic interpretation and the attack relation constrain each other? Shi, Smets, and Velázquez-Quesada [2017] seek to add argumentative structure to topological evidence models. I.e., given a topology over the universe, they interpret the open sets as arguments and aim at inducing an attack relation over them. For this purpose, they stipulate three requirements:

(A) For arguments a, b, a attacks b or b attacks a iff $a \cap b = \varnothing$.

(B) Let a, b, c be arguments. If a attacks b and $c \subseteq b$ then a attacks c.

(C) All arguments attack the empty argument and the empty argument attacks nothing.

The intuitive interpretation of (A) is that arguments are in conflict with each other iff there is no possible world in which both arguments hold. Then the whichever of the two arguments is judged to be the better argument should attack the other one. In addition, if an argument a is a subset of an argument b semantically it can be understood as a "strengthening" of that argument: syntactically, it is a superset of a thus having more demanding requirements. At the same time it supports more propositions making it more general. According to (B), an attacker of a general argument should also attack any more specific version of that argument. Finally, (C) demands that inconsistent arguments should be exposed in the sense that every argument attacks them and they do not succeed in attacking anything but themselves.

Now, we have the reverse goal as Shi et al. [2017], namely given an argumentation framework, assign semantics to arguments in terms of sets of worlds that allow us to define a notion of belief. In this setting slightly different conditions seem warranted. Firstly, it is easy to come up with examples where two arguments attack each other although there are worlds in which both of them are correct.

Example 8. Suppose Alice sees another news report stating that:

- a_3: The oracle of Delphi confirms that Atlantis will be inundated but due to the wrath of the gods.

Alice is not sure what to make of this. Clearly, the oracle is contradicting the scientists' report. Thus it seems like either the oracle or the scientists must be wrong. Or could it be the case that both of them are right? Maybe the gods will engineer the downfall of Atlantis through exactly the mechanism described by the Atlantian Scientific Society?

In light of this, it seems wise to require only half of (A) in our setting:

(A') a attacks b or b attacks a if $a \cap b = \varnothing$.

Secondly, in addition to (B), it seems also plausible that an argument should attack everything attacked by a weaker version of it.

Example 9. Later that day Alice reads an in depth piece on the Delphian oracle's prediction. She reads that:

- a_3': Not only did the oracle of Delphi confirm that Atlantis will be inundated and that this is due to the wrath of the gods but also it predicted that the catastrophe will occur in a single day.

Now this strengthened version of the argument is incompatible with the scientists' prediction which describes a gradual process taking place over years.

Therefore we strengthen (B) in the following way:

(B') Let a, b, c be arguments. If a attacks b and $c \subseteq b$ then a attacks c. Additionally, if a attacks b and $c \subseteq a$, then c attacks b.

Finally, we have to decide how to deal with inconsistent arguments, i.e. arguments that support only the empty set. Shi et al. [2017]'s condititión (C) is very much an artifact of the evidence topology: every topology includes the empty set, hence one better had a way of dealing with it. Coming from an argumentation framework however, the situation is less clear: most frameworks will not contain an argument that is self-attacking and attacked by all other arguments. In addition, condition (A') allows for non-empty, self-attacking arguments. Empirically speaking it appears plausible that people sometimes put forward inconsistent arguments due to the bounds of their rationality or for strategic reasons. Such arguments are not always recognised as inconsistent. Even if they were, it appears strange that any other argument would attack such an inconsistent argument. Rather, an argument that clearly points out the inconsistency would be needed. On the other hand condition (A') requires that there be an attack between disjoint arguments and the empty set is disjoint from any other argument. Thus it seems what would be needed here to dissolve the dilemma would be a notion of *impossible worlds*, that is worlds that allow for contradictions. An inconsistent argument would then support a set of impossible worlds. There is a long tradition of research into modal logics allowing for impossible worlds (see Berto [2013]), but we leave this topic to future work. For present purposes we will forgo dealing with inconsistent arguments by giving agents the benefit of the doubt and assuming their arguments to be consistent. In other words, we require that every argument supports some non-empty set of worlds:

(C') For every argument a, $a \neq \varnothing$.

4.3 Putting Things Together: Argument-Support Logic

Now that we have developed a language to characterize argumentation frameworks, a language for argument claims and a notion of support it remains to combine these ingredients in a unified framework. In the literature, there are two approaches to this problem: Shi et al. [2017] develop an approach based on topological semantics. In contrast, Grossi and van der Hoek [2014] use product logics to combine logic for argumentation with a logic for argument claims. In Shi et al. [2017]'s topological semantics arguments are the open sets of a topology over a space of epistemically possible worlds. As usually, formulas are evaluated locally at individual worlds. This setting is difficult to combine with a hybrid language for arguments: if an argument refers to a set of worlds the naming character of nominals is violated. Furthermore, as arguments are not necessarily disjoint, *Sym* will no longer hold: $@_i j \not\leftrightarrow @_j i$. Thus what we need are semantics that preserve the hybrid character of HAL on the one hand and allow for the interpretation of arguments as sets of worlds on the other. Intuitively, we need the ability to distinguish between world w when considering argument a and w when considering argument a'. I.e., we need an evaluation point for every argument-world *combination*. In other worlds our semantics should be based on the Cartesian product of the set of arguments and epistemic possibilities. This approach is explored by Grossi and van der Hoek [2014] under the name *Doxastic Argument Logic* (DA). However, in DA both arguments *and* belief are modelled as primitive. That is, semantically an argumentation framework is combined with a doxastic KD45-frame with a "readymade" doxastic accessibility relation. Contrary to this we seek to define belief in terms of argumentation. For this purpose we shall develop a logic for argument supported belief.

4.3.1 Product Logics

We present here a HAL-variant of the product logic DA presented by Grossi and van der Hoek [2014]. As before, we start with semantics. In general, a product frame is defined as follows.

Definition 20 (Product Frame)**.** [Gabbay and Shehtman [1998]] Let $\mathcal{A} = \{A, R\}$, $\mathcal{F} = \{W, R'\}$ be frames. Then the corresponding product frame is defined as $\mathcal{D}_A = \{A \times W, Q, S\}$ where

$$(a,w)Q(b,v) \text{ iff } aRb \text{ and } w = v$$
$$(a,w)S(b,v) \text{ iff } wR'v \text{ and } a = b$$

We extend this definition to classes of frames in the natural way: let $\mathfrak{F}$, $\mathfrak{F}'$ be classes of frames. Then their product class is defined as $\mathfrak{F} \times \mathfrak{F}' = \{\mathcal{F} \times \mathcal{F}' : \mathcal{F} \in \mathfrak{F}, \mathcal{F}' \in \mathfrak{F}'\}$. In turn, the logic of a product class $\mathfrak{F} \times \mathfrak{F}'$ is denoted by $L \times L'$ where L, L' are the logics of $\mathfrak{F}, \mathfrak{F}'$ respectively. We will discuss how to axiomatize such a logic in section 4.3.2. Then it seems like the logic $HAL \times S5$ (where

$S5$ is interpreted over a global relation) as a good candidate for our purposes. However, we still need to build in the support relation. On the language side, following Grossi and van der Hoek [2014] we will do this by adding a special operator σ. The logic $(HAL \times S5)_\sigma$ has the following language:

Definition 21 (Argument-Support Logic: Language)**.** Let P be a set of atoms, Nom be a set of nominals. The language of Hybrid Argument Support Logic is defined as follows:

$$\mathcal{L}(P) : \phi ::= i \mid p \mid \bot \mid \neg\phi \mid \phi \wedge \phi \mid \boxplus\phi \mid \boxplus^{+}_{odd}\phi \mid @_i\phi \mid \overline{K}\phi \mid \sigma$$

where $p \in P$ is an atom, $i \in Nom$ is a nominal.

σ denotes the nullary support operator. That is, $(a, w) \vDash \sigma$ is to be interpreted as "argument a supports world w" (or vice versa). Correspondingly we will now define product frames equipped with a unary relation to interpret σ.

Definition 22. An Argument-Support frame is a tuple $(\mathcal{N} \times W, R_\sigma)$ where $\mathcal{N}$ is a named argumentation framework, W a non-empty set and R_σ is a subset of $A \times W$. R_σ is constrained to respect the conditions (A') to (C') delineated in section 4.2.1:

- (A'): For all $a, b \in A$, if $\{w : (a, w) \in R_\sigma\} \cap \{w : (b, w) \in R_\sigma\} = \varnothing$ then $a \twoheadrightarrow b$ or $b \twoheadrightarrow a$.
- (B').1: For all $a, b, c \in A$, if $a \twoheadrightarrow b$ and $\{w : (c, w) \in R_\sigma\} \subseteq \{w : (b, w) \in R_\sigma\}$ then $a \twoheadrightarrow c$.
- (B').2: For all $a, b, c \in A$, if $a \twoheadrightarrow b$ and $\{w : (c, w) \in R_\sigma\} \subseteq \{w : (a, w) \in R_\sigma\}$ then $c \twoheadrightarrow b$.
- (C'): For all $a \in A$ there is $w \in W$ such that $(a, w) \in R_\sigma$

Now let $(\mathcal{N} \times W, R_\sigma)$ be an Argument-Support frame. Then we gain a model by equipping it with a valuation function $V_{base} : P \cup Nom \longrightarrow 2^{A \times W}$ such that:

- $V_{base}(i) = \{(a, w) : a \in D_i, w \in W\}$ for any $i \in Nom$;
- $V_{base}(p) = \{(a, w) : w \in V_{Prop}(p)\}$ for any $p \in Prop$;

Then we define satisfaction as follows:

- $\mathcal{AS}, (a, w) \vDash p$ iff $(a, w) \in V_{base}(p)$
- $\mathcal{AS}, (a, w) \vDash i$ iff $(a, w) \in V_{base}(i)$
- $\mathcal{AS}, (a, w) \vDash \sigma$ iff $(a, w) \in R_\sigma)$
- $\mathcal{AS}, (a, w) \vDash \neg\phi$ iff $\mathcal{AS}, (a, w) \nvDash \phi$.
- $\mathcal{AS}, (a, w) \vDash \phi \wedge \psi$ iff $\mathcal{AS}, (a, w) \vDash \phi$ and $\mathcal{AS}, (a, w) \vDash \psi$.

$\mathcal{AS}_1$	w_1	w_2	w_3	w_4	Ext
a_1	σ			σ	AD_1, CO_1, PR_1
a_2		σ	σ		AD_2, CO_2, PR_2
a_3	σ	σ			AD_3, CO_2, GR, PR_2
	$cure$, $rise$	$cure$, $\neg rise$	$\neg cure$, $\neg rise$	$\neg cure$, $rise$	

Figure 4.1: Alice's orginal Argument-Support Model $\mathcal{AS}_1$ with admissible extensions AD_1, AD_2, complete extensions CO_1, CO_2, preferred extensions PR_1, PR_2 and grounded extension GR.

- $\mathcal{AS}, (a, w) \vDash \boxbar\phi$ iff for all $a' \in A$ if $a \twoheadleftarrow a'$ then $\mathcal{AS}, (a', w) \vDash \phi$
- $\mathcal{AS}, (a, w) \vDash \boxbar_{odd}\phi$ iff for all $a' \in A$ s.t. $a \twoheadleftarrow_{odd} a'$ and $\mathcal{AS}, (a', w) \vDash \phi$ where $\twoheadleftarrow_{odd}$ is the odd transitive closure of $\twoheadleftarrow$.
- $\mathcal{AS}, (a, w) \vDash @_i\phi$ iff there is a' such that for all w, $(a', w) \in V_{base}(i)$ and $\mathcal{AS}, (a', w) \vDash \phi$.
- $\mathcal{AS}, (a, w) \vDash \overline{K}\phi$ iff $\mathcal{AS}, (a, w') \vDash \phi$ for all $w' \in W$.

Note that each nominal denotes a single argument and akin to "rigid designators", points to that arguments across all worlds. Hence the valuation of a nominal is now a set of argument-world pairs where the argument stays the same. Correspondingly, the satisfaction-operator - since it now refers to a set - demands that every argument-world pair satisfy the proposition in its scope (recall, that unlike in standard hybrid logic, not every nominal needs to be assigned). It is thus, in a way, a two-dimensional operator. In constrast, the modal operators are one-dimensional in that each modal operator "looks" along one axis of the product space only. I.e. $\overline{K}$ looks along worlds while $\boxbar$ looks along arguments. In addition to the above language, we also import all the abbreviations we defined for HAL. That is $\boxbar_U\phi$, $\downarrow x.\phi$, $\boxbar_{even}\phi$ are all defined as before with respect to the HAL component of the language.

Example 10. Figure 4.1 shows the $\mathcal{AS}$-model representing Alice's argumentation framework from Example 2 featuring in a product logic enriched with the support operator σ and atoms $cure$ and $rise$ for the claims made by the arguments (see the lower margins).

4.3.2 A Note on Axiomatization

Our conjecture is that Argument Support Logic can be axiomatized by the axiomatization of HAL for the vertical logic, the axiomatization of $S5$ for the horizontal logic and the following axioms:

$$\Diamond - Com \qquad \langle \overline{K} \rangle \diamondsuit^* \phi \leftrightarrow \diamondsuit^* \langle \overline{K} \rangle \phi \tag{4.1}$$

$$\Diamond - Con \qquad \diamondsuit^* \overline{K} \phi \leftrightarrow \overline{K} \diamondsuit^* \phi \tag{4.2}$$

$$(A') \qquad @_i \overline{K}(\sigma \rightarrow \boxminus_U (j \rightarrow \neg\sigma) \rightarrow @_i \diamondsuit j \vee @_j \diamondsuit i \tag{4.3}$$

$$(B'.1) \qquad @_i \diamondsuit k \wedge @_j \overline{K}(\sigma \rightarrow \boxminus_U (\sigma \rightarrow i) \rightarrow @_j \diamondsuit k \tag{4.4}$$

$$(B'.2) \qquad @_i \diamondsuit k \wedge @_j \overline{K}(\sigma \rightarrow \boxminus_U (\sigma \rightarrow k) \rightarrow @_i \diamondsuit j \tag{4.5}$$

$$(C') \qquad \langle \overline{K} \rangle \sigma \tag{4.6}$$

where $\diamondsuit^* \in \{\diamondsuit, \diamondsuit_{odd}\}$ and $\boxminus^*$ is defined accordingly. Here $\Diamond - Com$ and $\Diamond - Con$ are the axioms that govern the interactions between horizontal and vertical modalities (see Carnielli and Coniglio [2016], Gabbay and Shehtman [1998], section 7)). The axioms $(A') - (C')$ correspond to the conditions we placed on R_σ. For a large class of logics, completeness transfers from the component logics to the product logic (see Gabbay and Shehtman [1998], section 7). Namely, this is the case for normal modal logics axiomatized by axioms that either satisfy a condition called "pseudotransitivity" or do not contain propositional letters. It is known that the $S5$-axioms are pseudotransitive. In addition note that all axioms of HAL are free of propositional letters (since it is a frame language). So are the axioms $(A') - (C')$ but they are not part of the component logics. This suggests that the above axiomatization should be complete with respect to Argument-Support frames. However, we leave it to future work to prove this.

4.4 Belief in Terms of Argument Semantics

Argument-Support Logic provides us with a powerful language to express different notions of belief. In this section we will explore three such notions and investigate their properties when using different argument semantics as belief-makers.

Namely, we make the following distinctions:

Definition 23. Given a semantic S, we say that a proposition ϕ is

- credulously believed iff there is some argument a, such that $a \in E$ for some extension $E \in S$ and ϕ holds at every world that is supported by a (or to present it as a pseudo-second order expression:

$$\exists a, E.(E(a) \wedge S(E) \wedge \forall w.(\sigma(a,w) \rightarrow \phi(a)))).$$

- skeptically believed iff there is some argument a, such that $a \in E$ for all extensions $E \in S$ and ϕ holds at every world that is supported by a (or

$$\exists a. \forall E.((S(E) \rightarrow E(a)) \wedge \forall w.(\sigma(a,w) \rightarrow \phi(a)))).$$

- strongly believed iff for all arguments a, such that $a \in E$ for some extension $E \in S$, ϕ holds at every world that is supported by a (or

$$\forall a.\exists E.((E(a) \wedge S(E)) \rightarrow \forall w.(\sigma(a,w) \rightarrow \phi(a)))).$$

Let ψ be the formula scheme that characterizes S as per proposition 17[1] and let X be a set of arguments. Then, for finite argumentation frameworks, we can express these three notions of belief as follows in Argument-Support logic:

$$\begin{array}{ll} \textit{Credulous Belief} & \overline{CB}_S\phi := \Diamond\!\!\!\mid_U(\bigvee_{X\subseteq A}(\check{X} \wedge \psi(\check{X})) \wedge \overline{K}(\sigma \rightarrow \phi)) \\ \textit{Skeptical Belief} & \overline{SB}_S\phi := \Diamond\!\!\!\mid_U(\bigwedge_{S\subseteq A}(\psi(\check{X}) \rightarrow \check{X}) \wedge \overline{K}(\sigma \rightarrow \phi)) \\ \textit{Strong Belief} & \overline{STB}_S\phi := \boxminus_U(\bigvee_{S\subseteq A}(\check{X} \wedge \psi(\check{X})) \rightarrow \overline{K}(\sigma \rightarrow \phi)) \end{array}$$

Comparing these formulas to the second order expressions in Definition 23 is instructive: the vertical global modality serves as an existential/universal quantifier over arguments. The second order quantifier is represented by a big disjunction over all subsets of A. Finally and most interestingly, $\bar{K}$ takes the role of the universal horizontal quantifier. In other words, credulous and skeptical belief require that it is known that the considered argument supports the proposition in question whereas strong belief requires that this be the case for all accepted arguments.

As the reader may have noticed, credulous and skeptical belief follow the distinction between credulous and skeptical acceptance of arguments we encountered in chapter 2. We should note that on finite frames, "grounded belief" as proposed by Shi et al. [2017] is similar to credulous belief where S is the grounded semantic (not equivalent, as we shall see). On the other hand strong belief captures a situation when the agent has maximal certainty in a proposition: all the arguments she is aware of and willing to accept support the proposition. Another noteworthy property of these belief operators is that they are global. That is $\mathcal{AS},(a,w) \vDash \overline{CB}_S\phi$ ($\overline{SB}_S\phi$, $STB_S\phi$ respectively) iff $\mathcal{AS} \vDash \overline{CB}_S\phi$ ($\overline{SB}_S\phi$, $STB_S\phi$ respectively). This is so because they quantify over all arguments and the "row properties" bearing on the evaluating argument hold independently of the evaluation world.

Example 11. The reader is invited to check the following in Figure 4.1:

- Alice believes that bloodletting cures the cold credulously under admissible semantics complete semantics and preferred semantics.
- She believes that bloodletting cures the cold skeptically under grounded and preferred semantics but not under admissible or complete semantics.

[1] Strictly speaking, the version of the formula in the language of ASL, i.e. every $\Diamond$ has to be replaced by a $\Diamond\!\!\!\mid$ and so forth.

- She believes that bloodletting cures the cold strongly under grounded semantics but not under admissible, complete or preferred semantics.

We shall now study some general properties of credulous, skeptical and strong belief before turning to their behaviour under specific semantics. First, we should take note of the duals of the belief operators.

Definition 24. (Duals) Given an ASL formula ϕ denote by $\langle\overline{CB}\rangle\phi, \langle\overline{SB}\rangle\phi, \overline{STB}\phi$ respectively the formuals $\neg\overline{CB}\neg\phi$, $\neg\overline{SB}\neg\phi$, $\neg\overline{STB}\neg\phi$ which correspond to

- $\mathcal{AS} \vDash \langle\overline{CB}\rangle_S\phi \leftrightarrow \boxdot_U(\bigwedge_{X\subseteq A}(\check{X} \to \neg\psi(\check{X})) \vee \langle\overline{K}\rangle(\sigma \wedge \phi))$
- $\mathcal{AS} \vDash \langle\overline{SB}\rangle_S\phi \leftrightarrow \boxdot_U\phi(\bigvee_{X\subseteq A}(\psi(\check{X}) \wedge \neg\ \check{X}) \vee \langle\overline{K}\rangle(\sigma \wedge \phi))$
- $\mathcal{AS} \vDash \langle\overline{STB}\rangle_S\phi \leftrightarrow \Diamond_U(\bigvee_{X\subseteq A}(\check{X} \wedge \psi(\check{X})) \wedge \langle\overline{K}\rangle(\sigma \wedge \phi))$

Usually, the duals of belief operators are interpreted as the agent not ruling out a proposition or considering it a "serious possibility". However, in the case of credulous and skeptical belief this interpretation is unsuitable: as well shall see in the next section, depending on the semantic it is possible for an agent to believe both a proposition ϕ *and* its negation $\neg\phi$. As a consequence belief in $\neg\phi$ is not enough to rule out ϕ. Another difference between credulous and skeptical belief on the one hand and strong belief on the other is that the latter quantifies universally over arguments. As a consequence, if the only extension is the empty set, the agent trivially strongly believes everything. This is not the case for the existentially quantified credulous and skeptical belief. Thus the intuition that strong belief should imply the weaker forms of belief holds only in the case where there are non-empty extensions. More curiously, it is even possible to know something and yet to (credulously, skeptically) believe nothing - again this is the case if no argument is accepted. One could compare this situation to the extreme skepticism explored by Descartes: suppose the agent is not willing to accept any argument as it may only appear convincing to her due to the works of an evil demon. As a consequence she believes nothing. Yet it is not possible to know nothing as there are some things one knows, such as that one exists, that cannot be given up. This is also the point where Credulous Grounded Belief differs from Grounded Belief as defined by Shi et al. [2017]: Since the arguments in their setting are the open sets in a topology, the whole space is always an argument and it is always accepted. This ensures that what is known is also believed. Coming from an argumentation framework, there is no such guarantee. Finally, it is easy to see that if there is a unique extension, credulous and skeptical belief become equivalent. We summarise these facts in the following proposition:

Proposition 19. Let $\mathcal{AS}$ be a finite argument-support model and S a semantic. Then

1. $\mathcal{AS} \vDash K\phi \to \overline{STB}_S\phi$
2. For all $E \in S$, $E \neq \varnothing$ implies $\mathcal{AS} \vDash \overline{STB}_S\phi \to \overline{SB}_S\phi \to \overline{CB}_S\phi$

3. $|S| = 1$ implies $\mathcal{AS} \vDash \overline{CB}_S\phi \leftrightarrow \overline{SB}_S\phi$.

Proof. 19.1: Suppose $AS \vDash \phi$. Then for all $(a,w) \in A \times W$, $AS, (a,w) \vDash \phi$. But then a fortiori $AS, (a,w) \vDash \sigma \to \phi$ and thereby thereby $AS \vDash \overline{K}(\sigma \to \phi)$. This holds for all arguments and by weakening we get $AS \vDash \boxplus_U(\bigvee_{S\subseteq A}(\check{X} \wedge \psi(\check{X})) \to \overline{K}(\sigma \to \phi))$ as desired.

19.2: Suppose $E \neq \varnothing$ and $\mathcal{AS} \vDash \overline{STB}_S\phi$. Then there is at least one accepted argument and for every accepted argument a, $\mathcal{AS}, (a,w) \vDash \overline{K}(\sigma \to \phi))$ where W is any world. But then it follows that $(a,w) \vDash \bigwedge_{X\subseteq A}(\psi(\check{X}) \to \check{X})$. Hence $\mathcal{AS} \vDash \overline{STB}_S\phi \to \overline{SB}_S\phi$. Now clearly $(a,w) \vDash \bigwedge_{X\subseteq A}(\psi(\check{X}) \to \check{X})$ implies that $(a,w) \vDash \bigvee_{X\subseteq A}(\psi(\check{X}) \wedge \check{X})$ and hence $\mathcal{AS} \vDash \overline{STB}_S\phi \to \overline{SB}_S\phi \to \overline{CB}_S\phi$ as desired.

19.3,⇒: Suppose $|S| = 1$ and $\mathcal{AS} \vDash \overline{CB}_S\phi$. Let E be the unique extension in S. Then for some argument s.t $a \in E$ and $AS, (a,w) \vDash \overline{K}(\sigma \to \phi)$. But since E is the unique extension we have that for all $E \in S$, $a \in E$ and hence $\mathcal{AS} \vDash \overline{SB}_S\phi$ as desired. ⇐: similar.

□

Next we study how our notions of belief compare to a standard doxastic logic with an $S5$-knowledge operator and a $KD45$-belief operator. In doing so we consider the following axioms where B stands for a generic belief operator:

- N: $K\phi \to B\phi$
- D: $B\phi \to \langle B\rangle\phi$
- M: $B(\phi \wedge \psi) \to B\phi \wedge B\psi$
- C: $B\phi \wedge B\psi \to B(\phi \wedge \psi)$
- 4: $B\phi \to B(B\phi)$
- 5: $\neg B\phi \to B(\neg B\phi)$

Axiom N demands that knowledge imply belief. We have already seen that this axiom only holds for strong belief in our setting. M and C together are equivalent to axiom K. Finally D, 4, and 5 are the usual requirements on belief: D enforces consistency, 4 positive introspection and 5 negative introspection. Grounded belief as defined by [Shi et al., 2017] satisfies all of these axioms except C. We will now study which fragment of the standard axiomatization each of credulous, skeptical and strong belief fulfill.

4.4.1 Credulous Belief

We begin with credulous belief and investigate its behaviour when plugging in various semantics for S. Our focus is on the admissible, complete, grounded and preferred semantics. Except for the grounded semantics, an argumentation framework may have multiple extensions under any of these semantics. For credulous belief, this has the consequence that agents can be locally inconsistent in the sense that they can believe both a proposition and its negation. However, they are saved from full-blown inconsistency by the fact that there beliefs need not be closed under conjunction. A possible interpretation of Credulous Belief under non-unique argument semantics is the idea that people may have several, mutually incompatible yet coexisting mental models that allow them to have context-dependent beliefs. Under this interpretation, extensions play the role of mental models with consistency required within each extension but not between extensions. The result is a notion of bounded rationality that limits the degree to which agents check the consistency of their beliefs: in the case of admissible semantics, they only check that their model can defend itself and is internally consistent; if completeness is required, agents are assumed to be able to also derive all the argumentative consequences of their models; finally in the preferred case, the agents look for and are able to obtain maximally complete models.

Proposition 20. Let $\mathcal{AS}$ be an Argument-Support model and let $S \in \{\mathcal{AD}, \mathcal{CO}, \mathcal{PR}\}$. Then we have that

1. N: $\mathcal{AS} \nvDash K\phi \rightarrow \overline{CB}_S\phi$
2. D: $\mathcal{AS} \nvDash \overline{CB}_S\phi \rightarrow \langle\overline{CB}\rangle_S\phi$
3. M: $\mathcal{AS} \vDash \overline{CB}_S(\phi \wedge \psi) \rightarrow \overline{CB}_S\phi \wedge \overline{CB}_S\psi$
4. C: $\mathcal{AS} \nvDash \overline{CB}_S\phi \wedge \overline{CB}_S\psi \rightarrow \overline{CB}_S(\phi \wedge \psi)$
5. 4: $\mathcal{AS} \vDash \overline{CB}_S\phi \rightarrow \overline{CB}_S(\overline{CB}_S\phi)$
6. 5: $\mathcal{AS} \vDash \neg\overline{CB}_S\phi \rightarrow \overline{CB}_S(\neg\overline{CB}_S\phi)$

Proof. 20.1: Take as a counterexample any model where there is a unique empty extension such as the case of a single self-attacking argument.

20.2: Take as a counterexample any model where there are two extensions containing arguments supporting both ϕ and $\neg\phi$. See e.g. Figure 4.1.

20.3: Suppose $\mathcal{AS} \vDash \overline{CB}_S(\phi \wedge \psi)$. Then there exist a, $E \in S$ such that $a \in E$ and for all w, $\mathcal{AS}, (a, w) \vDash \sigma \rightarrow (\phi \wedge \psi)$. But then $\mathcal{AS}, (a, w) \vDash \sigma \rightarrow \phi$ and $\mathcal{AS}, (a, w) \vDash \sigma \rightarrow \psi$ as desired.

20.4: For a counterexample consider again 4.1. $\mathcal{AS}_1 \vDash \overline{CB}_S rise$ and $\mathcal{AS}_1 \vDash \overline{CB}_S \neg rise$ under admissible, complete and preferred semantics with a_1 supporting $rise$ and a_2 supporting $\neg rise$. But clearly there is no world w such that $\mathcal{AS}_1, (a, w) \vDash rise \wedge \neg rise$ for some a and hence $\mathcal{AS}_1 \nvDash \overline{CB}_S(rise \wedge \neg rise)$.

20.5: Suppose $\mathcal{AS} \vDash \overline{CB}_S\phi$. Then for all $(a, w) \in A \times W$ we have that $\mathcal{AS}, (a, w) \vDash \overline{CB}_S\phi$. This holds fortiori for all (a, w) such that $\mathcal{AS}, (a, w) \vDash \sigma$. But since there is some accepted argument that supports some world by assumption, there must be some accepted argument that supports $\overline{CB}_S\phi$. Hence $\mathcal{AS} \vDash \overline{CB}_S(\overline{CB}_S\phi)$ as desired.

20.6: The argument is analogous to the proof of 20.5.

□

If the semantics yield a unique extension as in the case of the grounded semantics, local inconsistency is not possible. However, closure under conjunction still fails. One might wonder what causes this. The reason becomes apparent when comparing the semantics of arguments in our setting to the setting of **?**. As we have noted, the arguments form a topology in their setting. This requires that every intersection of arguments is an argument. In our setting no such requirement is placed on the valuation of arguments. That is for any a, a', $\{w : (a,w) \cap (a',w)\}$ needn't correspond to an argument. Thus, even in a unique extension is not enough to ensure closure under conjunction.

Proposition 21. Let $\mathcal{N} \times \mathcal{F}$ be an argument-support frame and let $S = \mathcal{GR}$. Then we have that

1. N: $\mathcal{AS} \nvDash K\phi \to \overline{CB}_S\phi$
2. D: $\mathcal{AS} \vDash \overline{CB}_S\phi \to \langle\overline{CB}\rangle_S\phi$
3. M: $\mathcal{AS} \vDash \overline{CB}_S(\phi \wedge \psi) \to \overline{CB}_S\phi \wedge \overline{CB}\psi$
4. C: $\mathcal{AS} \nvDash \overline{CB}_S\phi \wedge \overline{CB}_S\psi \to \overline{CB}(\phi \wedge \psi)$
5. 4: $\mathcal{AS} \vDash \overline{CB}_S\phi \to \overline{CB}_S(\overline{CB}_S\phi)$
6. 5: $\mathcal{AS} \vDash \neg\overline{CB}_S\phi \to \overline{CB}_S(\neg\overline{CB}_S\phi)$

Proof. The proofs for $N, M, 4, 5$ are as in the proof of 20.

21.2: Suppose $\mathcal{AS} \vDash \overline{CB}_S\phi$. Then there is an argument a in the grounded extension such that for all w, $\mathcal{AS},(a,w) \vDash \sigma \to \phi$. Now suppose for contradiction there were an argument a' in the grounded extension such that for all w $\mathcal{AS},(a,w) \vDash \sigma \to \neg\phi$. Then from $V(\phi) \cap V(\neg\phi) = \varnothing$ it follows that $\{w : (a,w) \in R_\sigma\} \cap \{w : (a',w) \in R_\sigma\} = \varnothing$. Then by (B') $a \twoheadrightarrow a'$ or $a' \twoheadrightarrow a$. But then the grounded extension is not conflict-free and we get a contradiction. Hence $\mathcal{AS},(a,w) \nvDash \sigma \to \neg\phi$. It follows immediately that $\mathcal{AS} \nvDash CB_S\neg\phi$ and hence $\mathcal{AS} \vDash \langle\overline{CB}\rangle_S\phi$.

21.4: The counterexample from the proof of 20 fails now since the grounded extension cannot contain arguments supporting ϕ and $\neg\phi$ respectively (see the argument for 21.2). However, consider the example in Figure 4.2. Here we have in the grounded extension the arguments a_3 and a_4 which support *cure* and *rise* respectively. But neither of them supports *cure* and *rise*. ☐

Example 12. Alice trusts the Oracle of Delphi more than the Atlantian Scientific Society and so her new argumentation framework looks like in Figure 4.2.

$\mathcal{AS}_2$	w_1	w_2	w_3	w_4	Ext
a_1	σ			σ	
a_2		σ	σ		
a_3	σ	σ			AD_1, AD_2,CO, GR, PR
a_4	σ			σ	AD_2,AD_3,CO, GR, PR
	$cure$, $rise$	$cure$, $\neg rise$	$\neg cure$, $\neg rise$	$\neg cure$, $rise$	

Figure 4.2: Alice's Argument-Support Model after hearing the prophecy by the oracle of Delphi

4.4.2 Skeptical Belief

Under the skeptical interpretation of belief, the question is not so much whether the semantics admit multiple extensions but rather whether the intersection of those extensions is empty, a true intersection or equals the union of the extensions. In the first case, the agent will believe nothing: there is no argument that is in all the extensions. Consider for example admissible semantics: since the empty set is always admissible, the intersection of all admissible sets is always empty. Thereby under admissible semantics skeptical belief collapses: the agent will not skeptically believe anything.

Proposition 22. Let $\mathcal{AS}$ be an argument-support model and S the admissible semantic. Then we have that

$$\mathcal{AS} \vDash \neg\overline{SB}_S\phi$$

Proof. Suppose for contradiction $\mathcal{AS} \vDash \overline{SB}\phi$. Then there is an argument that is in all admissible sets. But the intersection of all admissible sets is empty. Hence we get a contradiction. □

On the other hand, if the intersection of the extensions equals their union skeptical belief collapses into credulous belief. This is the case if the extension is unique and thus for the grounded semantic:

Proposition 23. Let $\mathcal{AS}$ be an argument-support model and let S be the grounded semantic. Then we have that

$$\mathcal{AS} \vDash \overline{CB}_S\phi \leftrightarrow \overline{SB}_S\phi$$

Proof. This follows trivially from the observation that in the case of a unique extension, an argument being in all extensions is equivalent to it being in one extension. □

Finally, for the complete and preferred semantics, skeptical belief behaves similar to credulous belief under the grounded semantic. However, instead of one mental model we get several ones that are required to be "mutually consistent" with respect to the proposition in question. That is, skeptical belief enforces C even without a unique extension since it requires that each extension contains the same arguments supporting the proposition. By condition (C') of the support relation a proposition and its negation cannot both be supported by the same argument.

Proposition 24. Let $\mathcal{AS}$ be an argument-support model and $S \in \{\mathcal{CO}, \mathcal{PR}\}$ Then we have that

1. N: $\mathcal{AS} \nvDash K\phi \to \overline{SB}_S\phi$
2. D: $\mathcal{AS} \vDash \overline{SB}_S\phi \to \langle\overline{SB}\rangle_S\phi$
3. M: $\mathcal{AS} \vDash \overline{SB}_S(\phi \wedge \psi) \to \overline{SB}_S\phi \wedge \overline{SB}_S\psi$
4. C: $\mathcal{AS} \nvDash \overline{SB}_S\phi \wedge \overline{SB}_S\psi \to \overline{SB}_S(\phi \wedge \psi)$
5. 4: $\mathcal{AS} \vDash \overline{SB}_S\phi \to \overline{SB}_S(\overline{SB}_S\phi)$
6. 5: $\mathcal{AS} \vDash \neg\overline{SB}_S\phi \to \overline{SB}_S(\neg\overline{SB}_S\phi)$

Proof. 24.1: The same counterexamples as for 20.1 apply.

24.2: Suppose $\mathcal{AS} \vDash \overline{SB}_S\phi$. Then there is an argument a such that $a \in E$ for all extensions $E \in S$ and for all W, $\mathcal{AS}, (a, w) \vDash \sigma \to \phi$. Now the proof is analogous to the proof of 21.2 except that we consider all extensions under S.

24.3: Again the proof is analogous to the proof of 20.3, accounting for the universal quantification.

24.4: Consider again the counterexample presented in Figure 4.2.

19 and 19 follow as before from the globality of SB_S. □

4.4.3 Strong Belief

Finally, strong belief goes beyond credulous and skeptical belief in that it is implied by knowledge and enforces consistency and closure under conjunction yielding the classical KD45 notion of belief:

Proposition 25. Let $\mathcal{N} \times \mathcal{F}$ be an argument-support model and let $S \in \{\mathcal{AD}, \mathcal{CO}, \mathcal{GR}, \mathcal{PR}\}$. Then we have that

1. N: $\mathcal{AS} \vDash K\phi \to \overline{STB}_S\phi$
2. D: $\mathcal{AS} \vDash \overline{STB}_S\phi \to \overline{STB}_S\phi$
3. M: $\mathcal{AS} \vDash \overline{STB}_S(\phi \wedge \psi) \to \overline{STB}_S\phi \wedge \overline{STB}_S\psi$
4. C: $\mathcal{AS} \vDash \overline{STB}\phi \wedge \overline{STB}_S\psi \to \overline{STB}_S(\phi \wedge \psi)$
5. 4: $\mathcal{AS} \vDash \overline{STB}_S\phi \to \overline{STB}_S(\overline{STB}_S\phi)$
6. 5: $\mathcal{AS} \vDash \neg\overline{STB}_S\phi \to \overline{STB}_S(\neg\overline{STB}_S\phi)$

Proof. 25.1: Suppose $\mathcal{AS} \models K\phi$. Then for all $(a,w) \in A \times W$, $\mathcal{AS},(a,w) \models \phi$. A fortiori this holds for all (a,w) s.t. $\mathcal{AS},(a,w) \models \sigma$. But then it follows immediately that $\mathcal{AS} \models \overline{STB}_S\phi$ as desired.

25.2: As in the proofs of 21.2 and 24.2 this follows from the fact that two arguments supporting ϕ and $\neg\phi$ respectively cannot be in the same extension.

25.3: Almost the same as the proof of 20.3.

25.4:Suppose $\mathcal{AS} \models \overline{STB}\phi \wedge \overline{STB}_S\psi$. Then for any a, if $a \in E$ for some $E \in S$, then for all $w \in W$, $\mathcal{AS},(a,w) \models \sigma \to \phi$ and $\mathcal{AS},(a,w) \models \sigma \to \psi$. But then $\mathcal{AS},(a,w) \models \sigma \to (\phi \wedge \psi)$ as desired.

25.5 and 25.6 follow from the globality of STB_S.

□

Strong belief enforces closure under conjunction through "brute force": since all accepted arguments have to support each of the conjuncts, it follows from propositional logic for each of them that they have to support the conjunction. Thus STB_S is the best-behaved operator with respect to the standard doxastic axioms. In addition, this good behaviour is completely invariant with respect to the argument semantics. On the other hand strong belief is very demanding: speaking in terms of the mental models we discussed earlier, it requires that each theory contains only arguments supporting the proposition. That is, it is intolerant to neutral arguments as Example 13 demonstrates.

Example 13. In figure **??**, Alice has two arguments that are accepted under the grounded semantic. But since neither of them support both *rise* and *cure*, she does not believe either proposition.

While skeptical belief and credulous belief are more forgiving they are also less well-behaved with respect to the standard axioms. Depending on the use case or one's philosophical inclination this can be a defect or a strength (especially the lack of closure under conjunction). However, the failure of the usual knowledge-implies-belief condition suggests that in this setting a belief-first definition of knowledge in terms of arguments might be more fruitful. We leave this issue to future work.

Chapter 5

Dynamics for Argumentation Logic

We are now able to describe an agent's argumentation framework in the formal language provided by HAL. However, argumentation is an inherently dynamic process: agents devise, exchange, modify, withdraw, comment on, or deliberately ignore arguments to just name a few phenomena in the wild. Therefore it is perhaps surprising that the investigation of dynamics of argumentation frameworks have only attracted significant interest in the last ten years and is still in its infancy. An overview of this work is presented by Mailly [2015]. Most of the research on dynamics of argumentation is guided by one of three goals: characterising conditions that ensure *invariance* of semantics, extensions and the acceptance status of individual arguments under various dynamics; finding operations that *enforce* certain properties of argumentation frameworks; and, in a logical context, *expressing* properties of argumentation frameworks in a dynamic formal language. The second approach has synergies with the AGM paradigm of belief revision, especially when it poses the question of the minimal change to bring about the desired goal. In this way, the literature on enforcement pertains particularly to strategic argumentation. We will only touch on considerations of strategy in this work insofar as they affect agents' beliefs. Namely, we will investigate the degree of confidence an agent can have that his current beliefs will survive the confrontation with new arguments. While the third approach is important to this work insofar it concerns the expressibility of the logics used here, we will sidestep most of these issues by restricting ourselves to working with finite argumentation frameworks. It is the first approach that is most congenial to our work in this section which mainly consists in studying the behaviour of various semantics under operations that make sense from a doxastic viewpoint.

Such operations fall into the following categories (Boella et al. [2009], Boella et al. [2010]):

Definition 25 (Argumentation Framework Updates)**.** Let $\mathcal{A}$, $\mathcal{A}'$ be two argumentation frameworks. Then

- $\mathcal{A}'$ is an argument refinement of $\mathcal{A}$ iff $A \subseteq A'$ and for all (a, a'), $(a, a') \in \rightarrow'$ implies $(a, a') \in \rightarrow$.
- $\mathcal{A}'$ is an attack refinement of $\mathcal{A}$ iff $A = A'$ and $\rightarrow \subseteq \rightarrow'$.
- $\mathcal{A}'$ is an argument abstraction of $\mathcal{A}$ iff $A' \subseteq A$ and for all (a, a'), $(a, a') \in \rightarrow'$ implies $(a, a') \in \rightarrow$.
- $\mathcal{A}'$ is an attack abstraction of $\mathcal{A}$ iff $A = A'$ and $\rightarrow' \subseteq \rightarrow$.

Liao et al. [2011] define two operations that are capable of carrying out all of these changes:

Definition 26 (Addition)**.** Given an argumentation framework $\mathcal{A}$, an addition to $\mathcal{A}$ is defined as a tuple $(B, I_{A:B} \cup I_A)$ s.t. B is a set of arguments disjoint from A, $I_A \subseteq A \times A$, $I_{A:B} \subseteq \{(a,b),(b,a),(b,b) : a \in A, b \in B\}$. The argumentation framework resulting from $\mathcal{A}$ updated by $(B, I_{A:B} \cup I_A)$ is defined as

$$\mathcal{A}' = \mathcal{A} \bigoplus (B, I_{A:B} \cup I_A) = (A \cup B, R \cup (I_{A:B} \cup I_A))$$

Definition 27 (Deletion)**.** Given an argumentation framework $\mathcal{A}$, a deletion from $\mathcal{A}$ is defined as a tuple $(B, I_{A\setminus B:B} \cup I_{A\setminus B})$ s.t. $B \subseteq A$, $I_{A\setminus B} \subseteq A \setminus B \times A \setminus B$, $I_{A\setminus B:B} = R \cap (B \times A \cup A \times B)$. The argumentation framework resulting from $\mathcal{A}$ updated by $(B, I_{A\setminus B:B} \cup I_{A\setminus B})$ is defined as

$$\mathcal{A}' = \mathcal{A} \bigominus (B, I_{A\setminus B:B} \cup I_{A\setminus B}) = (A \setminus B, R \setminus (I_{A\setminus B:B} \cup I_{A\setminus B}))$$

5.1 Argumentation Framework Updates in HAL

5.1.1 Semantics

While the operations of addition and deletion provide us with the ability to add and remove arguments or attacks, they have a crucial drawback for our purposes. Ultimately, we would like to be able to use HAL as a language for argumentation dynamics that allows us to formally express dynamic properties of the kind "If $\mathcal{N}, a \vDash \phi$ and $\mathcal{A}' \vDash \psi$ then $\mathcal{A} \bigoplus \mathcal{A}' \vDash \gamma$. However, the notions of addition and deletion differ from argumentation frameworks in a crucial way: they contain relations over arguments that are not elements of their underlying space but of the original framework. Therefore, they cannot be understood as Kripke-frames and it is unclear how they could serve as semantics for HAL.

For this reason we will develop a different notion of addition and deletion of arguments. The basic idea is that both the original framework and the arguments and attacks to be added/removed should form full-blown, independent Kripke-frames. In addition, they should be named frames so that we can use them as semantics for HAL.

Hence let $\mathfrak{U} = (U, \twoheadrightarrow_{\mathfrak{U}}, \{D_i\}_{i \in Nom})$ be some $\twoheadrightarrow$-complete countably infinite, nice, named frame that will serve as our local universe. We call $\mathcal{N} = (A \subseteq$

$U, \twoheadrightarrow \subseteq \twoheadrightarrow_{\mathfrak{U}}, \{D_i\}_{i \in Assigned \subseteq Nom})$ a *subframe* of $\mathfrak{U}$. Then we define the following two operations.

Definition 28 (Argument Framework Union)**.** Let $\mathcal{N}$, $\mathcal{N}'$ be argumentation frameworks. Then the union of $\mathcal{N}$, $\mathcal{N}'$ is defined as

$$\mathcal{N} \cup \mathcal{N}' = \mathcal{N}''$$

where

- $A'' = A \cup A'$
- $R'' = R \cup R'$
- $D'' = \{D_i\}_{i \in Assigned \cup Assigned'}$

Definition 29 (Argument Framework Intersection)**.** Let $\mathcal{N}$, $\mathcal{N}'$ be argumentation frameworks. Then the intersection of $\mathcal{N}$, $\mathcal{N}'$ is defined as

$$\mathcal{N} \cap \mathcal{N}' = \mathcal{N}''$$

where

- $A'' = A \cap A'$
- $\to'' = \to \cap \to'$
- $D'' = \{D_i\}_{i \in Assigned \cap Assigned'}$

Note that the local universe $\mathfrak{U}$ plays an important role here: if $\mathcal{N}$, $\mathcal{N}'$ where not subset of $\mathfrak{U}$, it could be the case that they have the same name for two different arguments. Then $\mathcal{N} \cup \mathcal{N}'$ would no longer be a nice frame.

It is easy to prove that we can represent any addition and deletion using union and intersection update. For this purpose, denote by $\mathcal{A}_{\mathcal{N}}$ the argumentation framework underlying a partial named argumentation framework.

Proposition 26 (Representation)**.** Given a local universe $\mathfrak{U}$ and $\mathcal{N}$ a partial named subframe of $\mathfrak{U}$, for any addition $(B, I_{A:B} \cup I_A)$ where $B \subseteq U$, $I_{A:B} \cup I_A \subseteq \twoheadrightarrow_U$, there exists $\mathcal{N}'$ s.t. $\mathcal{A}_{\mathcal{N}} \oplus (B, I_{A:B} \cup I_A) = \mathcal{A}_{\mathcal{N}''}$ where $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$

Proof. Take $\mathcal{A}_{\mathcal{N}'} = (B \cup C, I_{A:B} \cup I_A)$ where $C = \{a \in A : (a,b), (b,a), (a,a'), (a',a) \in I_{A:B} \cup I_A \text{ for } b \in B, a' \in A\}$. Clearly, this is an argumentation framework. Since $B \cup C \subseteq U$, $\mathcal{A}_{\mathcal{N}'}$ corresponds to a unique named argumentation framework $\mathcal{N}'$. clearly $\mathcal{A}_{\mathcal{N}} \oplus (B, I_{A:B} \cup I_A) = \mathcal{A}_{\mathcal{N}} \cup \mathcal{A}_{\mathcal{N}'}$. But then clearly $\mathcal{A}_{\mathcal{N}} \oplus (B, I_{A:B} \cup I_A) = \mathcal{A}_{\mathcal{N}''}$ where $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$. □

The proof of proposition 26 makes it clear that nothing hinges on whether we work with named or unnamed argumentation frameworks, if we assume a local universe to take care of the naming. Hence in what follows we will exclusively use named argumentation frameworks.

In addition, it is clear that union and intersection update are "sound and complete" with respect to $\mathfrak{U}$ insofar as every frame obtained from a union update or a intersection update of subframes of $\mathfrak{U}$ is a subframe of $\mathfrak{U}$ and any subframe of $\mathfrak{U}$ can be obtained from any other subframe of $\mathfrak{U}$ through repeated union or intersection update.

Hence for any argument or attack refinement (abstraction) $\mathcal{N}''$ of $\mathcal{N}$ we can easily construct an argument framework $\mathcal{N}'$ s.t. $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$ ($\mathcal{N}'$ s.t. $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$).

Based on the directionality property Liao et al. [2011] develop a method to divide an argumentation framework under a semantic into three parts with respect to one of its sub-frameworks $\mathcal{B}$: a part unaffected by $\mathcal{B}$ (anything upstream from $\mathcal{B}$), a part affected by $\mathcal{B}$ ($\mathcal{B}$ itself and anything downstream from it) and a "conditioning" part (the subset of $\mathcal{B}$ that attacks any argument in the affected part). This division enables them to prove necessary and sufficient condition for monotony of semantics and argument status while updating only the affected part of the framework.

We will keep things slightly simpler here and only consider the unaffected and affected subframe of an update argumentation framework.

Definition 30 (Affected and Unaffected Part of an Argumentation Framework Union)**.** Let $\mathcal{N}$, $\mathcal{N}'$, $\mathcal{N}''$ be argumentation frameworks s.t. $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$. Then we define the following sub-frameworks of $\mathcal{N}''$: the affected arguments $(A''_a, \rightarrow''_a)$ and the unaffected arguments $(A''_u, \rightarrow''_u)$, where:

- $A''_a = A' \cup \{a \in A : A' \twoheadrightarrow''^{+} a\}$.
- $A''_u = A'' \setminus A''_a$.
- $\twoheadrightarrow''_a = \twoheadrightarrow'' \cap (A''_a \times A''_a)$.
- $\twoheadrightarrow''_u = \twoheadrightarrow'' \cap (A''_u \times A''_u)$.
- $D''_u = \{D_i\}_{i \in Assigned \cap Assigned'} \cap \{\{w\} : w \in A''_u\}$
- $D''_a = \{D_i\}_{i \in Assigned \cap Assigned'} \cap \{\{w\} : w \in A''_a\}$

In other words, the affected part of an Argumentation Framework comprises the arguments that are either from the update frame $\mathcal{N}'$ or are attacked - under the updated attack relation $\twoheadrightarrow''$ - by A'; the restriction of the updated attack relation to those arguments; and the names corresponding to those arguments. Of course we can make the same distinction for intersection update:

Definition 31 (Affected and Unaffected Part of an Argumentation Framework Intersection)**.** Let $\mathcal{N}$, $\mathcal{N}'$, $\mathcal{N}''$ be argumentation frameworks s.t. $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$. Then we define the following sub-frameworks of $\mathcal{N}''$: the affected arguments $(A''_a, \twoheadrightarrow''_a)$ and the unaffected arguments $(A''_u, \twoheadrightarrow''_u)$ where:

- $A''_a = \{a \in A : A' \cap A \twoheadrightarrow^{+} a\}$.

- $A''_u = A'' \setminus A''_a$.
- $\twoheadrightarrow''_a = \twoheadrightarrow'' \cap (A''_a \times A''_a)$.
- $\twoheadrightarrow''_u = \twoheadrightarrow'' \cap (A''_u \times A''_u)$.
- $D''_u = \{D_i\}_{i \in Assigned \cap Assigned'} \cap \{\{w\} : w \in A''_u\}$
- $D''_a = \{D_i\}_{i \in Assigned \cap Assigned'} \cap \{\{w\} : w \in A''_a\}$

Note while our definitions of the affected and unaffected arguments are equivalent to the ones given by Liao et al. [2011], they also consider what they call "conditioning arguments". However, the conditioning arguments in Liao et al. [2011]'s setting do not form a complete argumentation framework which is why we leave them out. A consequence is that taking the union of the unaffected and affected arguments would not yield back the orginal argumentation framework - namely it would be missing the attacks between them. However, we proceed to show that the results obtained by Liao et al. [2011] are still valid in our setting. For this purpose consider Liao et al. [2011]'s notion of the *conditioning arguments* of an updated argumentation framework:

Definition 32. Let $\mathcal{N}, \mathcal{N}', \mathcal{N}''$ be argumentation frameworks s.t. $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$ or $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$. The conditioning arguments $(A''_c, \twoheadrightarrow''_c, D'')$ are defined as follows:

- $A''_c = \{a \in A''_u : \text{ there is } b \in A''_a \text{ s.t. } a \twoheadrightarrow'' b\} \cup \{a \in A''_a : \text{ there is } b \in A''_u \text{ s.t. } b \twoheadrightarrow'' a\}$.
- $\twoheadrightarrow''_c = \twoheadrightarrow'' \cap (A''_c \times A''_a)$.
- $D''_c = \{D_i\}_{i \in Assigned \cap Assigned'} \cap \{\{w\} : w \in A''_c\}$

Definition 33. Given an argumentation framework $\mathcal{A}$ an assigned conditioned argumentation framework wrt $\mathcal{A}$ is a tuple $\mathcal{A}'_{\mathcal{A},E} = (\mathcal{A}', C(A), I, E)$ where

- $\mathcal{A}'$ is an argumentation framework s.t. $\mathcal{A} \cap \mathcal{A}' = \emptyset$
- $C(A) \subseteq A$ is non-empty
- I is a binary relation over $C(A) \times A'$ where for all $(a, b) \in I$, $a \in C(A)$, $b \in A'$
- E is an extension of $\mathcal{A}$ under some semantic S.

Semantics for assigned conditioned argumentation frameworks are defined as in standard argumentation framework except that the notion of acceptability of an argument is modified as follows:

- an argument $a \in \mathcal{A}'$ is acceptable in $\mathcal{A}'$ to a set B iff for all $b \in A'$ such that $b \twoheadrightarrow' a$ there exists $c \in E \cap C(A) \cup B$ s.t. $(c, b) \in I \cup \twoheadrightarrow'$ and for all $b' \in C(A)$, if $(b', a) \in I$ then $E \twoheadrightarrow b'$.

This definition of acceptability yields conditioned versions of the various semantics described in definition 2.

Again, conditioned argumentation frameworks are not frames in their own right. We will replace them with the following notion of acceptability given a set:

Definition 34. [Acceptability given a set of arguments] Let $\mathcal{N}$, $\mathcal{N}'$,$\mathcal{N}''$ be subframes of $\mathfrak{U}$s.t. $\mathcal{N}' \cap \mathcal{N}'' = \varnothing$, $\mathcal{N}' \subseteq \mathcal{N}$, $\mathcal{N}'' \subseteq \mathcal{N}$ and B, C sets of arguments s.t. $B \subseteq A'$, $C \subseteq A''$. An argument $a \in A'$ is called acceptable to B in $\mathcal{N}'$ given $\mathcal{N}''$, C iff for all $b \in A'$ s.t. $b \twoheadrightarrow' a$ we have $B \twoheadrightarrow' b$ or $C \twoheadrightarrow'' b$.

Definition 35. Let $\mathcal{N}$, $\mathcal{N}'$,$\mathcal{N}''$ be subframes of $\mathfrak{U}$s.t. $\mathcal{N}' \cap \mathcal{N}'' = \varnothing$, $\mathcal{N}' \subseteq \mathcal{N}$, $\mathcal{N}'' \subseteq \mathcal{N}$ and S a semantic. Denote by $S_{\mathcal{N}'',C}(\mathcal{N}')$ the set of extensions obtained by replacing acceptability by acceptability given $\mathcal{N}''$, C in the definition of S. If $Ext \in S_{\mathcal{N}'',C}(\mathcal{N}')$ we call Ext an extension of $\mathcal{N}'$ under S given $\mathcal{N}''$, C.

By replacing conditioned argumentation frameworks by normal argumentation frameworks plus acceptability given a set, we ensure that the division process can be expressed in the semantics of HAL. It remains to prove that we can really capture conditioned argumentation frameworks in this way.

Definition 36 (Completion of a Conditioned Argumentation Framework)**.** Given an assigned conditioned argumentation framework $\mathcal{N}'_{\mathcal{N},E} = (\mathcal{N}', C(A), I, E)$ wrt an argumentation framework $\mathcal{N}$ we call $\mathcal{N}'' = (A \cup A', \twoheadrightarrow \cup I \cup \twoheadrightarrow'')$ the *completion* of $\mathcal{N}'_{\mathcal{N},E}$.

Proposition 27 (Representation of CAFs)**.** Let $\mathfrak{U}$ be the local universe, $\mathcal{N}$ a subframe of $\mathfrak{U}$, S a semantic, $E \in S_{\mathcal{N}}$ an extension under S in $\mathcal{N}$ and $name'_{\mathcal{N},E} = (\mathcal{N}', C(A), I, E)$ an assigned conditioned argumentation framework wrt $\mathcal{N}$. Let $\mathcal{N}''$ be the completion of $\mathcal{N}'_{\mathcal{N},E}$. Then $a \in \mathcal{N}'$ is accepted wrt B in $\mathcal{N}'_E$ iff a is accepted wrt B in $\mathcal{N}'$ given $\mathcal{N}$, E. Likewise, E' is an extension of $\mathcal{A}'_E$ under S iff E' is an extension of $\mathcal{A}'$ given $\mathcal{N}$, E.

Proof. $\Rightarrow$ Suppose a is acceptable wrt B in $\mathcal{N}'$ given $\mathcal{N}, E$. Then for all $b \in A'$ s.t. $b \twoheadrightarrow a$ either $B \twoheadrightarrow' b$ or $E \twoheadrightarrow'' b$. In the latter case, since A and A' are disjoint, $b \in \{b : (a, b) \in I\}$. Since $b \notin A$, $E \not\twoheadrightarrow b$. Hence there must be $e \in E$ s.t. $(e, b) \in I$. Then by definition, $e \in E \cap C(A)$. Now suppose there were $c \in C(A)$ s.t. $(c, a) \in I$. Since all elements of I are of the form (x, y) where $x \in C(A)$ and $y \in A'$, $B \not\twoheadrightarrow c$ it must be the case that $E \twoheadrightarrow c$. Taking all this together we have that a is accepted wrt B in $\mathcal{N}'_E$ as desired. The case for extensions follows immediately from the definitions.

$\Leftarrow$ Suppose $a \in \mathcal{N}'$ is accepted wrt B in $\mathcal{N}'_E$. Then for all $b \in A'$ such that $b \twoheadrightarrow' a$ there exists $c \in E \cap C(A) \cup B$ s.t. $(c, b) \in I \cup \twoheadrightarrow'$ and for all $b' \in C(A)$, if $(b', a) \in I$ then $E \twoheadrightarrow b'$. Then either $c \in E \cap C(A)$ or $c \in B$. If $c \in E \cap C(A)$, then it is in E a fortiori and $(c, b) \in I$. But then $E \twoheadrightarrow'' b$. If $c \in B$, then $B \twoheadrightarrow' b$. But then a is accepted wrt B in $\mathcal{N}'$ given $\mathcal{N}''$, E as desired. The case for extensions is immediate from definition. □

This representation lemma tells us that if we choose the right superframe, namely the completion of a conditioned argumentation framework, conditioned acceptability and "acceptability given" are equivalent. The last remaining puzzle piece is to show that the completion of the affected arguments conditioned on the conditioning subset of the unaffected arguments of an argumentation framework union or intersection returns precisely the whole updated framework:

Proposition 28 (Representation of Completions)**.** Let $\mathcal{N}$, $\mathcal{N}'$,$\mathcal{N}''$ be subframes of $\mathfrak{U}$s.t. $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$ or $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$. Then the completion of $\mathcal{N}^a_{\mathcal{N}^c,E}$ is $\mathcal{N}''$.

Proof. We state only the case for union: $\mathcal{N}'' = (A \cup A', \twoheadrightarrow \cup \twoheadrightarrow', D \cup D') = (A''^a \cup A''^u, \twoheadrightarrow''_a \cup I \cup \twoheadrightarrow''_u, D \cup D'')$ since I is precisely the set of attacks between the affected and the unaffected part. □

Given the three representation lemmas we can finally transfer the an important result proved by Liao et al. [2011] for addition/deletion and conditioned argumentation frameworks to our setting:

Proposition 29. [Monotony for Semantics under Union and Intersection Update; Liao et al. [2011], Theorem 4,5, 6,7] Let $\mathcal{N}''$ be a subframe of $\mathfrak{U}$ and $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$ or $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$. Let $E \subseteq A''$. For every directional semantic S, $E \in S(\mathcal{N}'')$ iff $E \cap A_a \in S_{E'}(\mathcal{N}''_a)$, $E \cap A_u \in S(\mathcal{N}''_u)$ where E' is an extension of $\mathcal{N}_u$ under S.

5.1.2 Syntax

Union and Intersection Update provide us with the necessary semantics to add dynamic operators to the language of HAL. We define the resulting logic as follows:

Definition 37 (HAL with Argument Framework Union/Intersection)**.** The language is given by the following grammar:

$$\phi \coloneqq \mathcal{L} : \phi ::= i \mid \bot \mid \neg\phi \mid \phi \wedge \phi \mid \Box\phi \mid \Box_{odd}\phi \mid [\cup\mathcal{N}']\phi \mid [\cap\mathcal{N}']\phi$$

where $i \in Nom$ is a nominal and $\mathcal{N}'$ a subframe of $\mathfrak{U}$. Satisfaction is defined as in HAL plus the following clauses for the update operators is defined as follows:

- $\mathcal{N}, a \vDash [\cup\mathcal{N}']\psi$ iff $a \in \mathcal{N} \cup \mathcal{N}'$, $\mathcal{N} \cup \mathcal{N}', a \vDash \psi$
- $\mathcal{N}, a \vDash [\cap\mathcal{N}']\psi$ iff $a \in \mathcal{N} \cap \mathcal{N}'$, $\mathcal{N} \cap \mathcal{N}', a \vDash \psi$

We denote this language by $HAL^{\cap}_{\cup}$.

Undecidability of $HAL^{\cap}_{\cup}$

A useful technique to axiomatize dynamic logics is the use of reduction axioms. That is, axioms that allow the translation of any formula of the dynamic logic into a formula of the underlying static logic. Obviously, this is only possible if every formula of the dynamic logic is expressible in the underlying static logic. Can we reduce $HAL^{\cap}_{\cup}$ to HAL? No. For consider the binder operator we discussed in section 3.2.5. The binder operator works by naming the current argument. Thus $\downarrow x.\phi$ ensures that whenever the nominal x appears in ϕ, it refers to the current argument. We cannot do this syntactically in $HAL^{c}ap_{c}up$ but semantically, we can pick out the current state, say a, without a problem. But the dynamic modalities allow us to bring the semantics into the syntax. Hence we can use $[\cap\mathcal{N}']$ to simply delete every argument besides a. If x appears in ψ it has to denote a. Thus, x will be satisfied iff the current argument exists after deleting all arguments besides a. Thus we can express binder in $HAL^{\cap}_{\cup}$ as follows:

Proposition 30. Let ϕ be a $(\mathcal{H}_{odd}, @, \downarrow)$-formula of the form $\downarrow x.\psi$. Let $\mathcal{N}$ be a named frame. Then we have that $\mathcal{N}, a \vDash \phi$ iff $\mathcal{N}, a \vDash \psi[x := [\cap(\{a\}, \{\}, D_i = \{a\})]\top]$.

Proof. $\Rightarrow$ Suppose $\mathcal{N}, a \vDash \psi[x := [\cap(\{a\}, \{\}, D_i = \{a\})]\top]$. Then $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N} = (\{a\}, \{\}, D_i = \{a\})$. Hence $\mathcal{N}, b \vDash [\cap(\{a\}, \{\}, D_i = \{a\})]\top$ iff $a = b$. Now pick any fresh nominal y and set $D_y = \{a : \mathcal{N}, a \vDash [\cap(\{a\}, \{\}, D_i = \{a\})]\top\}$. Clearly, $D_y = \{a\}$. Denote by $\mathcal{N}_{D_y = \{a\}}$ the frame obtained from $\mathcal{N}$ by replacing $\{D_i\}_{i \in Assigned}$ by $\{D_i\}_{i \in Assigned \cup \{D_x\}}$. Clearly, $\mathcal{N}_{D_y=\{a\}}, a \vDash y$ if $\mathcal{N}, a \vDash [\cap(\{a\}, \{\}, D_i = \{a\})]\top$. Hence $\mathcal{N}_{D_y=\{a\}}, a \vDash \psi$. But then $\mathcal{N}, a \vDash \downarrow y.\psi$ as desired.

$\Leftarrow$ Suppose $\mathcal{N}, a \vDash \phi$. Then for $\mathcal{N}_{D_x=\{a\}} = (A, R, \{D_i\}_{i \in Assigned \cup D_x})$ where $D_x = \{a\}$ we have that $\mathcal{N}_{D_x=\{a\}}, a \vDash \psi$. Now recall from the proof of the other direction that $D_x = \{a\} = \{a : \mathcal{N}, a \vDash [\cap(\{a\}, \{\}, D_i = \{a\})]\top\}$. Hence if $\mathcal{N}_{D_x=\{a\}}, a \vDash x$ then $\mathcal{N} \vDash [\cap(\{a\}, \{\}, D_i = \{a\})]\top$ while for any x-free formula γ, $\mathcal{N} \vDash \gamma$ if $\mathcal{N}_{D_x=\{a\}}, a \vDash \gamma$. But then if $\mathcal{N}_{D_x=\{a\}}, a \vDash \psi$ then $\mathcal{N}, a \vDash \psi[x := [\cap(\{a\}, \{\}, D_i = \{a\})]\top]$ as desired. ☐

Thus adding union and intersection update raises the expressiveness of HAL to at least the level of HAL with binder. This implies that we cannot hope to find an axiomatization for $HAL^{\cap}_{\cup}$ based on reduction axioms to HAL. One might then ask, is HAL with binder enough? Again the answer is negative: if removing all worlds but one gives us added expressivity, what can we do by removing any number of worlds? The answer is, we can mimmick propositional letters. Semantically, a propositional letter is just a set of worlds. But we can pick out every set of worlds using $[\cap\mathcal{N}']$.

Proposition 31. Let $Prop$ be a set of propositional letters and $\mathcal{N}, V_{Prop}$ a model with a valuation function $V_{Prop} : Prop \longrightarrow 2^{A}$. Then for any $p \in Prop$, $\mathcal{N}, V_{Prop}, a \vDash p$ iff $\mathcal{N}, a \vDash [\cap\mathcal{N}']\top$ where $\mathcal{N}'$ is the restriction of $\mathcal{N}$ to $V_{Prop}(p)$.

Proof. $\Rightarrow$ Suppose $\mathcal{N}, a \vDash [\cap\mathcal{N}']\top$. Define $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$. $A'' = V_{Prop}(p)$ by definition of $\mathcal{N}'$. Hence $a \in V_{Prop}(p)$. But then $\mathcal{N}, V_{Prop}, a \vDash p$ as desired.

$\Leftarrow$ Suppose $\mathcal{N}, V_{Prop}, a \vDash p$. Then $a \in V_{Prop}(p)$ by its definition. Let $\mathcal{N}'$ be the restriction of $\mathcal{N}$ to $V_{Prop}(p)$. By the definition of satisfaction for $[\cap\mathcal{N}']$, $\mathcal{N}, a \vDash [\cap\mathcal{N}']\top$. □

Thus argumentation framework union and intersection turn HAL into a full-blown Hybrid logic with propositional letters. And we are not done yet: so far we have only used the expressivity yielded by removing states. What happens if we consider the manipulations of the attack relation that $HAL^{\cap}_{\cup}$ allows?

Proposition 32. Let ϕ be $(\mathcal{H}_{odd}, \Diamond_U)$-formula of the form $\Box_U\psi$ and $\mathcal{N}$ a named frame. Then $\mathcal{N}, a \vDash \phi$ iff $\mathcal{N}, a \vDash [\cup\mathcal{N}'] \Box [\cap\mathcal{N}]\psi$ where $\mathcal{N}'$ is the complete and reflexive graph over A.

Proof. $\mathcal{N}'$ is an $S5$-frame with a single non-empty equivalence class and hence its $\Box$ is equivalent to the global modality. Clearly, $\mathcal{N}', a' \vDash \Box[\cap\mathcal{N}]\psi$ iff $\mathcal{N}, a' \vDash \psi$. Hence for any a', $\mathcal{N}, a' \vDash \psi$ iff $\mathcal{N}, a \vDash [\cup\mathcal{N}'] \Box [\cap\mathcal{N}]\psi$. But then $\mathcal{N}, a \vDash \phi$ iff $\mathcal{N}, a \vDash [\cup\mathcal{N}'] \Box [\cap\mathcal{N}]\psi$ as desired. □

Now we have reached a decisive point: Recall the identity *first-order logic = modal logic + nominals + global modality + interpolation* we discussed in chapter 3. It is known that the Hybrid logic with binder has the interpolation property (see e.g. Cate [2005], Theorem 9.51). It follows that $HAL^{\cap}_{\cup}$ is at least as expressive as first-order logic (we have not considered the transitive closure modality). The main result of this section follow immediately:

Proposition 33. $HAL^{\cap}_{\cup}$ is undecidable.

What has happened here? We started with a very impoverished language that did not even contain propositional letters and ended up with the full expressive power of first order logic. This goes to show that frame union and intersection are extremely expressive operations. This ought to be expected since at the beginning of this chapter we noted that union and intersection are "complete" with respect to the local universe. That is, any subframe of the local universe can be constructed from any other frame using (combinations of) union and intersection. Hence we can modify the semantics as we please and in addition "import" the semantics into the syntax through the union and intersection operators. This combination allows us to express any frame condition using the union and intersection operators which yields the enormous expressive power.

Another logic which offers similar abilities is Local Graph Modifier Logic ($LGML$) as described by Aucher et al. [2009]. $LGML$ has similar expressivity to $HAL^{\cap}_{\cup}$, except for *models*, not frames: $LGML$ provides various operators to construct any model by adding or removing states, modifying the accessibility relation and modifying the valuation function. As we have seen, we can express propositional letters in $HAL^{\cap}_{\cup}$ and thus one might speculate that it is possible to reduce $LGML$ to $HAL^{\cap}_{\cup}$. On the other hand, Aucher et al. [2009] provide a reduction from Hybrid logic with binder and global modality to Local Graph

Modifier Logic. Thus, it would be interesting to investigate the exact relationship between $HAL^{\cap}_{\cup}$ and $LGML$. We leave this question as well as the quest for an axiomatization of $HAL^{\cap}_{\cup}$ to future work.

Expressing Argumentation

Based on the semantics developed in the previous sections, we can use the language of $HAL^{\cap}_{\cup}$ to express the results developed by [Liao et al., 2011]. We begin with the notion of (un)affectedness which can be expressed in the language of HAL in the following way on the set and argument level:

Proposition 34. Let $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$ or $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$ and $B \subseteq \mathcal{N}''$. Then

1. B is the unaffected set iff $\mathcal{N} \vDash [\cup \mathcal{N}'](\check{B} \leftrightarrow \neg \Diamond^* \check{A}')$ or $\mathcal{N} \vDash [\cap \mathcal{N}'](\check{B} \leftrightarrow \neg \Diamond^* \check{A}')$ respectively.
2. B is the affected set iff $\mathcal{N} \vDash [\cup \mathcal{N}'](\check{B} \leftrightarrow \Diamond^* \check{A}')$ or $\mathcal{N} \vDash [\cap \mathcal{N}'](\check{B} \leftrightarrow \Diamond^* \check{A}')$ respectively

Proof. 34.1: $\Rightarrow$: Suppose $\mathcal{N} \vDash [\cup \mathcal{N}'](\check{B} \leftrightarrow \neg \Diamond^* \check{A}')$. Then $\mathcal{N}'' \vDash \check{B}$ iff $\mathcal{N}'' \vDash \neg \Diamond^* \check{A}'$ iff $\mathcal{N}'' \vDash \neg \check{A}' \wedge \neg \Diamond^+ \check{A}'$. Hence $a \in B$ iff $a \notin A'$ and a is not reachable from A' as desired. The argument is the same for the intersection case. $\Leftarrow$: similar. 34.2 follows immediately from from 34.1. ☐

Proposition 35. Let $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$ or $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$ and $a \in \mathcal{N}''$. Then

- a is unaffected iff $\mathcal{N} \vDash [\cup \mathcal{N}']@_a \neg \Diamond^* \check{A}'$, $\mathcal{N} \vDash [\cap \mathcal{N}']@_a \neg \Diamond^* \check{A}'$ respectively.
- a is affected iff $\mathcal{N} \vDash [\cup \mathcal{N}']@_a \Diamond^* \check{A}'$, $\mathcal{N} \vDash [\cap \mathcal{N}']@_a \neg \Diamond^* \check{A}'$ respectively.

Proof. An immediate corollary of 34. ☐

Next we consider "acceptability given" as defined in definition 34. In the language of HAL this is expressed as follows:

Proposition 36. Let $\mathcal{N}$, $\mathcal{N}'$,$\mathcal{N}''$ be subframes of $\mathfrak{U}$s.t. $\mathcal{N}' \cap \mathcal{N}'' = \varnothing$, $\mathcal{N}' \subseteq \mathcal{N}$, $\mathcal{N}'' \subseteq \mathcal{N}$ and B, C sets of arguments s.t. $B \subseteq A'$, $C \subseteq A''$. Then $a \in A'$ is acceptable to B in $\mathcal{N}'$ given $\mathcal{N}''$, C iff

$$\mathcal{N} \vDash @_a \Box (\check{A}' \rightarrow ([\cap \mathcal{N}'] \Diamond \check{B} \vee [\cap \mathcal{N}''] \Diamond \check{C}))$$

Proof. $\Rightarrow$: Suppose $\mathcal{N} \vDash @_a \Box (A' \rightarrow ([\cap \mathcal{N}'] \Diamond B \vee [\cap \mathcal{N}''] \Diamond C))$. It is easy to see that then for every attacker b of a under $\twoheadrightarrow'$, there is an attacker c such that either $c \in A'$ and $c \twoheadrightarrow' b$ or $c \in A''$ and $c \twoheadrightarrow'' b$. But that is just the desired definition. $\Leftarrow$: similar. ☐

As a corollary we also get a $HAL^{\cap}_{\cup}$-version of Definition 35

Proposition 37. Let $\mathcal{N}$, $\mathcal{N}'$,$\mathcal{N}''$ be subframes of $\mathfrak{U}$s.t. $\mathcal{N}' \cap \mathcal{N}'' = \varnothing$, $\mathcal{N}' \subseteq \mathcal{N}$, $\mathcal{N}'' \subseteq \mathcal{N}$, $C \subseteq A''$, S a semantic and ϕ the formula scheme characterising S as per Proposition 17. Define $\phi_{\mathcal{N}'',C} = \phi[\Box \Diamond \psi := \Box(\check{A}' \to ([\cap \mathcal{N}'] \Diamond \psi \vee [\cap \mathcal{N}''] \Diamond \check{C}))]$ to be the formula scheme where all instances of the scheme $\Box \Diamond \psi$ are substituted by $\Box(\check{A}' \to ([\cap \mathcal{N}'] \Diamond B \vee [\cap \mathcal{N}''] \Diamond \check{C}))$. Then X is an extension of $\mathcal{N}'$ under S given $\mathcal{N}'', C$ iff $\mathcal{N} \vDash \phi_C(\check{X})$.

Proof. This is an immediate corrolary of 36. ☐

Proposition 38 (Monotony for Semantics under Union/Intersection Update in $HAL^{\cap}_{\cup}$)**.** Let $\mathcal{N}''$ be a subframe of $\mathfrak{U}$ and $\mathcal{N}'' = \mathcal{N} \cup \mathcal{N}'$ or $\mathcal{N}'' = \mathcal{N} \cap \mathcal{N}'$. Let $E \subseteq A''$, S be a semantic and ϕ be the formula scheme characterising S. Then $\mathcal{N}'', a \vDash \phi(\check{E})$ iff $\mathcal{N}_a, a \vDash (\phi_{E'}(\check{E} \wedge \check{A}_a))$ and $\mathcal{N}_u, a \vDash (\phi(\check{E} \wedge \check{A}_u))$ where E' is an extension of $\mathcal{N}_u$ under S.

Proof. This is a corollary of 29 and 37. ☐

Chapter 6

Argument Dynamics and Belief

We are now in a position to tackle the problem we set out to investigate in the introduction: when does a new argument change an agent's opinion? We scrutinize this question in the framework of Argument Support Logic enriched with argument dynamics. That is, in the product logic, the place of HAL is now taken by $HAL^{\cap}_{\cup}$. We will restrict ourselves to the interactions between one kind of opinion change and one kind of update here: losing a belief (or belief contraction as it is called in the belief revision literature) and argumentation framework union. This is not to say that it is not of interest how argumentation framework intersection update influences belief and that there are not different ways in which update operations can change belief. Intersection update is interestingly not symmetric in its effects to union: for example, removing an argument or an attack can be a more effective way of causing a belief change than adding arguments and attacks since it allows for example the splitting of extensions. Likewise it is obvious that belief gain is of equal interest than belief loss and there are other interesting questions that can be asked with respect to belief change. For example, it would be of interest under which circumstances an agent's beliefs can be rendered (in)consistent through updates of their argumentation framework; or how their beliefs can be weakened and strengthened in this way. However, we will leave these additional questions to future work.

There are two ways in which union update can affect an agent's beliefs. Firstly, through manipulating (adding and changing the attack relation on) arguments that affect accepted arguments in the original argumentation framework. We call this kind of change *destructive*. Secondly, by adding arguments that leave the acceptance status of arguments from the original argumentation framework unaffected. This latter approach may seem strange but we will see that skeptical agents may lose a belief when they are provided with alternative "theories" that neither contradict nor support or when their theories are "watered down" through additions.We call this kind of change *expansive*.

We start by looking at the conditions under which an update of an argumentation framework is safe for a credulous belief an agent holds.

6.1 Dynamics of Credulous Belief

We begin with destructive loss of belief. It is intuitively clear that supplying someone with a single counterargument against an argument for a belief they have will not generally suffice to make them lose that belief. People usually have more than just one argument for an opinion they hold. In addition, they may have counterarguments that defend their original argument against the new attack.

For credulous belief, this means that the new arguments have to kick out every argument that supports the belief from all extensions of the argumentation framework. This can be achieved in two ways: either by attacking the arguments themselves; or, if they are attacked, by attacking their defenders.

Either option may be easier to achieve, depending on the situation: an argument a supporting a belief may be at the root of an inverse tree under indirect defense. In that case it has multiple defenders and each of its defenders has several defenders and so forth. Obviously, then the better strategy is to attack arguments directly. On the other extreme each of the supporting arguments may be a leaf of an indirect defense tree. In that case, the belief can be lost due to a single attack on the root of the tree.

We pin these ideas down in terms of a notion of "entrenchment" which we define in what follows. Since entrenchment depends on the number of defenders an argument has, we first need to introduce counting versions of the $\Diamond$-operator. We do this making use of the expressive power of nominals under our usual assumption that the argumentation framework is finite.

Definition 38 (Counting Operators)**.** Given an Argument-Support model $\mathcal{AS}$, abbreviate by $\Diamond^n\phi$ the formula $\bigvee_{\binom{n}{|A|}}(\Diamond(i_1 \wedge \phi)\wedge, ... \wedge \Diamond(i_n \wedge \phi))$ where $\bigvee_{\binom{n}{|A|}}$ denotes the disjunction over all combinations of n arguments from A. Correspondingly, define by $\Diamond^{\geq n}\phi$ the disjunction $\Diamond^n\phi \vee \Diamond^{n+1}\phi \vee ... \vee \Diamond^{|A|}\phi$ and by $\Diamond^{\leq n}\phi$ the disjunction $\neg\Diamond\phi \vee \Diamond\phi \vee ... \vee \Diamond^n\phi$. The strict inequality modalities $\Diamond^{>n}$, $\Diamond^{<n}$ are defined accordingly and the definitions are extended in the obvious way to $\Diamond_U$, $\Diamond_{odd}$.

We put these counting modalities to use to define a counting version of acceptance.

Definition 39. Given an named argumentation framework $\mathcal{N}$, an argument a is accepted to degree n wrt a set $X \subseteq A$ iff every attacker of a is attacked by n arguments from X or in other words

$$\mathcal{N}, a \vDash \Box \Diamond^n \check{X}$$

We say that a is strongly accepted to degree n wrt X iff a and every (indirect) defender of a are accepted to degree n, that is:

$$\mathcal{N}, a \vDash \boxminus_{odd} \diamondminus^n \check{X}$$

We are now ready to make precise what we mean by entrenchment:

Definition 40. Let $\mathcal{AS}$ be a finite argument-support model, S a semantic, and ψ the formula scheme characterising that semantic as per Proposition 17. Then we say that ϕ is *credulously entrenched* to degree x iff it is supported by at least x arguments that are credulously strongly accepted to degree x under S or in other words:

$$\mathcal{AS}, (a, w) \vDash \diamondminus_U^{\geq x}(\bigvee_{X \subseteq A} (\psi(\check{X}) \wedge \check{X}) \wedge \overline{K}(\sigma \rightarrow \phi) \wedge (\boxminus_{odd} \diamondminus^{\geq x} \check{X}))$$

We say that ϕ is *grounded credulously entrenched* to degree x iff it is *credulously entrenched* to degree x and in addition, each of the considered supporters of ϕ is indirectly defended by at least x unattacked arguments:

$$\mathcal{AS}, (a, w) \vDash \diamondminus_U^{\geq x}(\bigvee_{X \subseteq A} (\psi(\check{X}) \wedge \check{X}) \wedge \overline{K}(\sigma \rightarrow \phi) \wedge (\boxminus_{odd} \diamondminus^{\geq x} \check{X}) \wedge (\boxminus_{odd} \diamondminus_{odd} \neg \diamondminus_{odd} \top)$$

Finally, ϕ is credulously entrenched to degree x given $\mathcal{N}', C$ in $\mathcal{N}''$ where $\mathcal{N}' \cap \mathcal{N}'' = \varnothing$ and $\mathcal{N}' \subseteq \mathcal{N}$, $\mathcal{N}'' \subseteq \mathcal{N}$ iff

$$\mathcal{AS}, (a, w) \vDash \diamondminus_U^{\geq x}(\bigvee_{X \subseteq A} (\psi_{\mathcal{N}, C}(\check{X}) \wedge \check{X}) \wedge \overline{K}(\sigma \rightarrow \phi) \wedge (\boxminus_{odd} \diamondminus^{\geq x} ([\cap \mathcal{N}'] \check{X} \vee [\cap \mathcal{N}''] \check{X})))$$

That is, ϕ is supported by at least x accepted arguments such that for every indirect attacker of a there is an attacker that is either in A' or in A''.

A proposition is thus credulously entrenched to degree x if it is supported by at least x accepted arguments and those arguments have at least x accepted defenders against every of their indirect attackers.

In addition to the notion of entrenchment we need a dynamic-oriented version of belief, i.e. a kind of belief operator that tells us what would happen if we were to implement a certain change to the argumentation framework. Recall the definitions of acceptability given a set (Definition 34, Definition 36) and extensions given a set (Definition 35, Definition 37. We define belief given a set of arguments as follows:

Definition 41 (Belief given a Set of Arguments)**.** Let $Bel \in \{CB, SB, STB\}$ and $\mathcal{AS}, \mathcal{AS}', \mathcal{AS}''$ be Argument-Support models such that $W = W' = W''$, $A' \subseteq A$, $A'' \subseteq A$, $A' \cap A'' = \varnothing$, $C \subseteq A''$, S a semantic and ψ the formula that characterises S as per proposition 17. Then ϕ is believed in $\mathcal{AS}'$ given $\mathcal{N}'', C$ iff $\mathcal{AS} \vDash Bel_{S, \mathcal{N}'', C}\phi$ where $Bel_{S, \mathcal{N}'', C}\phi := Bel_S \phi[\phi := \phi_{\mathcal{N}'', C}]$.

One can think of this kind of belief in terms of the situation a jury faces in a court case. There is a body of evidence the general public knows about.

The jurors however are only provided the evidence the court deems admissible. Thus their argumentation frameworks constitute subframes of the overall argumentation framework representing the body of evidence. The evidence that is presented to them at any given time is a subset of the admissible evidence. They don't necessarily know whether the evidence in this set is accepted among the admissible evidence. However, during the court proceedings it becomes apparent to them how each piece of evidence supports or undermines the presumption of innocence of the accused and which other pieces of evidence it attacks or defends in *their* argumentation frameworks. Hence they believe in the innocence or guilt of the accused given a set of evidence from the argumentation framework represented by the admissible evidence which is in turn a subframe of the overall body of evidence.

Next we define a corresponding notion of degrees for updates. In words an argumentation framework union is of degree x if both the number of attacks from the update framework to the original framework *and* the number of vice versa attacks are smaller than x. For intersection update we require that the update remove less or equal to x attacks and arguments from the original framework. More formally:

Definition 42. Let $\mathcal{AS}$ be a finite argument-support model. We call $[\cup\mathcal{N}']$ as of degree x to $\mathcal{AS}$ if $\mathcal{AS} \vDash \boxdot_U [\cup\mathcal{N}'] \diamondplus_U^{\leq x} (\check{A} \wedge \diamondplus \check{A}') \wedge \diamondplus_U^{\leq x} (\check{A}' \wedge \diamondplus \check{A})$. We call $[\cap\mathcal{N}']$ as of degree x to $\mathcal{AS}$ if $\mathcal{AS} \vDash \boxdot_U((\diamondplus^{>x} \check{A}' \to [\cap\mathcal{N}'] \diamondplus \check{A}') \wedge \diamondplus^{\leq x}[\cap\mathcal{N}']\bot$

Together with propositions 29, 38 and this allows us to derive the following result:

Proposition 39. Let $\mathcal{AS}$ be an Argument-Support model and S a semantic such that $\mathcal{AS} \vDash CB_S\phi$. Then $CB_S\phi$ is preserved under an update $[\cup\mathcal{N}']$ if any of the following conditions hold:

1. $\mathcal{AS}''_u \vDash CB_S\phi$ or $\mathcal{AS}''_a \vDash CB_{S,\mathcal{N}''_u,E}\phi$ where $E \in S(\mathcal{N}''_u)$.

2. for $S \in \{\mathcal{AD}, \mathcal{CO}, \mathcal{PR}\}$ ($S = \mathcal{GR}$), ϕ is (grounded) credulously entrenched to degree x in $\mathcal{N}$ or $\mathcal{N}'$ and $[\cup\mathcal{N}']$ is of degree $< x$ to $\mathcal{N}$.

3. for $S \in \{\mathcal{AD}, \mathcal{CO}, \mathcal{PR}\}$ ($S = \mathcal{GR}$), $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$ and ϕ is (grounded) credulously entrenched to degree x in $\mathcal{N} \cap \mathcal{N}''_a$ given $\mathcal{N}''_u, E$ where $E \in S(\mathcal{N}''_u)$ or (grounded) credulously entrenched to degree x in $\mathcal{N}' \cap \mathcal{N}''_a$.

Proof. 39.1 is an easy corollary of Proposition 38.

39.2: Consider first the case where $S \in \{\mathcal{AD}, \mathcal{CO}, \mathcal{PR}\}$ and suppose ϕ is credulously entrenched to degree x in $\mathcal{N}$ or $\mathcal{N}'$ and $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$. In the first case, there are at least x accepted supporters of ϕ in A which are defended against any indirect attacker by at least x accepted arguments. Since $\mathcal{N}'$ is of degree $< x$, $A' \twoheadrightarrow'' a$ for at most at most $x-1$ arguments a in $\mathcal{N}$. But then there must be $a \in A''$ s.t. a supports ϕ and a has at least one accepted defender a' against any indirect attacker. But then a is credulously accepted under $\mathcal{AD}$,

$\mathcal{CO}$ and $\mathcal{PR}$. It follows immediately that $\mathcal{N}'' \models CB_S\phi$. The second case is analogous. As for the grounded semantic, note that by grounded entrenchment each of the x considered accepted arguments supporting ϕ is indirectly defended by at least x unattacked arguments. But then there is $a \in A''$ s.t. a supports ϕ and a has at least one accepted defender a' against any indirect attacker and in addition is indirectly defended against any indirect attacker by at least one unattacked argument. But then by Proposition 17.5, a is accepted under the grounded semantic. Hence $\mathcal{N}'' \models CB_{\mathcal{GR}}\phi$ as desired.

39.3: Again consider first the case where $S \in \{\mathcal{AD}, \mathcal{CO}, \mathcal{PR}\}$ and suppose $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$ and ϕ is (grounded) credulously entrenched to degree x in $\mathcal{N} \cap \mathcal{N}''_a$ given $\mathcal{N}''_u, E$ where $E \in S(\mathcal{N}''_u)$ or (grounded) credulously entrenched to degree x in $\mathcal{N}' \cap \mathcal{N}''_a$. In the first case, there are at least x accepted supporters a of ϕ in A which are defended against any indirect attacker by at least x arguments that are accepted given $\mathcal{N}''_u, E$. Since $\mathcal{N}'$ is of degree $< x$, $A' \twoheadrightarrow'' a$ for at most at most $x - 1$ arguments a in $\mathcal{N}$. Then there must be $a \in A''_a$ s.t. a supports ϕ and a has at least one defender a' that is accepted given $\mathcal{N}''_u, E$ against any indirect attacker. Hence a is accepted in $\mathcal{N}''_a$ given $\mathcal{N}''_u, E$. Then by proposition 29, any such a is accepted in $\mathcal{N}''$ and hence $\mathcal{N}'' \models CB_S\phi$. The argument for the second case as well as the grounded extension follow like before.

□

In other words, if credulous belief in ϕ (given an extension of the unaffected arguments) obtains for either of the sub-frameworks $\mathcal{N}''_u$ or $\mathcal{N}''_a$ of the updated argumentation framework then $\mathcal{N}'' \models \overline{CB}\phi$. Furthermore, if a credulous belief is credulously entrenched to degree x then any union-update with an argumentation framework that attacks less than x arguments in the original argumentation framework will not change the credulous belief. Finally, it also suffices that ϕ be credulously entrenched to degree x given the unaffected arguments in only the affected arguments within $\mathcal{N}$ or $\mathcal{N}'$. This kind of belief change falls in the destructive category we outlined earlier. Notably, credulous belief is safe from any expansive update: adding new extensions never affects the acceptance status of arguments under the original extensions. That is, for a credulous believer more evidence always leads to more believed propositions (under semantics that allow multiple extensions). Coming back to the interpretation of extensions as mental models or theories, the credulous believer maximizes: she accepts any new theory regardless of whether it is consistent with her other theories.

6.2 Dynamics of Skeptical Belief

The natural question is now whether the conditions we devised for credulous belief also suffice for the preservation of skeptical and strong belief. As far as the grounded extension and skeptical belief are concerned, the answer is yes since under the grounded extension credulous and skeptical belief are equivalent. Hence we immediately get as a corollary of Proposition 39 that the same sufficient conditions for preservation apply to skeptical belief under the grounded extension:

Proposition 40. Let $\mathcal{AS}$ be an Argument-Support model such that $\mathcal{AS} \models SB\phi$ and $\mathcal{GR}$ be the grounded semantic. Then $SB_{\mathcal{GR}}\phi$ is preserved under an update $[\cup\mathcal{N}']$ if any of the following conditions hold:

1. $\mathcal{AS}''_u \models SB_{\mathcal{GR}}\phi$ or $\mathcal{AS}''_a \models SB_{\mathcal{GR},\mathcal{N}''_u,E}\phi$ where $E \in \mathcal{GR}(\mathcal{N}''_u)$.

2. ϕ is grounded credulously entrenched to degree x in $\mathcal{N}$ or $\mathcal{N}'$ and $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$.

3. $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$ and ϕ is grounded credulously entrenched to degree x in $\mathcal{N} \cap \mathcal{N}''_a$ given $\mathcal{N}''_u, E$ where $E \in \mathcal{GR}(\mathcal{N}''_u)$ or grounded credulously entrenched to degree x in $\mathcal{N}' \cap \mathcal{N}''_a$.

Proof. This immediately follows from Propositions 39 and 23 which states that under grounded semantics credulous and skeptical belief are equivalent. □

However, the picture looks entirely different for the other semantics. For admissible semantics, we do not need to investigate invariance of skeptical belief since, as we have seen in Proposition 22, no proposition is ever skeptically admissibly believed. For complete and preferred semantics it suffices that a single extension contains no supporter of ϕ for the agent to lose belief in ϕ. Thus the crucial question is how many supporters of ϕ *each* extension contains, how deeply entrenched they are *and* whether the update creates any new extensions without supporters of ϕ. In other words, skeptical belief is vulnerable to expansive updates. Updates that add a new candidate to the set of acceptable theories or mental models at the agent's disposal can shake her confidence in her previous beliefs if the new candidate does not bear out the proposition in question. To account for this we need to modify our notion of entrenchment:

Definition 43. Let $\mathcal{N} \times \mathcal{F}$ be a finite argument-support model, S a semantic, ψ the formula scheme characterising that semantic. A formula ϕ is skeptically entrenched to degree x iff a) every extension under S contains at least x arguments that support ϕ which are defended against any indirect attacker by at least x arguments from the extension and b) every argument a that is rejected wrt S and does not support ϕ is undefended and has at least x accepted attackers or every indirect defender of a has at least x accepted attackers. In other words ϕ is skeptically entrenched to degree x iff

a) it is strongly entrenched to degree x, that is $\mathcal{AS} \models \bigwedge_{X\subseteq A} \psi(\check{X}) \to \Diamond_U^{\geq x}(\check{X} \wedge \overline{K}(\sigma \to \phi) \wedge (\boxminus_{odd} \Diamond^{\geq x} \check{X})))$

b) and it is reverse entrenched to degree x, that is $\mathcal{AS} \models \boxminus_U(\bigwedge_{X\subseteq A}(\psi(\check{X}) \to \neg\check{X}) \wedge \langle\overline{K}\rangle(\sigma \wedge \neg\phi) \to \Diamond^{\geq x}\top \vee \boxminus_{even} \Diamond^{\geq x} \top$

Then we obtain the following invariance result for skeptical belief under destructive change:

Proposition 41. Let $\mathcal{AS}$ be an Argument-Support model and $S \in \{\mathcal{CO}, \mathcal{PR}\}$ a semantic such that $\mathcal{AS} \vDash SB_S\phi$. Then $SB_S\phi$ is preserved under an update $[\cup\mathcal{N}']$ if any of the following conditions hold:

1. $\mathcal{AS}''_u \vDash SB_S\phi$ or $\mathcal{AS}''_a \vDash SB_{S,\mathcal{N}''_u,E}\phi$ where $E \in S(\mathcal{N}''_u)$.

2. ϕ is skeptically entrenched to degree x in $\mathcal{N}$ and $\mathcal{N}'$ and $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$.

3. $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$ and ϕ is skeptically entrenched to degree x in $\mathcal{N} \cap \mathcal{N}''_a$ given $\mathcal{N}''_u, E$ where $E \in S(\mathcal{N}''_u)$ or skeptically entrenched to degree x in $\mathcal{N}' \cap \mathcal{N}''_a$.

Proof. 42.1: This again carries over from Proposition 38.

42.2: Suppose ϕ is skeptically entrenched to degree x in $\mathcal{N}$ and $\mathcal{N}'$ and $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$. Then every extension of $\mathcal{N}$, $\mathcal{N}'$ under S contains at least x arguments that support ϕ which are defended against any indirect attacker by at least x arguments from the extension and b) every argument a that is rejected wrt S in $\mathcal{N}$, $\mathcal{N}'$ respectively and does not support ϕ is either undefended and has at least x accepted attackers or every indirect defender of a has at least x accepted attackers. Since $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$, only at most $x - 1$ arguments from A can be attacked by A' and vice versa. Now suppose for contradiction there were an extension $E \in S(\mathcal{N}'')$ such that E does not contain any argument that supports ϕ. By assumption, E can not be an extension of either $\mathcal{N}$ or $name'$. In addition, E cannot be obtained by removing all arguments that support ϕ from an existing extension E' of $\mathcal{N},\mathcal{N}'$ since $x - 1$ attacks do not suffice for that. Adding arguments to E' does not affect the acceptance status of members of E' for complete and preferred semantics. Hence the only remaining option is that E is a new extension. But then E must contain a formerly not accepted argument that does not support ϕ. But since any such argument is skeptically entrenched to degree x, this is not possible: $\mathcal{N}'$ is only of degree $x - 1$. We have derived a contradiction.

42.3: Suppose $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$ and ϕ is skeptically entrenched to degree x in $\mathcal{N} \cap \mathcal{N}''_a$ given $\mathcal{N}''_u, E$ where $E \in S(\mathcal{N}''_u)$ or skeptically entrenched to degree x in $\mathcal{N}' \cap \mathcal{N}''_a$. In the first case, every extension E' of $\mathcal{N}$ contains at least x arguments which are defended against any indirect attacker by at least x arguments that are accepted under E' given $\mathcal{N}''_u, E$. In addition, any rejected argument is either undefended and attacked by at least x arguments from A, E or each of its indirect defenders is attacked by at least x arguments from A, E. Since $\mathcal{N}'$ contains maximally $x - 1$ attacks on arguments in $\mathcal{N}_a$, every extension of $\mathcal{N}''_a$ must contain at least one argument that supports ϕ. But then by Proposition 29, any such a is accepted in $\mathcal{N}''$ and hence $\mathcal{N}'' \vDash SB_S\phi$. The argument for the second case is analogous. □

6.3 Dynamics of Strong Belief

Finally, we consider strong belief. Here we also have to require that no new arguments be added and no new extensions arise but luckily we don't have to worry about entrenchment: removing an argument from an extension will never lead to loss of a strong belief as *every* accepted argument is required to support ϕ. Thus even if we reduce the argumentation framework's extensions to empty sets, an existing strong belief will not be lost (however,the agent will then believe everything). In this respect strong belief is the dynamic mirror image of credulous belief: it is immune to destructive updates but vulnerable to expansive ones. For the strong believer, the content of her belief is so strongly embedded in her mental models that only the rise of a new theory or an expansion of her existing theories can destroy her beliefs.

Proposition 42. Let $\mathcal{AS}$ be an Argument-Support model and $S \in \{\mathcal{AD}, \mathcal{CO}, \mathcal{GR}, \mathcal{PR}\}$ a semantic such that $\mathcal{AS} \models STB_S\phi$. Then $STB_S\phi$ is preserved under an update $[\cup\mathcal{N}']$ if any of the following conditions hold:

1. $\mathcal{AS}''_u \models STB_S\phi$ or $\mathcal{AS}''_a \models STB_{S,\mathcal{N}''_u,E}\phi$ where $E \in S(\mathcal{N}''_u)$.
2. ϕ is reverse entrenched to degree x in $\mathcal{N}$ and $\mathcal{N}'$ and $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$.
3. $\mathcal{N}'$ is of degree $< x$ to $\mathcal{N}$ and ϕ is reverse entrenched to degree x in $\mathcal{N} \cap \mathcal{N}''_a$ given $\mathcal{N}''_u, E$ where $E \in S(\mathcal{N}''_u)$ or reverse entrenched to degree x in $\mathcal{N}' \cap \mathcal{N}''_a$.

Proof. Essentially the same argument as for Proposition 41 insofar as it pertains to reverse entrenchment. □

Chapter 7

Conclusion

We set out to study how agents can form and change opinions based on the process of argumentation. What follows gives a brief recap of what this work contributed to this area.

7.1 Summary

In chapter 2 we assembled a toolkit from the literature on abstract argumentation theory that we used throughout this work. Chapter 3 introduced a basic modal logic for argumentation developed by Grossi [2010] and built on it by equipping it with nominals and (odd) eventualities to yield Hybrid Argumentation Logic (HAL).

HAL is one of the main contributions of this work in that it allows us to express many many of the concepts many of the concepts from argumentation theory, including reference to individual arguments, the grounded extension and the presence of odd and even cycles. In Chapters 4 and 5 we put HAL to work to develop logics for argument based belief and argument dynamics.

For the first task we took inspiration from Grossi and van der Hoek [2014] in combining HAL with a logic for the content of propositions in a product logic. For this purpose we also developed a theory of support between arguments and propositions. We place three conditions on this support relation: Firstly, arguments that support jointly inconsistent propositions are in conflict; secondly, when an argument a attacks another argument b then a also attacks any stronger version of b and any stronger version of a also attacks b; thirdly, every argument has to support something. Integrating this notion of support in the product logic yields Argument Support Logic (ASL). We then defined three notions of argument based belief in ASL that reflect how agents might form beliefs from the arguments they are given: credulous believers require only that one accepted argument support the proposition in question; skeptical believers demand that such an argument be a part of every theory (extension) they consider possible; strong believers do not believe a proposition if not all of the arguments they accept point to it. We studied the well-behavedness of these belief operators with respect to the classical $KD45$ axioms for belief: credulous belief and skeptical belief turned out to be weaker than those axioms require, strong belief satisfies all of them. Thus credulous believers may be inconsistent, skeptical believers' do not close their beliefs under conjunctions and both may believe nothing at all, even if they *know* something.

In Chapter 5 we developed argument framework union and argument framework intersection as a semantic for dynamics of HAL. We showed that these two operations cover all of the notions found in the literature on argumentation dynamics. We then added two dynamic modalities based on these semantics to HAL to obtain $HAL^{\cap}_{\cup}$. We proved that $HAL^{\cap}_{\cup}$ is very expressive - at least as expressive as first order logic - and concluded that it is undecidable.

Finally, we employ the theory developed in chapters 3 to 5 to study the dynamic behaviour of argument based belief. We find that different strategies are needed to destroy credulous, skeptical and strong beliefs: credulous belief is never lost in the face of a new alternative theory. On the other hand strong belief is immune to attacks on individual accepted arguments. Skeptical belief is vulnerable to both approaches. In addition, we defined three notions of entrenchment (weak entrenchment, strong entrenchment, reverse entrenchment) that describe how difficult it is to change an agent's opinion in terms of the minimum number of arguments needed.

7.2 Future Work

On the other hand, there are a great many questions that we could not answer here due to time and space constrains. Most importantly, we conjectured but had to leave open the question of completeness of HAL. In future work we hope to complete the proof sketch we provided. As for the completeness of ASL an investigation of the behaviour of hybrid logics in products of logics would be required. We also left aside the questions of decidability and complexity of the satisfaction problem for these logics.

Likewise, finding an axiomatization for $HAL^{\cap}_{\cup}$, its precise relation to Local Graph Modifier Logic (LGML) (Aucher et al. [2009]) and its second order correspondence language would be an interesting investigation in its own right. Such a study could go beyond the context of argumentation and take a wider perspective of HAL as a language to describe general (directed) graphs. Akin to the Global Graph Modifier Logic, a fragment of $LGML$ that can be reduced to PDL without eventualities Aucher et al. [2009], one would look for decidable fragments of full $HAL^{\cap}_{\cup}$ by placing restrictions on the expressive power of the union and intersection operators.

A different direction of research might explore the added value of propositional variables ranging over arguments give to argumentation: we have seen that propositional variables allow one to pick out all arguments with a certain properties. In the context of dynamics, this could allow one to equip arguments with dynamic properties, akin to the preconditions used in action models of

Dynamic Epistemic Logic (see Baltag and Renne [2016]). E.g. an argument a might attack all arguments with property p. Thus whenever an update adds such an argument b that satisfies p to the argumentation framework, an attack from a to b also has to be added. Such dynamic properties would allow even further expressiveness: for example, a might have a dynamic property $\overleftarrow{\Diamond} i$, meaning that it will attack every added argument that attacks the argument named by i.

Finally, deepening the study of dynamics of argument based belief conducted in Chapter 6 would be of interest. This would entail not only exploring different kinds of belief change such as belief gain but also obtaining not only sufficient but also necessary conditions for preservation of belief. Such an investigation would certainly have to start by extending the known invariance results of argumentation dynamics - a worthy topic for a study in its own right.

7.3 The Fate of Atlantis

After thorough debate the Atlantians decided to base their politics on their scientists' recommendations. They elected new leaders and conducted great efforts to halt the process of sea level rise threatening their nation. Driving enormous rods deep into the earth they stabilized their island's geology. Alas, the Oracle of Delphi is never wrong: Displeased with the hubris of mortals attempting to bind the forces of nature, Zeus decided to punish Atlantis with a more traditional disaster. Luckily, such divine catastrophies are usually preceded by bad omens such as heavy rains of various sorts of amphibians. Thus the Atlantians were able to escape in time and bestow their story to the world. Atlantis however was swallowed by the sea in a single day, just as predicted. And so castles made of sand slip into the sea, eventually.

Bibliography

Carlos Areces and Balder ten Cate. 14 hybrid logics. In Patrick Blackburn, Johan Van Benthem, and Frank Wolter, editors, *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*, pages 821 – 868. Elsevier, 2007. doi: https://doi.org/10.1016/S1570-2464(07)80017-6. URL `http://www.sciencedirect.com/science/article/pii/S1570246407800176`.

Guillaume Aucher, Philippe Balbiani, Luis Fariñas Del Cerro, and Andreas Herzig. Global and local graph modifiers. *Electronic Notes in Theoretical Computer Science*, 231(C):293–307, 2009. ISSN 1571-0661.

Alexandru Baltag and Bryan Renne. Dynamic epistemic logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016.

Alexandru Baltag, Bryan Renne, and Sonja Smets. The logic of justified belief, explicit knowledge, and conclusive evidence. *Annals of Pure and Applied Logic*, 165(1):49–81, 2014.

Alexandru Baltag, Nick Bezhanishvili, Aybüke Özgün, and Sonja Smets. Justified belief and the topology of evidence. In Jouko Väänänen, Åsa Hirvonen, and Ruy de Queiroz, editors, *Logic, Language, Information, and Computation*, pages 83–103, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg. ISBN 978-3-662-52921-8.

Pietro Baroni and Massimiliano Giacomin. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence*, 171(10): 675 – 700, 2007. ISSN 0004-3702. doi: https://doi.org/10.1016/j.artint.2007.04.004. URL `http://www.sciencedirect.com/science/article/pii/S0004370207000744`. Argumentation in Artificial Intelligence.

Francesco Belardinelli, Davide Grossi, and Nicolas Maudet. A Formal Analysis of Dialogues on Infinite Argumentation Frameworks. In *24th International Joint Conference on Artificial Intelligence (IJCAI-15)*, pages 861–867, Buenos Aires, Argentina, July 2015. URL `https://hal.archives-ouvertes.fr/hal-01167908`.

Francesco Berto. Impossible worlds. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2013 edition, 2013.

Patrick Blackburn. *Modal logic / Patrick Blackburn, Maarten de Rijke, Yde Venema*. Cambridge tracts in theoretical computer science 53. Cambridge University Press, Cambridge [etc., 2001. ISBN 0521802008.

Patrick Blackburn and Jerry Seligman. What are hybrid languages? In *Advances in Modal Logic*, 1996.

Guido Boella, Souhila Kaci, and Leendert van der Torre. Dynamics in argumentation with single extensions: Abstraction principles and the grounded extension. In Claudio Sossai and Gaetano Chemello, editors, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 107–118, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. ISBN 978-3-642-02906-6.

Guido Boella, Souhila Kaci, and Leendert van der Torre. Dynamics in argumentation with single extensions: Attack refinement and the grounded extension (extended version). In Peter McBurney, Iyad Rahwan, Simon Parsons, and Nicolas Maudet, editors, *Argumentation in Multi-Agent Systems*, pages 150–159, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. ISBN 978-3-642-12805-9.

Walter Carnielli and Marcelo Esteban Coniglio. Combining logics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016.

Balder ten Cate. *Model theory for extended modal languages / Balder David ten Cate*. ILLC dissertation series 2005-01. Institute for Logic, Language and Computation, Amsterdam, 2005. ISBN 9057761335.

Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321 – 357, 1995. ISSN 0004-3702. doi: https://doi.org/10.1016/0004-3702(94)00041-X. URL `http://www.sciencedirect.com/science/article/pii/000437029400041X`.

Paul E. Dunne and T.J.M. Bench-Capon. Complexity and combinatorial properties of argument systems. Technical report, 2001.

Paul E. Dunne and T.J.M. Bench-Capon. Coherence in finite argument systems. *Artificial Intelligence*, 141(1):187 – 203, 2002. ISSN 0004-3702. doi: https://doi.org/10.1016/S0004-3702(02)00261-8. URL `http://www.sciencedirect.com/science/article/pii/S0004370202002618`.

Dov M. Gabbay and Valentin B. Shehtman. Products of modal logics, part 1. *Logic Journal of the IGPL*, 6(1):73–146, 1998. doi: 10.1093/jigpal/6.1.73. URL `http://dx.doi.org/10.1093/jigpal/6.1.73`.

Davide Grossi. On the logic of argumentation theory. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 - Volume 1*, AAMAS '10, pages 409–416, Richland, SC, 2010. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 978-0-9826571-1-9. URL http://dl.acm.org/citation.cfm?id=1838206.1838264.

Davide Grossi. Argumentation in the view of modal logic. In Peter McBurney, Iyad Rahwan, and Simon Parsons, editors, *Argumentation in Multi-Agent Systems*, pages 190–208, Berlin, Heidelberg, 2011a. Springer Berlin Heidelberg. ISBN 978-3-642-21940-5.

Davide Grossi. An application of model checking games to abstract argumentation. In Hans van Ditmarsch, Jérôme Lang, and Shier Ju, editors, *Logic, Rationality, and Interaction*, pages 74–86, Berlin, Heidelberg, 2011b. Springer Berlin Heidelberg. ISBN 978-3-642-24130-7.

Davide Grossi. Fixpoints and iterated updates in abstract argumentation. In *Proceedings of the Thirteenth International Conference on Principles of Knowledge Representation and Reasoning*, KR'12, pages 65–74. AAAI Press, 2012. ISBN 978-1-57735-560-1. URL http://dl.acm.org/citation.cfm?id=3031843.3031852.

Davide Grossi. Abstract argument games via modal logic. *Synthese*, 190(1): 5–29, Dec 2013. ISSN 1573-0964. doi: 10.1007/s11229-012-0237-1. URL https://doi.org/10.1007/s11229-012-0237-1.

Davide Grossi and Wiebe van der Hoek. Justified beliefs by justified arguments. In *Proceedings of the Fourteenth International Conference on Principles of Knowledge Representation and Reasoning*, KR'14, pages 131–140. AAAI Press, 2014. ISBN 1-57735-657-8, 978-1-57735-657-8. URL http://dl.acm.org/citation.cfm?id=3031929.3031947.

Jens Ulrik Hansen. A hybrid public announcement logic with distributed knowledge. *Electronic Notes in Theoretical Computer Science*, 273:33 – 50, 2011. ISSN 1571-0661. doi: https://doi.org/10.1016/j.entcs.2011.06.011. URL http://www.sciencedirect.com/science/article/pii/S157106611100082X. International Workshop on Hybrid Logic and Applications 2010 (HyLo 2010).

Barteld Kooi, Gerard Renardel de Lavalette, and Rineke Verbrugge. Hybrid logics with infinitary proof systems. *Journal of Logic and Computation*, 16 (2):161–175, April 2006. ISSN 0955-792X.

Beishui Liao, Li Jin, and Robert C. Koons. Dynamics of argumentation systems: A division-based method. *Artificial Intelligence*, 175(11):1790 – 1814, 2011. ISSN 0004-3702. doi: https://doi.org/10.1016/j.artint.2011.03.006. URL http://www.sciencedirect.com/science/article/pii/S0004370211000518.

Jean-Guy Mailly. *Dynamics of argumentation frameworks*. PhD thesis, 2015. URL http://www.theses.fr/2015ARTO0402. Thse de doctorat dirige par Konieczny, Sbastien Informatique Artois 2015.

Chenwei Shi, Sonja Smets, and Fernando R. Velázquez-Quesada. Argument-based belief in topological structures. 251, 2017. ISSN 2075-2180.

J. van Benthem and E. Pacuit. Dynamic logics of evidence-based beliefs. *Studia Logica: An International Journal for Symbolic Logic*, 99(1/3):61–92, 2011. ISSN 00393215, 15728730. URL http://www.jstor.org/stable/41475196.