Formalising human argumentation about action proposals: a case-study

Masters thesis, 45 ECTS

Onno Treep

Supervisor and first reviewer: prof. dr. mr. H. Prakken
Second reviewer: dr. G.A.W. Vreeswijk

March 2015

Contents

1 Introduction **2**

2 Modelling Techniques for Argumentation **5**
2.1 Formalising argumentation in general . . . 5
2.1.1 Abstract Argumentation . . . 5
2.1.2 Structured Argumentation . . . 6
2.2 Formalising argumentation about action proposals . . . 10
2.2.1 Computational Representation of Practical Argument . . . 11
2.2.2 Argumentation about action proposals based on $ASPIC^+$ . . . 12
2.2.3 A modelling technique based on Assumption-Based Argumentation 13

3 Formal Reconstruction of a Debate **16**
3.1 Introduction to the debate and online sources . . . 16
3.2 Interpretation and logical structure . . . 18
3.3 Formalisations on the basis of $ASPIC^+$ . . . 23
3.3.1 A formalisation in $ASPIC^+$ . . . 23
3.3.2 Reconstructing the formalisation in ABA . . . 24
3.3.3 Reconstructing the formalisation in Classical-Logic Argumentation 28
3.3.4 Extensions . . . 33
3.4 A formalisation based on Assumption-Based Argumentation . . . 37
3.4.1 Extended decision framework . . . 37
3.4.2 ABA framework . . . 38

4 Discussion **41**
4.1 Defeasibility of conclusions . . . 41
4.1.1 Methods of formalising . . . 41
4.1.2 Relevance . . . 44
4.2 Defeasibility of rules . . . 44
4.2.1 Methods of formalising . . . 44
4.2.2 Relevance . . . 46
4.3 Use of Preferences . . . 47
4.3.1 Approaches to formalising preferences . . . 47
4.3.2 Relevance . . . 48
4.4 Summary . . . 48

5 Conclusion **50**

References **52**

A A Text Version of the Structured Climate Debate **54**

B $ASPIC^+$ Formalisation and Reconstructions **59**

1 Introduction

One of many fields of research in AI concerns the process of formalising reasoning and argumentation. In this process, reasons for conclusions that are drawn are made explicit and the way conflicting information is handled is specified. In the past decades, numerous formal theories of argumentation have been developed. A formal theory of argumentation can be used in designing autonomous systems. Those systems can be, for example, artificial agents that reason and argue with each other. Other theories provide a method for formally representing natural arguments in human argumentation. This last category will be the topic of research in this thesis. Considering a debate or discussion in natural language, it is interesting to try and represent that debate or discussion in a formal way. Doing this can help to create an abstract view on the discussion. That can be helpful by providing more insight in how arguments relate to each other and which positions in the debate can be held.

A distinction is made between abstract and structured argumentation [3]. In structured argumentation, arguments are constructed from knowledge represented in a formal language. In the majority of approaches in structured argumentation, the notion of argument is more or less the same: an argument consists of a claim (or conclusion) and a number of premises by which the claim is supported. Furthermore, a definition of attack is provided, indicating which arguments are in conflict. In abstract argumentation, neither the structure of an argument nor the way arguments attack each other is specified. By assuming this information is given, abstract argumentation can focus on determining which arguments are acceptable.

Current debate There is disagreement when it comes to the best way of representing arguments and attacks. According to some theories, an argument can only be disputed if its support is contradicted [18, 10]. In other words, an attack has to be directed at the support. If one accepts the support of an argument, one has to accept the conclusion, as well. Others want to leave the possibility open to disagreement with an argument without denying (part of) its support [11]. According to those theories, an argument can be disputed based solely on an arguments' conclusion or on its inference (from the support to the conclusion).

Another topic of debate is whether or not to use preference relations over arguments. With preference relations, it is possible to state which arguments are more important than others. A formalisation including a notion of preference is said to represent argumentation more realistically than a formalisation without such a notion [11]. However, there is no consensus on this subject: it is also claimed that taking preference into account is less attractive from a computational point of view [18], or only solves one problem (dealing with conflicting information) at the cost of creating another (determining the 'right' preference ordering) [4].

Reasoning about action proposals In earlier stages, research on argumentation focused mostly on reasoning about beliefs. More recently, reasoning about what to do (practical reasoning) has also been a topic of research. Reasoning about actions is an important part of human reasoning in general, since humans reason about what is best to do in a large variety of situations. These can be personal choices in everyday life ('What should I have for lunch?' or 'Which route should I take to work?') as well as larger situations in, for example, politics ('What should we do about the situation in Ukraine?'). Thus in an attempt to model human argumentation, it is sensible to try and formalise practical reasoning as well as reasoning about beliefs.

Although formalising practical reasoning requires a different approach than reasoning about beliefs, the same topics of debate play a role considering the proposed solutions. There is no consensus on which kinds of attacks should be possible, and neither on the use of preferences.

Thesis The variety of modelling techniques gives rise to certain questions. Why is there such a variety? Are the different techniques equally suitable for constructing argumentation models? If that is not the case, does it depend on the application which technique is most suitable?

The aim of this thesis is to investigate which way of formalising is most suitable for constructing a realistic model of argumentation about action proposals. More specifically, we will investigate for a number of features whether a realistic model includes them or not. We will focus on human argumentation. Furthermore, we will not focus on models which describe argumentation; we are interested in models that can help human debaters in an actual discussion.

The research in this thesis will be focused on the following questions:

1. Is the most realistic model created by a modelling technique in which *attacks on conclusions* are reduced to premise-attacks, or by one which does not, and why?
2. Is the most realistic model created by a modelling technique in which *attacks on inferences* are reduced to premise-attacks, or by one which does not, and why?
3. Is the most realistic model created by a modelling technique which includes *preferences*, or by one which does not, and why?

In order to answer the questions, a debate about proposals for action will be formalised using different techniques. In particular, formalisations in $ASPIC^+$[11], assumption-based argumentation (ABA) [18], and a form of classical-logic argumentation [10] will be discussed. The different models will be compared, with a focus on the features mentioned in the research questions above. Attention will be given to the following criteria:

- At least all of the implicit elements of arguments should be made explicit in a formalisation. The more implicit elements are not made explicit, the less we consider a formalisation natural or realistic.

- The number of elements of arguments that were not part of an original debate but added in the formalising process, should be minimised. The more elements are added, the less we consider a formalisation natural or realistic.
- A formalisation should be a logically correct or sound representation of the original debate. All arguments in the original debate should be formalised as an argument following the notion of argument from the chosen technique. Furthermore, sub-argument relations as well as conflicts should be between the same pairs of arguments as in the original debate. The less this is the case, the less we consider a formalisation natural or realistic.

Another criterium could be the complexity of the underlying logics. However, whether the complexity of the underlying logics is more important than the previously mentioned criteria depends on the applications of the system: if the goal is to implement (autonomous) software, this might be a more important issue, since it is easier to implement a model based on less complex logics. However, since our focus is human argumentation, and our goal is to find the best way to formalise natural arguments, the complexity of the underlying logics is of less importance than the two criteria mentioned above.

This thesis is structured as follows. In section 2, several formalisation techniques will be discussed. First some background is provided about techniques for formalising practical reasoning, as these are more general and form the basis for the techniques described next. What follows is the description of some theoretical frameworks for formalising reasoning about action proposals. Section 3 considers the discussion of an online debate on how to respond to climate change. It is formalised using the different techniques described in section 2. An $ASPIC^{+}$ formalisation is discussed in section 3.3, as well as reconstructions in ABA and classical logic based argumentation. Section 3.4 discusses an alternative ABA-based formalisation, constructed using the techniques from section 2.2.3. In section 4 the models will be discussed on different aspects. The question whether the possibility of attacks on conclusions makes for better models, is discussed in section 4.1. Section 4.2 discusses whether the possibility of defeasible rules leads to more realistic models. Section 4.3 will be directed at the question whether the use of preferences is a valuable addition to such a model. We will conclude in section 5.

2 Modelling Techniques for Argumentation

In this section, various methods to formalise argumentation will be presented. Reasoning about actions is different from reasoning about beliefs, and requires a different kind of formalisation. Since acting means modifying the environment, reasoning about which action is best in a given situation means reasoning about which possible modification in the environment is most desirable. This is different from reasoning about what is true in the current state of the environment. For example, there is a notion of choice: deciding which action is best, based on available information, is a matter of choice. In reasoning about beliefs, choice does not play this role. Logics for reasoning about beliefs are not directly suitable for formalising reasoning about actions. However, these logics are still useful and form a basis for further development in approaches to this problem. Therefore, we will first discuss some general formalising methods. Next, we will discuss several modelling techniques for reasoning about actions.

2.1 Formalising argumentation in general

This part is structured as follows: first an introduction will be given to Dung's work on abstract argumentation frameworks. After that, structured argumentation will be discussed, as well as some models in that field.

2.1.1 Abstract Argumentation

Abstract argumentation is abstract in the sense that the internal structure of arguments is ignored and arguments are viewed as atoms. An important influence are Dung's abstract argumentation frameworks (AF) [6]. Dung's frameworks define, for given sets of arguments and defeat relations, several semantics that sets of arguments should satisfy to be defensible. AFs are visualised by directed graphs in which nodes represent arguments and arcs represent attack relations between these arguments - the latter indicating conflicts. Using AFs, it is possible to compare the formal systems discussed in section 2.1.2. This is because the output of these systems can be translated into AFs. To be able to explain this, first some definitions will be given.

Definition 2.1. An *argumentation framework* is a pair $\langle Args, defeat \rangle$, where *Args* is a set of arguments and *defeat*[1] is a binary relation on *Args*.

An argument *A strictly defeats B* iff *A* defeats *B* and *B* does not defeat *A*. A set *S* of arguments defeats an argument *A* iff some argument in *S* defeats *A*. *S* defeats a set *S'* of arguments iff it defeats a member of *S'*.

Definition 2.2.

[1] Originally, *defeat* is called *attacks*. It is renamed to avoid confusion in the next sections.

- A set $S \in Args$ is *conflict-free* if there are no arguments $A, B \in S$ such that A defeats B.
- An argument A is *acceptable* with respect to a set S iff for each argument B holds: if B defeats A then B is defeated by S.

Then a conflict-free set $S \subseteq Args$ is:

- an *admissible extension* iff each $X \in S$ is acceptable with respect to S.
- a *complete extension* iff each $X \in S$, whenever X is acceptable with respect to S.
- a *preferred extension* iff S is a set inclusion maximal complete extension.
- a *stable extension* iff S is preferred and $\forall Y \notin S, \exists X \in S$s.t.$(X, Y) \in$ *defeat*.
- a *grounded extension* iff S is the set inclusion minimal complete extension.

An argument is *sceptically* justified under some semantics iff it belongs to all of the corresponding extensions. It is *credulously* justified under the semantics iff it belongs to at least one of the corresponding extensions.

2.1.2 Structured Argumentation

As mentioned before, in structured argumentation the aim is to give a notion of argument and of attack, with which argumentation can be formalised. Different developed systems for formalising argumentation will be described below. The systems are built around a logical language and a set of inference rules defined over this language. Dung's semantics are used to determine which sets of arguments are winning. Differences between the systems can be found in the notions of argument, the use of preferences, and the types of attack possible between arguments.

***ASPIC*$^+$** The *ASPIC*$^+$ framework [11] differs from other approaches in that it makes a distinction between two kinds of inference rules: strict rules and defeasible rules. *Strict* rules are rules that cannot be disputed. If there is a strict inference from a set of premises and one accepts the premises, then the conclusion has to be accepted, as well. *Defeasible* rules express what usually (but not always) can be inferred. Hence, if a defeasible inference is made from a set of premises and one accepts the premises, it is still possible to deny the conclusion. A second difference in the *ASPIC*$^+$ framework compared to other approaches is the use of preference relations over arguments. If an argument is more preferred than others, then some categories of attack against this argument will never be successful. The *ASPIC*$^+$ framework is meant to generate abstract argumentation frameworks. Winning sets of arguments can then be determined as in [6].

Definition 2.3. An *argumentation system* is a tuple $AS = \langle \mathcal{L}, \mathcal{R}, n, ^{-} \rangle$, where:

- $\mathcal{L}$ is a logical language.
- $\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_d$ is a set of strict ($\mathcal{R}_s$) and defeasible ($\mathcal{R}_d$) rules of the form $\varphi_1, ..., \varphi_n \to \varphi$ and $\varphi_1, ..., \varphi_n \Rightarrow \varphi$, respectively (where φ_i and φ are meta-variables ranging over wff in $\mathcal{L}$), and $\mathcal{R}_s \cap \mathcal{R}_d = \varnothing$;
- n is a partial function such that $n : \mathcal{R}_d \to \mathcal{L}$.
- $^-$ is a function from $\mathcal{L}$ to $2^{\mathcal{L}}$ such that:
  - φ is a *contrary* of ψ if $\varphi \in \overline{\psi}, \psi \notin \overline{\varphi}$.
  - φ is a *contradictory* of ψ (denoted by '$\varphi = -\psi$') if $\varphi \in \overline{\psi}, \psi \in \overline{\varphi}$.
  - each $\varphi \in \mathcal{L}$ has at least one contradictory.

$n(r)$ is a wff in $\mathcal{L}$ which says that the defeasible rule $r \in \mathcal{R}$ is applicable. This makes attacks on rules possible: an argument claiming $\overline{n(r)}$ attacks the inference step on the corresponding rule.

Definition 2.4. Given an argumentation system $AS = \langle \mathcal{L}, \mathcal{R}, n, ^- \rangle$, a *knowledge base* $\mathcal{K} \subseteq \mathcal{L}$ consists of two disjoint subsets of axioms $\mathcal{K}_n$ and ordinary premises $\mathcal{K}_p$.

Definition 2.5. An *argumentation theory* is a tuple $AT = \langle AS, \mathcal{K} \rangle$, where AS is an argumentation system and $\mathcal{K}$ is a knowledge base in AS.

In the following definitions, $\texttt{Prem}(A)$, $\texttt{Conc}(A)$ and $\texttt{Sub}(A)$ indicate the premises, conclusion and sub-arguments of A, respectively.

Definition 2.6. An *argument* A based on an $AT = \langle AS, \mathcal{K} \rangle$ where $AS = \langle \mathcal{L}, \mathcal{R}, n, ^- \rangle$, is:

- φ if $\varphi \in K$ with: $\texttt{Prem}(A) = \{\varphi\}, \texttt{Conc}(A) = \varphi, \texttt{Sub}(A) = \{\varphi\}$
- $A_1, ..., A_n \to / \Rightarrow \psi$ if $A_1, ..., A_n$ are arguments such that there exists a strict/defeasible rule $\texttt{Conc}(A_1), ..., \texttt{Conc}(A_n) \to / \Rightarrow \psi$.
  $\texttt{Prem}(A) = \texttt{Prem}(A_1) \cup ... \cup \texttt{Prem}(A_n)$,
  $\texttt{Conc}(A) = \psi$,
  $\texttt{Sub}(A) = \texttt{Sub}(A_1) \cup ... \cup \texttt{Sub}(A_n) \cup \{A\}$.

A distinction is made between attack and defeat: attacks indicate conflict between arguments, defeats are successful attacks. Defeat here is the same as attack in [6]. Attacking arguments is possible in three different ways: by an *undermining attack* on an ordinary premise, a *rebutting attack* on a conclusion of a defeasible inference step, or an *undercutting attack* on a defeasible inference step.

Definition 2.7. An argument A attacks an argument B iff one of the following statements holds:

- A undercuts B (on B') iff $\texttt{Conc}(A) \in \overline{n(r)}$ for some $B' \in \texttt{Sub}(B)$ such that B''s top rule r is defeasible.

- A rebuts B (on B') iff $\texttt{Conc}(A) \in \overline{\varphi}$ for some $B' \in \texttt{Sub}(B)$ of the form $B_1'', ..., B_n'' \Rightarrow \varphi$.
- A undermines B (on φ) iff $\texttt{Conc}(A) \in \overline{\varphi}$ for some $\varphi \in K_p$ of B.

While undercutting attacks always succeed as a defeat, rebuttals and underminers rely on a preference ordering to succeed. This preference ordering, which is given by the user, is a binary ordering on the set of arguments and is used to determine whether attacks are successful.

Definition 2.8 (Successful rebuttal, successful undermining and defeat)**.**

- A *successfully rebuts* B if A rebuts B on B' and A is at least as preferred as B'.
- A *successfully undermines* B if A undermines B on φ and A is at least as preferred as φ.
- A *defeats* B if A undercuts, successfully rebuts or successfully undermines B.

Definition 2.9. Given some AT, a *structured argumentation framework* SAF is a tuple $\langle \mathcal{A}, \mathcal{C}, \preceq \rangle$ where:

- $\mathcal{A}$ is the smallest set of all finite arguments constructed from KB in AS;
- $\preceq$ is an ordering on $\mathcal{A}$;
- $(X, Y) \in \mathcal{C}$ for $X, Y \in \mathcal{A}$ iff X attacks Y.

Given an $SAF = \langle \mathcal{A},\mathcal{C},\preceq \rangle$, an $AF = \langle \mathit{Args,defeat} \rangle$ can be defined, taking as $Args$ and *defeat* the arguments and defeat relation as determined by definition 2.9.

Definition 2.10. An $AF = \langle \mathit{Args,defeat} \rangle$ corresponds to a $SAF = \langle \mathcal{A},\mathcal{C},\preceq \rangle$ iff *defeat* is the defeat relation on $\mathcal{A}$ determined by $\langle \mathcal{A}, \mathcal{C}, \preceq \rangle$.

Now it is possible to justify statements in a debate on the basis of the semantics defined by Dung.

Definition 2.11. A wff $\varphi \in \mathcal{L}$ is sceptically justified if it is the conclusion of a sceptically justified argument, and credulously justified if it is not sceptically justified and is the conclusion of a credulously justified argument.

Assumption-Based Argumentation In Assumption-Based Argumentation (ABA) frameworks [18], rules in a deductive system together with assumptions and their contraries form the input. Arguments and defeat relations are derived from these deductive rules. Instead of determining which sets of arguments are winning sets, it provides means of determining which sets of *assumptions* are winning sets. There is an equivalence in the sense that winning sets of arguments are those that can be derived from winning sets of assumptions.

Definition 2.12. An ABA framework is a tuple $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C} \rangle$, where:

- $\langle \mathcal{L}, \mathcal{R} \rangle$ is a deductive system consisting of a language L and a set of rules R;
- $\mathcal{A} \subseteq \mathcal{L}$ is a set of assumptions;
- $\mathcal{C}$ is a total mapping from $\mathcal{A}$ into $2^{\mathcal{L}}$, where $\mathcal{C}(\alpha)$ is the *contrary* of α.

Note, there is no consensus about the contrary relation in the literature on ABA: in some publications [18, 7], $\mathcal{C}$ is defined as a total mapping from $\mathcal{A}$ into $\mathcal{L}$ (i.e. each assumption can only have one contrary), while in [9] it is a total mapping from $\mathcal{A}$ into $2^{\mathcal{L}}$. We choose to use the latter definition, since it is the same as the one we use in *ASPIC*$^+$, which makes the models easier to compare.

Rules are of the form $p \rightarrow q$. Facts are represented by rules with an empty body. For example, if '$\rightarrow r$' is a rule, then r is derivable from an empty set of assumptions and therefore always derivable. A framework is *flat* if and only if no assumption is the head of a rule.

Definition 2.13. A *deduction* for $\sigma \in \mathcal{L}$ supported by $S \subseteq \mathcal{L}$ and $R \subseteq \mathcal{R}$, denoted $S \vdash^R \sigma$, is a finite tree with nodes labelled by sentences in $\mathcal{L}$ or by τ[2], the root labelled by σ, leaves either τ or sentences in S, non-leaves σ' with, as children, the elements of the body of some rule in $\mathcal{R}$ with head σ', and R the set of all such rules.

Definition 2.14. An *argument* for $\sigma \in \mathcal{L}$ supported by $A \subset \mathcal{A}$, denoted $A \vdash \sigma$, is a deduction for σ supported by A (and some $R \subseteq \mathcal{R}$).

Definition 2.15. An argument $A_1 \vdash \sigma$ *attacks* an argument $A_2 \vdash \sigma_2$ iff σ_1 is a contrary of one the assumptions in A_2.

Relation to *ASPIC*$^+$ The ABA framework as such can be seen as an instantiation of *ASPIC*$^+$, where:

- all axioms a are strict rules of the form '$\rightarrow a$' (hence $\mathcal{K}_n = \varnothing$)
- $\mathcal{R}_d = \varnothing$

Classical Logic-based Argumentation A number of argumentation systems is based on classical logic. In the system described in [4], well-known elements from classical logic are used to represent arguments and conflicts between them. Arguments are defined as logical proofs: a claim is deductively entailed by a consistent set of premises, the support. We will refer to this form of argumentation as classical argumentation. The most common definition of an argument in classical argumentation (and the one we will use) is the following:

Definition 2.16. An *argument* is a pair $\langle \Phi, \alpha \rangle$ such that:

- $\Phi \nvdash \perp$;

[2] $\tau \notin \mathcal{L}$ represents 'true' and stands for the empty body of rules.

- $\Phi \vdash \alpha$;
- Φ is a minimal set satisfying the above.

The set of premises must be consistent in order to exclude arguments of the form $\langle\{a, \neg a\}, b\rangle$, since these arguments do not correspond to intuitive human reasoning. The third condition prevents unnecessary premises from being part of an argument, and therefore preserves a resemblance to reality as well.

Several attack relations have been defined in the literature on classical argumentation. In [10], Gorogiannis and Hunter analyse seven possible attack relations that have been defined. They set out a number of postulates with which to characterise these attack relations. Furthermore, since their aim is to create an instantiation of Dung's *AF*s that uses classical logic as its language, they propose a number of postulates which an attack relation should satisfy in any semantics. On the basis of these postulates, they argue which attack relations should be included in an argumentation system to obtain desirable behaviour.

Included in this last set of postulates are four consistency postulates. There are only two attack relations that satisfy all consistency postulates in any semantics: the *direct defeater* and the *direct undercut*. For that reason, these are the only attack relations we will consider.

Definition 2.17. A *direct defeater* for an argument $\langle\Phi, \alpha\rangle$ is an argument $\langle\Psi, \beta\rangle$ such that $\beta \vdash \neg\varphi$ for some $\varphi \in \Phi$.

Definition 2.18. A *direct undercut* for an argument $\langle\Phi, \alpha\rangle$ is an argument $\langle\Psi, \beta\rangle$ such that $\beta \equiv \neg\varphi$ for some $\varphi \in \Phi$.

Since direct undercut is a special case of direct defeat (i.e. for each pair of arguments A, B it holds that if A directly undercuts B, then A directly defeats B), the definition of direct defeater will be used as definition for attack.

Definition 2.19. An argument A *attacks* an argument B iff A is directly defeats B.

Classical argumentation form of argumentation can be seen as an instantiation of *ASPIC*$^+$, with:

- $\mathcal{R}_d = \varnothing$
- $\mathcal{K}_n = \varnothing$
- a connective $\neg$ in $\mathcal{L}$ instead of a contrariness relation

2.2 Formalising argumentation about action proposals

In this part, several approaches to formalise reasoning about actions will be discussed. First, a traditional view called the *practical syllogism* will be discussed. Then, a technique based on $ASPIC^+$ will be introduced. Finally, we will discuss a model based on assumption-based argumentation.

2.2.1 Computational Representation of Practical Argument

One possible way to formalise reasoning about actions is using the *practical syllogism*. This is a traditional method of viewing practical reasoning, which has the form of an abduction:

I have goal G

If I perform action A I will reach goal G

Therefore I will perform action A

The problem with this method is the possibility to accept the premises but deny the conclusion. For it to be useful, there must be a method of comparing all alternative actions and goals, and choosing the best of these alternatives. However, in general it is impossible to consider all alternatives. It is also not made explicit what it means if some alternative is the *best* alternative.

Atkinson et al. [1] regard practical reasoning as presumptive argumentation. An argument like the one above can be seen as a presumptive reason for performing an action. However, this presumption can be challenged by considering alternatives and additional consequences of the action. If a better alternative is found, or if some additional consequence of the action makes other goals impossible to reach, the presumption might have to be withdrawn. To see if this is the case, the answers to certain *critical questions* have to be examined.

Atkinson et al. make a distinction between *states*, *goals* and *values*: states are sets of propositions about the world, goals are propositional formulae on those sets of propositions, and values are functions on goals. While states represent the consequences of an action, goals represent the consequences that the agent wanted to achieve. Values represent the reasons for wanting to achieve these goals.

Having made this distinction, the following is the proposed argument scheme:

In the circumstances R

we should perform action A

to achieve new circumstances S

which will realise some goal G

which will promote some value V.

For this to be valid, there are four statements which must hold:

1. R is the case.

2. Performing A in state R results in state S.
3. G is true in S.
4. G promotes V.

Possible attacks on an argument are grouped in the following categories:

- Denial of premises
  One of the four premises above is denied. By denying one of the premises, the whole presumption is challenged. The attack can be simply a denial, or a denial with additional information. For example: 'R is not the case and some other state Q is the case'.
- Alternative ways to satisfy the same value
  An alternative way is proposed to achieve the same desired value. Thereby it shows that action A is not necessary.
- Side effects of the action
  Another goal is reached, which promotes (or demotes) another value, and therefore needs to be considered. Example: 'Another goal H is true in S, which demotes the value V'.
- Interference with other actions
  It is stated that performing the action makes the promotion of another value impossible. Example: 'Performing A makes another action B impossible, while performing B results in a state in which a goal is true that promotes another value W'.
- Disagreements relating to impossibility
  It is denied that a certain element of the argument scheme can possibly exist. For example: 'A is never a possible action'.

2.2.2 Argumentation about action proposals based on $ASPIC^+$

$ASPIC^+$ can be used to argue about action proposals [13]. Arguments in favour of and against action proposals are raised following an argument scheme that bears close resemblance to that in [1]. These arguments refer to sets of good or bad consequences of the proposed actions.

Scheme from good consequences:

If: Action A results in C_1
...
Action A results in C_n
C_1 is good
...

	C_n is good
Then:	Action A is good

Scheme from bad consequences:

If:	Action A results in C_1
	...
	Action A results in C_m
	C_1 is bad
	...
	C_n is bad
Then:	Action A is bad

There are three types of critical questions to ask considering an argument of this scheme from good or bad consequences. These questions correspond to the three types of attacks in $ASPIC^+$ described earlier.

1) Does action A result in $C_1, ..., C_n/C_m$?	(Underminer)
2) Does action A also result in something which is bad (good)?	(Rebuttal)
3) Is there another way to realise C_n/C_m?	(Undercutter)

To decide which of two conflicting action proposals is the better alternative, no utility function has to be used. Instead, preferences can be stated that pref-attack attack relations between the proposals. By pref-attack we mean an attack on an attack of some argument A on some argument B, based on a preference of B over A. An argument stating the preference $B \succeq A$ attacks the attack of A on B. To state a preference between action proposals, the following scheme from [2] can be used:

If:	P_1 has good consequences $C_1^+ = \{c_1, ..., c_m\}$
	P_2 has good consequences $C_2^+ = \{c_n, ..., c_p\}$
	P_1 has bad consequences $C_1^- = \{c_q, ..., c_r\}$
	P_2 has bad consequences $C_2^- = \{c_s, ..., c_t\}$
	$(C_1^+, C_1^-) < (C_2^+, C_2^-)$
Then:	Action proposal P_2 is preferred over P_1

2.2.3 A modelling technique based on Assumption-Based Argumentation

Another way of formalising argumentation about actions is making use of assumption-based argumentation frameworks. Fan et al. [9] present a decision-making framework, in which decisions and the goals they meet are formalised as well as preferences over goals or sets of goals. *Extended decision functions* are defined to select good decisions. A mapping is then created from the defined frameworks and functions into ABA frameworks. Selected decisions are conclusions of arguments in an admissible extension in the corresponding ABA framework. Therefore good decisions can be found by computing admissible arguments.

Decision-Making Framework In the approach of Fan et al. an *extended decision framework* (*edf*) is a tuple $\langle D, A, G, DA, GA, P\rangle$, where:

- D is a non-empty set of possible decisions;
- A is a non-empty set of attributes;
- G is a non-empty set of goals;
- DA is a table in which is indicated which decision has which attribute;
- GA is a table in which is indicated which goal is satisfied by which attribute;
- P is a partial order over goals, representing their preference ranking.

A decision d meets a goal g if and only if there is an attribute a of d that meets g, as indicated by the tables.

To select good decisions, a *most-preferred extended decision function* is defined. This function selects the set of decisions that meet the more preferred goals which no other decisions meet. It does not select decisions for which it holds that there is another decision that meets a more preferred goal.

Mapping into ABA Given an $edf = \langle D, A, G, DA, GA, P\rangle$, a corresponding *most-preferred ABA framework* $ABAF$ can be built. The components of $ABAF = \langle \mathcal{L},\mathcal{R},\mathcal{A},\mathcal{C}\rangle$ are defined as follows:

- $\mathcal{R}$ contains:
  - $da \leftarrow$
    for each decision-attribute pair d,a for which holds that a is an attribute of d, according to table DA;
  - $ga \leftarrow$
    for each decision-attribute pair g,a for which holds that a satisfies g, according to table GA;
  - $dg \leftarrow da, ga$
    for each $d \in D, a \in A, g \in G$, dg meaning that d meets g;
  - $Pg_1g_2 \leftarrow$
    for each pair of goals $g_1, g_2 \in G$ for which g_1 is preferred over g_2;
  - $Nd_2 \leftarrow d_1g, Nd_2g, NX_g^{d_1d_2}$
    for each combination of two decisions $d_1, d_2 \in D$ and a goal $g \in G$;

- $X^{d_1 d_2}_{g_2} \leftarrow d_2 g_1, N d_1 g_1, P g_1 g_2$
  for each combination of two decisions $d_1, d_2 \in D$ and two goals $g_1, g_2 \in G$.

- $\mathcal{A}$ contains:
  - each decision $d \in D$;
  - $NX^{d_1 d_2}_{g}$
    for each combination of two decisions $d_1, d_2 \in D$ and a goal $g \in G$;
  - Ndg
    for each $d \in D, g \in G$.
- $\mathcal{C}$ is such that:
  - $\mathcal{C}(d) = \{Nd\}$;
  - $\mathcal{C}(NX^{d_1 d_2}_{g}) = \{X^{d_1 d_2}_{g}\}$;
  - $\mathcal{C}(Ndg) = \{dg\}$;

The following notation is used:

d, $d \in \mathcal{A}$	Select decision d
$X^{d_1 d_2}_{g}$	There is a goal g', more preferred than g, that is met by d_2 but not by d_1
N	It is not the case that

Given a most-preferred ABA framework $ABAF$, corresponding to an extended decision framework $edf = \langle D, A, G, DA, GA, P\rangle$, for all decisions $d \in D$ holds: $d \in \psi^{E}_{x}(edf)$ iff $\{d\} \in d$ belongs to an admissible set in $ABAF$.

3 Formal Reconstruction of a Debate

In this section an online debate will be interpreted and formalised using various methods from the previous section. We will first introduce the debate and the online source from which we retrieved this debate. Then we will explain our interpretation of the debate. Finally, we will discuss several ways to formalise the debate, which are to be discussed in section 4.

3.1 Introduction to the debate and online sources

Debate: climate change The topic of the debate considered here is climate change caused by humans. Three different positions are argued in the online debate concerning climate change: (A) to act immediately, (B) to prioritise challenges first or (C) to not act at all. In some arguments, the IPCC is mentioned. This is the Intergovernmental Panel on Climate Change. Another abbreviation, AGW, stands for Anthropogenic Global Warming (global warming caused by humans). The original text can be found on debategraph.org[3].

DebateGraph In the tree from DebateGraph, blue boxes contain different positions concerning the issue being discussed in the debate. Green boxes contain supporting statements and red boxes contain attacking statements. Orange boxes refer to other issues discussed on the website, while pale yellow nodes are cross-links from a specific category. We will ignore the last two, since otherwise the debate would become too complex and cluttered with merely indirectly relevant information. Darker blue boxes indicate what is called a *component: a distinct part of a complex position; identified separately and analysed on its own merits.* Since there is only one of this kind in this debate and it is supporting its parent, we will simply interpret it as a normal supporting statement. Finally, there are pink nodes. In DebateGraph, they are confusingly labelled *premise*, which in fact they are not; several pink nodes have children containing supporting statements. The description given for this kind of node is *a co-premise that works with other co-premises to support an argument or conclusion.* Pink nodes that are children of the same parent together form the support of a supporting argument[4].

Rationale Each statement in Rationale is numbered. A statement numbered `3B-b` is the second statement (b) in the support of the second argument from the left (B) on the third level from the root in the argument tree (3). We will give names of statements the following `font`, to keep a more clear distinction from names of arguments.

[3] `http://debategraph.org/Stream.aspx?nid=41848&vt=spacetree&dc=1`
(Default view: `http://debategraph.org/Stream.aspx?nid=41848&vt=bubble&dc=focus`)

[4] In cases in which there are more than three pink children of the same parent, it is not clear if they should all be combined in one argument or in multiple, distinct arguments. However, these cases do not occur in this debate, hence no assumptions have to be made on this matter.

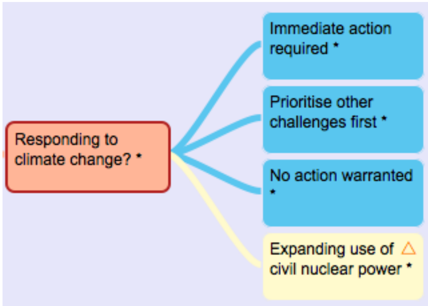


Figure 1: The three positions in DebateGraph

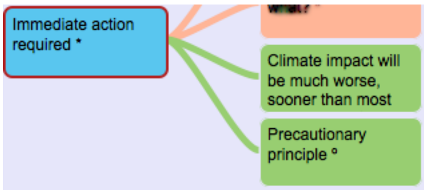


Figure 2: The first position in DebateGraph

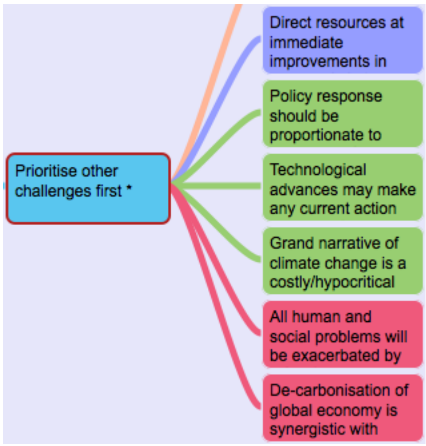


Figure 3: The second position in DebateGraph

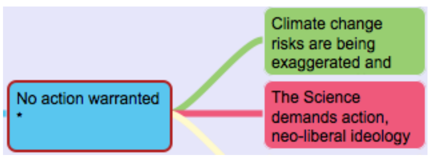


Figure 4: The third position in DebateGraph

3.2 Interpretation and logical structure

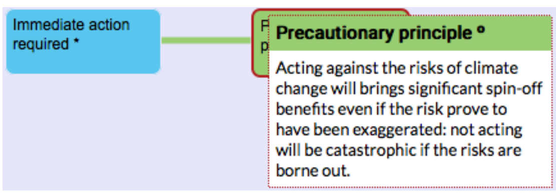
It is possible to view this debate as two debates in one. The first is about the question whether or not it is necessary to act on climate change. While there are no actual arguments directly against acting, there are arguments that dispute (sub-)arguments *for* acting, hence there is discussion about this point. The second topic of discussion presumes that it is necessary to act on climate change, and concerns the timing of action. On the one hand it is argued that we have to act as soon as possible, on the other hand there are arguments for waiting and prioritising several challenges first.

In order to obtain a logical argument structure, a number of arguments are re-interpreted. Also, several implicit arguments are made explicit. Below is an explanation of how the debate from DebateGraph is re-interpreted and visualised in Rationale. (Note that the three positions attack each other, even though there are no attack links between the three positions in the Rationale visualisation.) A number of statements from DebateGraph are listed, and the interpretation is explained right below each statement. The statements in the structured version in this thesis are kept the same as in DebateGraph as much as possible.

In general:

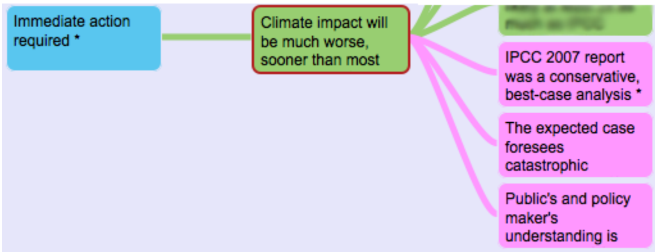
- The three positions in DebateGraph are *Immediate action required*, *Prioritise other challenges first*, and *No action warranted*. These are rephrased as *We should act immediately*, *We should prioritise our challenges before acting*, and *We should not act at all*, respectively, in the structured debate in this thesis. This emphasises the fact that the three positions in the debate express actions to perform, as well as the mutually exclusive and jointly exhaustive nature of the positions. The three positions will be referred to as position A (1A-a), position B (1B-a), and position C (1C-a), below.
- We want to keep the structure in which each argument for one of the positions opposes the other two positions. There are several arguments in favour of *Prioritise other challenges first* in DebateGraph, which oppose *Immediate action required* but not *No action warranted*. These are restructured as counterarguments against position A.

Arguments for and counterarguments against *Immediate action required*:



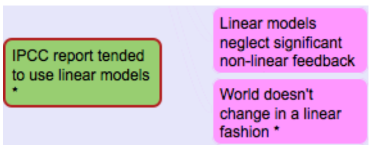
- *Precautionary principle: Acting against the risks of climate change will bring significant spin-off benefits even if the risks prove to have been exaggerated; not acting*

will be catastrophic if the risks are borne out.
This is interpreted as a statement that the degree in which bad consequences of the third position are bad, is higher than the degree in which its good consequences are good. This is expressed in the structured debate as *Continuing the current influence on climate is bad* (2H-b) being stronger than *Saving money, time and resources is good* (2I-b), in the sense that the degree in which continuing the current influence on climate is bad is higher than the degree in which saving money, time and resources is good.



- *IPCC 2007 report was a conservative, best-case analysis*,
  The expected case foresees catastrophic consequences, and
  Public's and policy maker's understanding is based on IPCC reports.
  These three statements, which are in pink boxes, are said to support *Climate impact will be much worse, sooner than most people think* together. However, it is not necessary to state *The expected case foresees catastrophic consequences* to make this point, hence it is left out.
  Furthermore, several steps are left implicit in the inference between these arguments in DebateGraph. *IPCC 2007 report was a conservative, best-case analysis* is interpreted as a sub-argument (7A-a) of *The IPCC reports give a too optimistic view on the expected climate change impact* (6A-b), which, together with *Public's and policy maker's understanding is based on IPCC reports* (6A-a), forms a sub-argument of *What most people think is based on something which gives a too optimistic view on climate change impact* (5A-a) in the interpretation in Rationale.

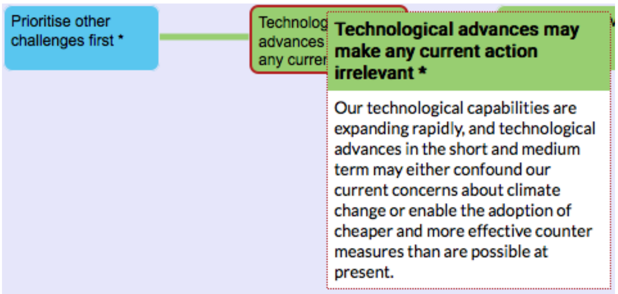
- *The IPCC report tended to use linear models.*
  Left implicit here is the idea that a report based on linear models is a conservative, best-case analysis. This is made explicit (8C-b).



- *The world does not change in a linear fashion.*

In combination with *Linear models neglect significant non-linear feedback*, this should lead to *The IPCC report tended to use linear models*. The interpretation given here is that it holds an implicit idea that linear models give a more positive view than other models and that it implies *A report based on linear models is a conservative best-case analysis*. Therefore, *The world ... fashion* is made a sub-argument (9G-a) of 8C-b, in combination with the implicit idea which is made explicit in 9G-b.

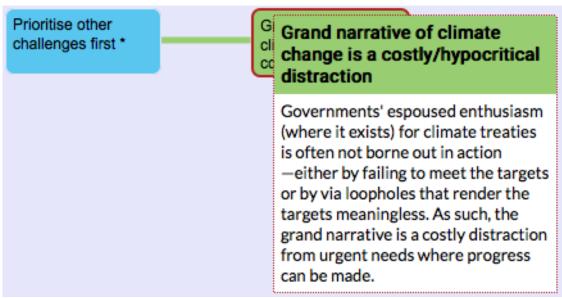
Arguments for and counterarguments against *Prioritise other challenges first*:



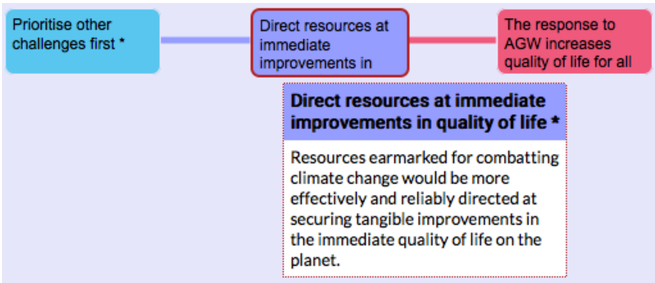
- *Technological advances may make any current action irrelevant. Our technological capabilities are expanding rapidly, and technological advances in the short and medium term may either confound our current concerns about climate change or enable the adoption of cheaper and more effective counter measures than are possible at present.*

  1) In DebateGraph, this argument is an argument for position B. But while this statement opposes position A, it does not oppose position C. Since we want all arguments for position B to be counterarguments against the other two positions, this argument is interpreted as a counterargument against position A (3D-a).

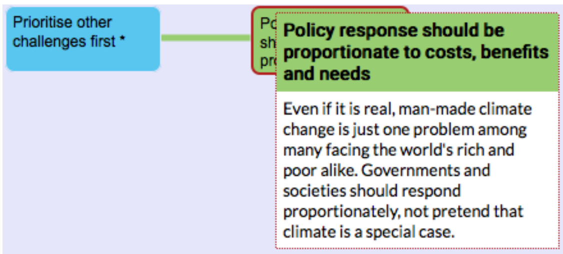
  2) *Confounding our current concerns about climate change* and *enabling the adoption of cheaper and more effective counter measures than are possible at present*, while stated together, are two different effects that do not influence each other. Therefore, these are separated into two sub-arguments (4B-a and 4C-a).



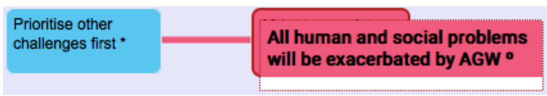
- *Grand narrative of climate change is a costly/hypocritical distraction. Governments' espoused enthusiasm (where it exists) for climate treaties is often not borne out in action – either by failing to meet the targets or via loopholes that render the targets meaningless. As such, the grand narrative is a costly distraction from urgent needs where progress can be made.*
  This is used as an argument for position B. However, while it is a counterargument against position A, it does not counter position C. Since we interpret arguments of one position as countering the other two positions, this argument cannot be an argument for position B, and is interpreted as a counterargument against position A (3E-a).



- *Resources earmarked for combatting climate change would be more effectively and reliably directed at securing tangible improvements in the immediate quality of life on the planet.*
  Again, this statement is interpreted as a normal supportive argument. Left implicit is that one needs to prioritise challenges in order to change the way resources are used, and that changing this way in a positive way is good. This is made explicit (see 2G-a, 2G-b and 3F-a).

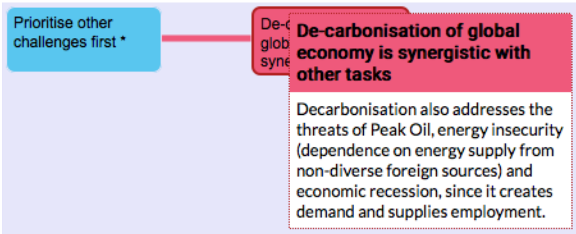


- *Policy response should be proportionate to costs, benefits and needs.*
  Left implicit is that prioritising challenges first leads to a more proportionate response. This is made explicit (2F-a).



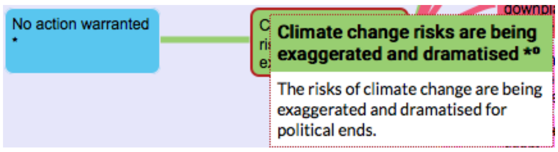
- *All human and social problems will be exacerbated by AGW.*
  This does not attack position B directly, and it can be seen as an attack on position

C. It is interpreted as an argument for position A, which attacks positions B and C indirectly (see 2C-a, 2C-b and 3C-a).

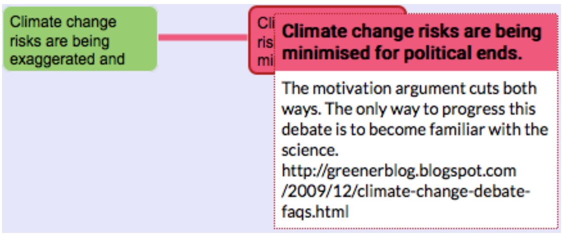


- *Decarbonisation also addresses the threats of Peak Oil, energy insecurity (dependence on energy supply from non-diverse foreign sources) and economic recession, since it creates demand and supplies employment.*
  The same holds for this argument. It is interpreted as an argument for position A, which attacks positions B and C indirectly (see 2B-a, 2B-b, 3B-a and 3B-b).

<u>Arguments for and counterarguments against *No action warranted*:</u>



- *The risks of climate change are being exaggerated and dramatised for political ends.*
  This argument is not an argument for position C: it implicitly assumes a counterargument against position C and attacks this argument. The implicit counterargument is that doing nothing causes a further temperature rise, which is dangerous. This is made explicit (4F-a).



- *The motivation argument cuts both ways: climate change risks are also being minimised for political ends. The only way to progress this debate is to become familiar with the science.*
  In the graph, this is a counterargument against the proposition that the consequences of climate change are being exaggerated, but it does not attack by contradicting the argument. It is interpreted as stating that political arguments are not useful in this debate.

- Implicit in the whole debate is that, whether prioritising other challenges or not, acting on climate change costs money, time, and resources. The most important reason to argue for doing nothing is to save all this. To have a more complete view on the debate as a whole, this reason is made explicit in the form of an argument for position C (2H-a and 2H-b).

A visualisation of the complete structure is made in Rationale[5]. See appendix A for a text version of the structured debate.

3.3 Formalisations on the basis of $ASPIC^+$

Three formalisations will be discussed in this section: first the debate will be formalised using $ASPIC^+$, as discussed in section 2.2.2. This formalisation will then be reconstructed in assumption-based argumentation, using a method from Dung and Thang [8]. This reconstruction will in turn be reconstructed in classical-logic argumentation.

3.3.1 A formalisation in $ASPIC^+$

First, the debate is formalised in $ASPIC^+$. The elements of $AS = \langle \mathcal{L}, \mathcal{R}, n, ^- \rangle$ are specified below. The complete formalisation can be found in Appendix B.1.

$\mathcal{L}$ contains:

- each of the statements in the structured version;
- the categorised premises such as literature and expert opinions.

$\mathcal{R}$ is such that:

- $\mathcal{R}_s$ consists of the rules `r6A` and `r8C`;
- $\mathcal{R}_d$ consists of the rest of the rules.

Furthermore, $\mathcal{K}$ is such that:

- $\mathcal{K}_n$ consists of the categorised leaves that refer to literature, expert opinions, websites or common belief.
- $\mathcal{K}_p$ consists of the rest of the leaves.

Since (A), (B) and (C) are different positions in the debate, they are considered each others contradictories in $\mathcal{L}$.

[5] http://www.rationaleonline.com

Arguments The arguments in the argumentation framework are the following:

- All premises, named $\mathrm{P}_{1\mathrm{Aa}}$, $\mathrm{P}_{1\mathrm{Bb}}$, etc.
- Several categorised facts:

  CBelief Common belief.

  CD'09 The Copenhagen Diagnosis, a report written by a group of climate scientists in 2009.

  ExpField An expert opinion of professor Chris Field on BBC News.

  WebMiller A website by Dan Miller.

  IPCC'07 The findings of the Intergovernmental Panel on Climate Change, which reports on relevant climate research literature to member governments, as stated in their report from 2007.

  MIT Predictions from MIT.

- There are several arguments of the form of DMP as used in [13]: $\mathrm{DMP}_{3\mathrm{A}^1}$ to $\mathrm{DMP}_{3\mathrm{A}^{11}}$, $\mathrm{DMP}_{9\mathrm{I}}$, $\mathrm{DMP}_{10\mathrm{F}}$.
  '$\mathrm{DMP}_{9\mathrm{I}}$' is the name for the argument with support 9I-a and 9I-b.
- Arguments from good consequences: $\mathrm{GC}_{2\mathrm{A}}$, $\mathrm{GC}_{2\mathrm{B}}$, $\mathrm{GC}_{2\mathrm{C}}$, $\mathrm{GC}_{2\mathrm{AB}}$, $\mathrm{GC}_{2\mathrm{AC}}$, $\mathrm{GC}_{2\mathrm{BC}}$, $\mathrm{GC}_{2\mathrm{ABC}}$, $\mathrm{GC}_{2\mathrm{F}}$, $\mathrm{GC}_{2\mathrm{G}}$, $\mathrm{GC}_{2\mathrm{FG}}$, $\mathrm{GC}_{2\mathrm{H}}$.
  Here, '$\mathrm{GC}_{2\mathrm{AB}}$' is the name of the argument from a scheme from good consequences concerning 2A and 2B.
- Arguments from bad consequences: $\mathrm{BC}_{2\mathrm{D}}$, $\mathrm{BC}_{2\mathrm{E}}$, $\mathrm{BC}_{2\mathrm{DE}}$, $\mathrm{BC}_{2\mathrm{I}}$.
  Here, '$\mathrm{BC}_{2\mathrm{DE}}$' is the name of the argument from a scheme from bad consequences concerning 2D and 2E.
- Analogy: 5E
- 6A and 8C are strict inferences

Where not specified otherwise, conclusions of arguments are regarded to be defeasibly implied by their support.

Attacks A list of attacks can be found in appendix B.1.

3.3.2 Reconstructing the formalisation in ABA

In [8], Dung and Thang define a method to translate an $ASPIC^+$ model into an ABA model. The above $ASPIC^+$ formalisation will be translated using their definition. The components of the resulting $ABAF = \langle \mathcal{L},\mathcal{R},\mathcal{A},\mathcal{C} \rangle$ will be given in appendix B.2. Some examples to illustrate the formalisation are worked out below.

In [8], features of $ASPIC^+$ that are not present in ABA are translated, in order to simulate these properties in ABA.

- The possibility to rebut a defeasible argument is simulated by adding a negation-as-failure component $not\_\neg \mathtt{h}$ to the body of each defeasible rule, where $\mathtt{h}$ is the head of the rule[6]. $not\_\mathtt{x}$ is defined as the contrary of $\mathtt{x}$. As a result, for every rebut directed at conclusion $\mathtt{c}$, there is a corresponding attack directed at assumption $not\_\neg\mathtt{c}$.
- The possibility to undercut a defeasible argument is simulated by adding a component $Oj(\boldsymbol{r})$ to the body of each defeasible rule $\boldsymbol{r}$, where $Oj(\boldsymbol{r})$ means that $\boldsymbol{r}$ is applicable. $Oj(\boldsymbol{r})$ is defined as the contrary of $\neg Oj(\boldsymbol{r})$. As a result, for every undercut directed at rule $\boldsymbol{r}$, there is an attack directed at assumption $Oj(\boldsymbol{r})$.

Dung and Thang's method only applies to a specific version of $ASPIC^+$. First of all, while $ASPIC^+$ contains an asymmetric contrariness function to indicate conflict, the translation described in [8] is from $ASPIC^+$ instantiated with the more specific classical negation. Furthermore, the translation is based on a system in which axioms and premises are represented by strict and defeasible rules with empty antecedents, respectively. This causes premises to be translated as rules with $Oj(\boldsymbol{r})$ and $not\_\neg\mathtt{h}$ assumptions in the body. In a much simpler transition, ordinary premises are translated directly as assumptions, as argued in [11].

Below, a modification of the definition by Dung and Thang will be given. It is meant to translate $ASPIC^+$ models in which axioms and premises are elements in a knowledge base. Furthermore, it assumes that conflicts are not indicated by classical negation, but by a more general contrariness relation as in [11]. Instead of an $Oj(\boldsymbol{r})$ component, the existing $n(\boldsymbol{r})$ from $ASPIC^+$ is used as an assumption to indicate that a rule $\boldsymbol{r}$ is applicable. Ordinary premises are translated as assumptions.

Definition 3.1. An $ABAF = \langle \mathcal{L}', \mathcal{R}', \mathcal{A}, \mathcal{C} \rangle$ corresponding to an $AS = \langle \mathcal{L}, \mathcal{R}, n, ^- \rangle$ is such that:

$\mathcal{A} =$ $\{\mathtt{p} | \mathtt{p} \in \mathcal{K}_p\} \cup \{n(\boldsymbol{r}) | \boldsymbol{r} \in \mathcal{R}_d\} \cup \{not\_\mathtt{l} | \mathtt{l}$ is a contrary or contradictory of the head of some rule in $\mathcal{R}_d\}$
where $n(\boldsymbol{r})$ indicates that rule $\boldsymbol{r}$ is applicable and $not\_\mathtt{l}$ indicates that there is no acceptable argument with conclusion $\mathtt{l}$

$\mathcal{R}' =$ $\{\rightarrow \mathtt{p} | \mathtt{p} \in \mathcal{K}_n\} \cup \mathcal{R}_s \cup \{\mathrm{Tr}(\boldsymbol{r}) | \boldsymbol{r} \in \mathcal{R}_d\}$
where $\mathrm{Tr}(\boldsymbol{r})$ is of the form $n(\boldsymbol{r}), not\_\mathtt{l}_1, ..., not\_\mathtt{l}_m, \lambda_1, ..., \lambda_n \rightarrow \mathtt{h}$
if $\boldsymbol{r}$ is of the form $\lambda_1, ..., \lambda_n \Rightarrow \mathtt{h}$
and $\mathtt{l}_1,...,\mathtt{l}_m$ are all the contraries and contradictories of $\mathtt{h}$

[6]In definition 20 in [8], the authors state that for every defeasible rule with head $\mathtt{h}$, $not\_\mathtt{h}$ instead of $not\_\neg\mathtt{h}$ is included in the set of assumptions, and that $not\_\mathtt{h}$ states there is no evidence to the contrary of $\mathtt{h}$. Since this matches neither the intuition about the negation-as-failure assumption, nor example 9 below the definition, this is regarded as a typographical error.

$\mathcal{C}$: $\texttt{l} \in \mathcal{C}(\texttt{not\_l})$
$\overline{\texttt{x}} = \mathcal{C}(\texttt{x})$ if x is a statement in $\mathcal{L}$

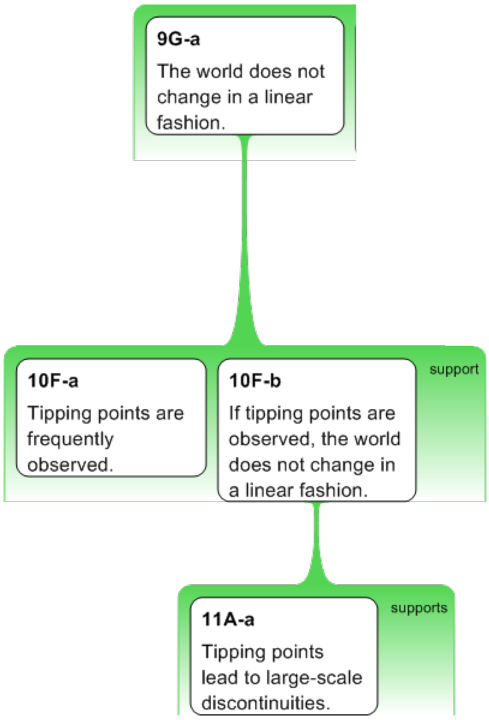


Figure 5: Argument 10F

Some examples are worked out to illustrate the model.

Example 3.1. This is a formal representation of argument 10F with conclusion 9G-a. See figure 5 for the Rationale visualisation of this part of the debate. $ABAF = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C} \rangle$ is as follows:

$\mathcal{A}$

11A-a	(since 11A-a $\in \mathcal{K}_p$)
10F-a	(since 10F-a $\in \mathcal{K}_p$)
r11A	(since *r11A* $\in \mathcal{R}_d$)
r10F	(since *r10F* $\in \mathcal{R}_d$)

$\mathcal{R}$

r11A: *r11A*, 11A-a $\rightarrow$ 10F-b	(since *r11A*: 11A-a, 11A-b $\Rightarrow$ 10F-b)
r10F: *r10F*, 10F-a, 10F-b $\rightarrow$ 9G-a	(since *r10F*: 10F-a, 10F-b $\Rightarrow$ 9G-a)

The arguments that can be constructed are the following. The left column contains the arguments from the ABA formalisation, on the right are the corresponding arguments from the $ASPIC^+$ formalisation.

ABA	$ASPIC^+$
A_{r11A}: $r11A \vdash r11A$	-
A_{r10F}: $r10F \vdash r10F$	-
P_{11Aa}: 11A-a $\vdash$ 11A-a	P_{11Aa}: 11A-a
P_{10Fa}: 10F-a $\vdash$ 10F-a	P_{10Fa}: 10F-a
11A: 11A-a, $r11A$ $\vdash$ 10F-b	11A: P_{11Aa}, 11A $\Rightarrow$ 9G-a
DMP_{10F}: 10F-a, 11A-a, $r11A$, $r10F$ $\vdash$ 9G-a	DMP_{10F}: P_{10Fa}, 11A $\Rightarrow$ 9G-a

There are no attacks in this example.

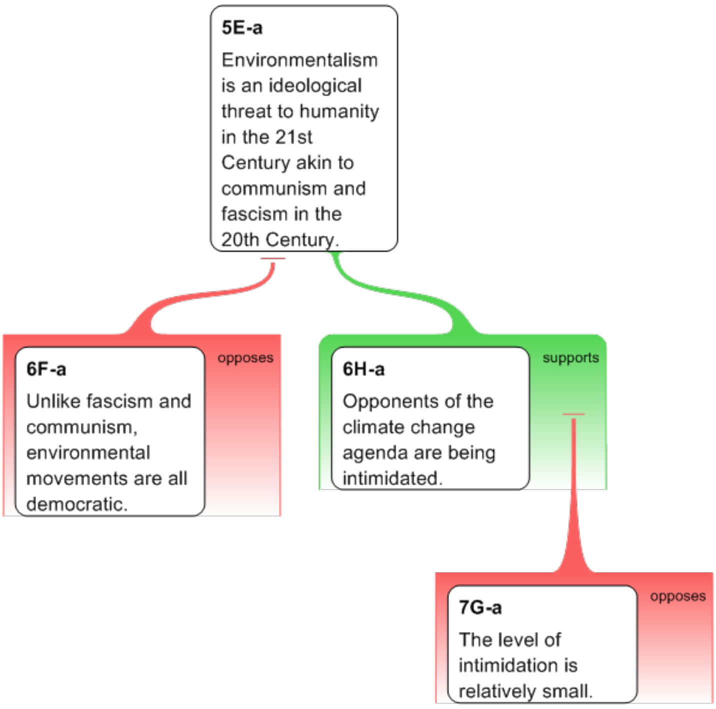


Figure 6: Argument 6H

Example 3.2. In this example, the formal argument for 5E-a is shown, as well as two of its attackers; one rebuttal (P_{6Fa}) and one undercut (P_{7Ga}). For the Rationale visualisation of this part, see figure 6. The $ASPIC^+$ and ABA formalisations are visualised in figures 7a and 7b.

$\mathcal{A}$

6H-a (since 6H-a $\in \mathcal{K}_p$ in de $ASPIC^+$ formalisation)

6F-a	(since 6F-a $\in \mathcal{K}_p$)
7G-a	(since 7G-a $\in \mathcal{K}_p$)
r6H	(since *r6H* $\in \mathcal{R}_d$)
*not*_6F-a	(since 5E-a $= -$6F-a and 5E-a is the head of a rule in $\mathcal{R}_d$)

$\mathcal{R}$

r6H: *r6H*, *not*_6F-a, *r6H*, 6H-a $\rightarrow$ 5E-a	(since *r6H*: 6H-a $\Rightarrow$ 5E-a and 6F-a $= -$5E-a and 7G-a $\in \overline{r6H}$)

$\mathcal{C}$

6F-a $\in \overline{not\_6F\text{-}a}$	(by definition)
7G-a $\in \overline{r6H}$	(directly from *ASPIC*$^+$)
5E-a $= -$6F-a	(directly from *ASPIC*$^+$)

The following arguments can be constructed. Again, the right column indicates to which *ASPIC*$^+$ arguments the ABA arguments correspond.

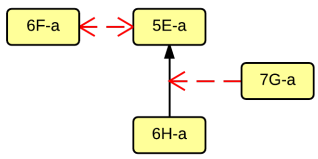
ABA	*ASPIC*$^+$
A_{r6H}: *r6H* $\vdash$ *r6H*	-
$A_{not\_6F\text{-}a}$: *not*_6F-a $\vdash$ *not*_6F-a	-
P_{6Ha}: 6H-a $\vdash$ 6H-a	P_{6Ha}: 6H-a
P_{6Fa}: 6F-a $\vdash$ 6F-a	P_{6Fa}: 6F-a
P_{7Ga}: 7G-a $\vdash$ 7G-a	P_{7Ga}: 7G-a
6H: 6H-a, *r6H*, *not*_6F-a $\vdash$ 5E-a	6H: $P_{6Ha} \Rightarrow$ 5E-a

In this example there are some attacks:

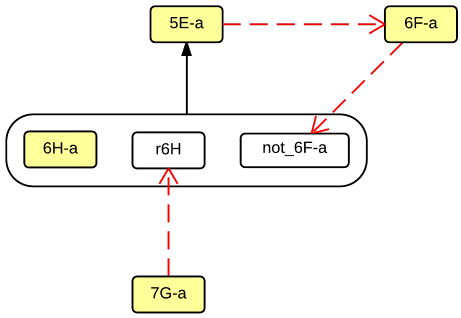
- P_{7Ga} attacks A_{r6H} and 6H on *r6H*.
  This corresponds to the undercut of the *ASPIC*$^+$ argument P_{7Ga} on 6H.
- P_{6Fa} attacks $A_{not\_6F\text{-}a}$ and 6H on *not*_6F-a.
  6H attacks P_{6Fa} on 6F-a.
  This corresponds to the formalisation in *ASPIC*$^+$, in which the arguments P_{6Fa} and 6H rebut each other on their conclusions (6F-a and 5E-a).

3.3.3 Reconstructing the formalisation in Classical-Logic Argumentation

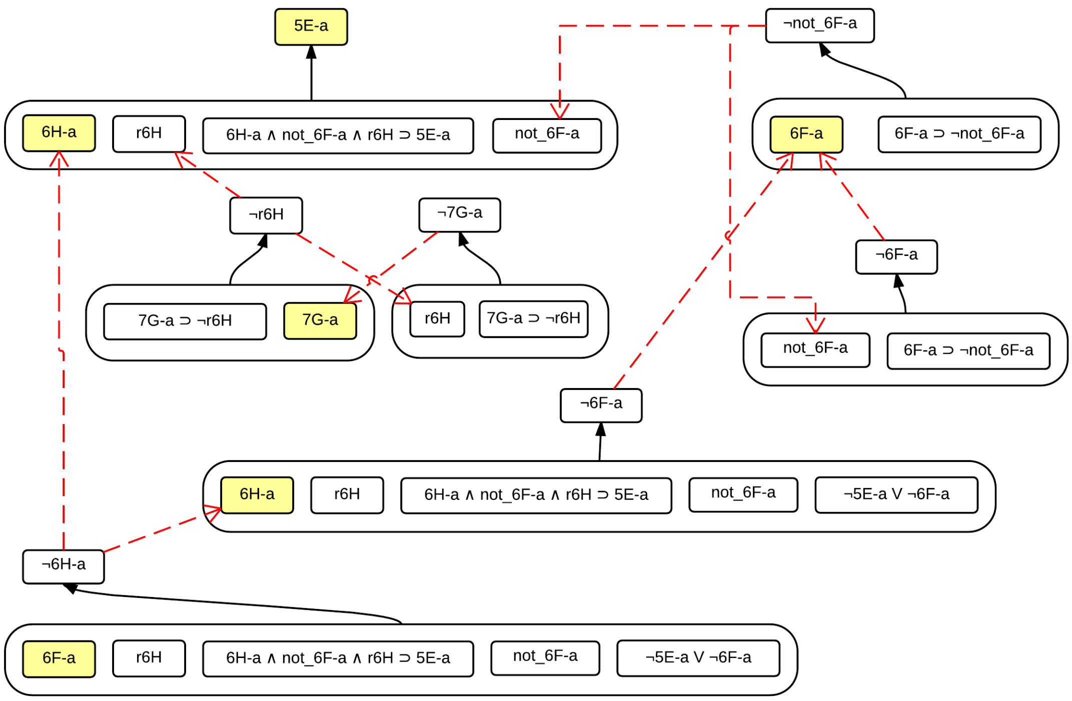
To reconstruct the ABA formalisation in an argumentation system based on classical logic, all assumptions, rules and the contrary function have to be interpreted as classical



(a) Formalisation in $ASPIC^+$



(b) Reconstruction in ABA



(c) Reconstruction in classical argumentation

Figure 7: Different formalisations of argument 6H and its attackers. Black arrows indicate support relations, red dashed lines indicate conflicts.

formulae. These formulae will form the knowledge base from which arguments can be constructed. Assumptions can be directly interpreted as elements of the knowledge base. A possible way to represent rules is to see them as material implications. Conflict will be indicated by classical negation.

One difference from the ABA formalisation is the behaviour of the negation-as-failure component $not\_\texttt{x}$. Recall that in the ABA formalisation, the 'meaning' of $not\_\texttt{x}$ was that there is no acceptable argument with x as its conclusion. In the context of classical logic, this cannot be defined that simple. Since $not\_\texttt{x}$ and x do not negate each other directly, it should be made explicit that they are not allowed together in the support of an argument. Examples 3.1 and 3.2 are reconstructed below as examples 3.3 and 3.4 respectively.

Example 3.3. This knowledge base consists of all assumptions and rules in example 3.1, where the rules are now material implications.

Δ

$\mathit{r11A}$	(since $\mathit{r11A} \in \mathcal{A}$)
$\mathit{r10F}$	(since $\mathit{r10F} \in \mathcal{A}$)
$\texttt{11A-a}$	(since $\texttt{11A-a} \in \mathcal{A}$)
$\texttt{10F-a}$	(since $\texttt{10F-a} \in \mathcal{A}$)
$\texttt{11A-a} \wedge \mathit{r11A} \supset \texttt{10F-b}$	(since $\mathit{r11A}, \texttt{11A-a} \rightarrow \texttt{10F-b} \in \mathcal{R}$)
$\texttt{10F-a} \wedge \texttt{10F-b} \wedge \mathit{r10A} \supset \texttt{9G-a}$	(since $\mathit{r10F}, \texttt{10F-a}, \texttt{10F-b} \rightarrow \texttt{9G-a} \in \mathcal{R}$)

Below are the relevant arguments to construct from the knowledge base. Compared to example 3.1, P1 is new. For the rest, the corresponding ABA argument is given in the right column.

Classical-logic argumentation	ABA
$\mathrm{A}_{\mathit{r11A}}$: $\mathit{r11A} \vdash \mathit{r11A}$	$\mathrm{A}_{\mathit{r11A}}$: $\mathit{r11A} \vdash \mathit{r11A}$
$\mathrm{A}_{\mathit{r10F}}$: $\mathit{r10F} \vdash \mathit{r10F}$	$\mathrm{A}_{\mathit{r10F}}$: $\mathit{r10F} \vdash \mathit{r10F}$
$\mathrm{P}_{11\mathrm{Aa}}$: $\texttt{11A-a} \vdash \texttt{11A-a}$	$\mathrm{P}_{11\mathrm{Aa}}$: $\texttt{11A-a} \vdash \texttt{11A-a}$
$\mathrm{P}_{10\mathrm{Fa}}$: $\texttt{10F-a} \vdash \texttt{10F-a}$	$\mathrm{P}_{10\mathrm{Fa}}$: $\texttt{10F-a} \vdash \texttt{10F-a}$
R11A: $\texttt{11A-a} \wedge \mathit{r11A} \supset \texttt{10F-a} \vdash \texttt{11A-a} \wedge \mathit{r11A} \supset \texttt{10F-a}$	-
R10F: $\texttt{10F-a} \wedge \texttt{10F-b} \wedge \mathit{r10F} \supset \texttt{9G-a} \vdash \texttt{10F-a} \wedge \texttt{10F-b} \wedge \mathit{r10F} \supset \texttt{9G-a}$	-
11A: $\texttt{11A-a}, \mathit{r11A} \vdash \texttt{10F-b}$	11A: $\texttt{11A-a}, \mathit{r11A} \vdash \texttt{10F-b}$
$\mathrm{DMP}_{10\mathrm{F}}$: $\texttt{10F-a}, \texttt{11A-a}, \mathit{r11A}, \mathit{r10F} \vdash \texttt{9G-b}$	$\mathrm{DMP}_{10\mathrm{F}}$: $\texttt{10F-a}, \texttt{11A-a}, \mathit{r11A}, \mathit{r10F} \vdash \texttt{9G-a}$

Example 3.4. This example is a conversion of example 3.2. The knowledge base consists of all assumptions and the rule from the ABA model. In addition, material implications are added for each element of the contrariness function. A visualisation of this example

can be found in figure 7c.

Δ

6H-a	(since 6H-a $\in \mathcal{A}$)
6F-a	(since 6F-a $\in \mathcal{A}$)
7G-a	(since 7G-a $\in \mathcal{A}$)
r6H	(since *r6H* $\in \mathcal{A}$)
*not*_6F-a	(since *not*_6F-a $\in \mathcal{A}$)
r6H $\wedge$ *not*_6F-a $\wedge$ 6H-a $\supset$ 5E-a	(since *r6H*, *not*_6F-a, 6H-a $\rightarrow$ 5E-a $\in \mathcal{R}$)
6F-a $\supset \neg$*not*_6F-a	(since 6F-a $\in \overline{not\_6F\text{-}a}$)
7G-a $\supset \neg$*r6H*	(since 7G-a $\in \overline{r6H}$)
$\neg$5E-a $\vee \neg$6F-a	(since 5E-a $= -$6F-a)

The following relevant arguments can be constructed. P1 is new, as are R1-R3. These lead to the new arguments A3, B2, C2, and D2, which are needed to create attacks.

Classical-logic argumentation	ABA
N1: *not*_6F-a $\vdash$ *not*_6F-a	$A_{not\_6F\text{-}a}$: *not*_6F-a $\vdash$ *not*_6F-a
O1: *r6H* $\vdash$ *r6H*	A_{r6H}: *r6H* $\vdash$ *r6H*
P1: *r6H* $\wedge$ *not*_6F-a $\wedge$ 6H-a $\supset$ 5E-a $\vdash$ *r6H* $\wedge$ *not*_6F-a $\wedge$ 6H-a $\supset$ 5E-a	-
R1: 6F-a $\supset \neg$*not*_6F-a $\vdash$ 6F-a $\supset \neg$*not*_6F-a	-
R2: 7G-a $\supset \neg$*r6H* $\vdash$ 7G-a $\supset \neg$*r6H*	-
R3: $\neg$5E-a $\vee \neg$6F-a $\vdash \neg$5E-a $\vee \neg$6F-a	-
A1: 6H-a $\vdash$ 6H-a	P_{6Ha}: 6H-a $\vdash$ 6H-a
A2: N1, O1, P1, A1 $\vdash$ 5E-a	6H: 6H-a, *r6H* $\vdash$ 5E-a
A3: O1, N1, P1, A1, R3 $\vdash \neg$6F-a	-
B1: 6F-a $\vdash$ 6F-a	P_{6Fa}: 6F-a $\vdash$ 6F-a
B2: B1, R1 $\vdash \neg$*not*_6F-a	-
C2: B1, R3 $\vdash \neg$5E-a	-
D1: 7G-a $\vdash$ 7G-a	P_{7Ga}: 7G-a $\vdash$ 7G-a
D2: D1, R2 $\vdash \neg$*r6H*	-

However, the following arguments (and more) are also possible to construct.

α1: N1, R1 $\vdash \neg$6F-a
α2: O1, R2 $\vdash \neg$7G-a

Attacks:

- A3 attacks B1, B2, and C2

- B2 attacks N1, A2, and α1
- D2 attacks O1, A2, and α2
- α1 attacks B1, B2, and C2
- α2 attacks D1 and D2

This is problematic. It is possible to construct an attack against `6F-a`, without making use of an argument for `5E-a`. This would not be possible in the *ASPIC*$^+$ formalisation nor in the ABA formalisation. Moreover, an attack against `7G-a` can be constructed, while in ABA and *ASPIC*$^+$, it is not attacked at all. Since the following set is a complete extension, `5E-a` is credulously justified under complete semantics: {`6H-a`, *r6H*, *not*_`6F-a`, *r6H* $\wedge$ *not*_`6F-a` $\wedge$ `6H-a` $\supset$ `5E-a`, `6F-a` $\supset$ $\neg$*not*_`6F-a`, `7G-a` $\supset$ $\neg$*r6H*, $\neg$`5E-a` $\vee$ $\neg$`6F-a`}. In the *ASPIC*$^+$ and ABA formalisations, `5E-a` is neither sceptically nor credulously justified.

To solve this, a preference ordering is applied to the knowledge base. Since there are no defeasible rules, there is no difference between last-link and weakest-link ordering. To compare arguments comes down to comparing two sets of premises, in this case.

The ordering in the knowledge base is to make sure that each *not*_`x` and *r* is less preferred than all other elements. This preference ordering is then lifted the Elitist way [11] to sets. A set is as strong as its weakest element, or as preferred as its least preferred element. This way, sets of premises including *not*_`x` or *r* will always be less preferred than sets without these elements. Translated to arguments: arguments supported by *not*_`x` or *r* will always be less preferred than arguments not supported by these elements.

Example 3.5 below illustrates how the use of such an ordering works.

Example 3.5. Consider the following ordinary premises:

Δ

$\mathtt{p} \wedge \neg ab \supset \mathtt{q}$
$\mathtt{p}$
$\mathtt{s} \supset ab$
$\mathtt{s}$
$\neg ab$

A preference ordering is applied in which the following holds:

$\leq$

$\neg ab < \mathtt{p} \wedge \neg ab \supset \mathtt{q}$
$\neg ab < \neg \mathtt{s} \vee ab$
$\neg ab < \mathtt{p}$
$\neg ab < \mathtt{s}$

The following arguments can be constructed:

P1 : $\texttt{p} \wedge \neg ab \supset \texttt{q}$
P2 : $\texttt{p}$
P3 : $\texttt{s} \supset ab$
P4 : $\texttt{s}$
P5 : $\neg ab$
A : $P1, P2, P5 \vdash \texttt{q}$
B : $P3, P4 \vdash ab$
β : $P3, P5 \vdash \neg\texttt{s}$

Here, B attacks P5, A and β, while β attacks P4 and B. Since $\neg ab < \neg\texttt{s} \vee ab$ and $\neg ab < \texttt{s}$, B is more preferred then β. Hence B's attack on β is successful and β's attacks are not. In all semantics, q is neither sceptically, nor credulously justified.

However, this solution only works for cases where s is a premise, and not the conclusion of some argument resting on an abnormality assumption. In that case, as argued in [11], the arguments would not defeat each other unless there was an ordering over the abnormality assumptions; an ordering that cannot be objectively determined.

To return to example 3.4: a preference ordering is defined on the knowledge base in which *not*_6F-a and ***r6H*** are less preferred than all other elements. The relevant resulting preferences are the following:

$\leq$

*not*_6F-a < 6F-a
r6H < 7G-a

Now $\alpha 1$ and $\alpha 2$ are less preferred then B2 and D2, respectively. Their attacks on arguments B1 and D1 are not successful. There is a unique extension containing 6H-a, 6F-a and 7G-a (among elements that were added to the knowledge base in the reconstruction), and not 5E-a. 5E-a is neither sceptically, nor credulously justified. Again, however, this only works because 6F-a and 7G-a are premises. In the full debate, where each non-premise argument is built up with explicit assumptions that rules are applicable, the support of each non-premise argument contains an ***r*** element and is therefore incapable of defeating arguments like $\alpha 1$ and $\alpha 2$.

3.3.4 Extensions

ASPIC$^+$

The preferred and stable extensions are defined by the union of the following sets:

- S_1 where $S_1 = \texttt{Sub}(\text{GC}_{2\text{A}}) \cup \texttt{Sub}(\text{GC}_{2\text{B}}) \cup \texttt{Sub}(\text{GC}_{2\text{C}}) \cup \texttt{Sub}(\text{GC}_{2\text{D}}) \cup \texttt{Sub}(\text{GC}_{2\text{E}}) \setminus$

$\{GC_{2A}, GC_{2B}, GC_{2C}, GC_{2D}, GC_{2E}\}$

- S_2 where $S_2 = \{P_{2Fa}, P_{2Fb}, P_{2Gb}, P_{2Ha}, P_{2Hb}, P_{2Ia}, P_{4Da}, P_{6Cb}, P_{6Fa}, P_{6Ga}, P_{6Ha}, P_{7Ga}, P_{8Ia}, 8I\}$
- S_3 where $S_3 \in \{\{8H, P_{8Ja}\}, \{8H, P_{7Fb}, 7F\}, \{P_{8Ja}, P_{6Ca}, 6C\}, \{P_{6Ca}, P_{7Fb}, 6C\}, \{P_{6Ca}, P_{7Fb}, 7F\}\}$
- S_4 where $S_4 \in \{\{P_{5Da}\}, \{P_{6Ea}\}\}$
- S_5 where S_5 is one of the following three sets, and $T \in \{\{BC_{2I}, 3G, 4E, P_{6Ba}\}, \{P_{5Ba}\}\}$:
  1. $\{GC_{2A}, GC_{2B}, GC_{2C}, GC_{2AB}, GC_{2AC}, GC_{2BC}, GC_{2ABC}\} \cup T$
  2. $\{GC_{2F}, BC_{2D}, BC_{2E}, BC_{2DE}\} \cup T$
  3. $\{GC_{2H}, P_{5Ba}, BC_{2D}, BC_{2E}, BC_{2DE}\}$

The grounded extension consists of $S_1 \cup S_2$ (see figures 9 and 8 for a visualisation). The following arguments are neither sceptically nor credulously justified under any semantics:

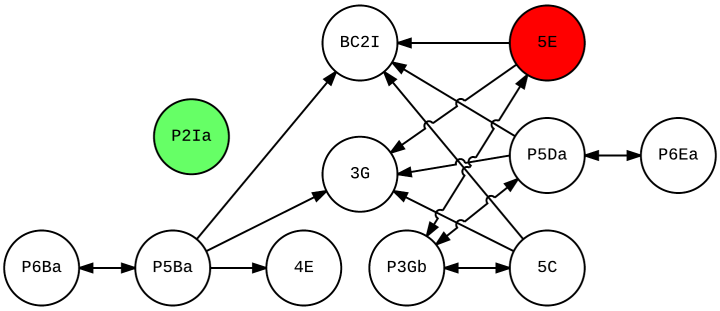
- GC_{2G}, GC_{2FG}, 3F, P_{3Fa}, 5C, 5D, 5E, 6H

To draw conclusions on the outcome of the debate, specifically: each of the three positions (A), (B) and (C) is credulously justified under preferred and stable semantics. In order to obtain a unique extension under any semantics, the following preferences should at least be in $\preceq$:

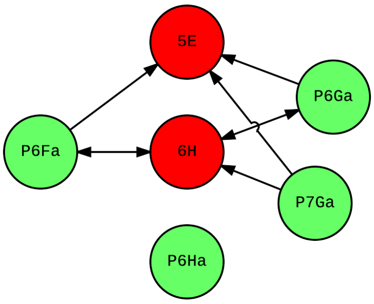
- $8H \succ P_{6Ca}$ and $P_{8Ja} \succ P_{7Fb}$
  or $8H \succ P_{6Ca}$ and $P_{8Ja} \prec P_{7Fb}$
  or $8H \prec P_{6Ca}$ and $P_{8Ja} \succ P_{7Fb}$
  or $8H \prec P_{6Ca}$ and $P_{8Ja} \prec P_{7Fb}$ and $6C \succ 7F$
  or $8H \prec P_{6Ca}$ and $P_{8Ja} \prec P_{7Fb}$ and $6C \prec 7F$
- $P_{5Da} \succ P_{6Ea}$ or $P_{5Da} \prec P_{6Ea}$

Furthermore, to obtain a unique extension, one of the three positions should be sceptically justified.

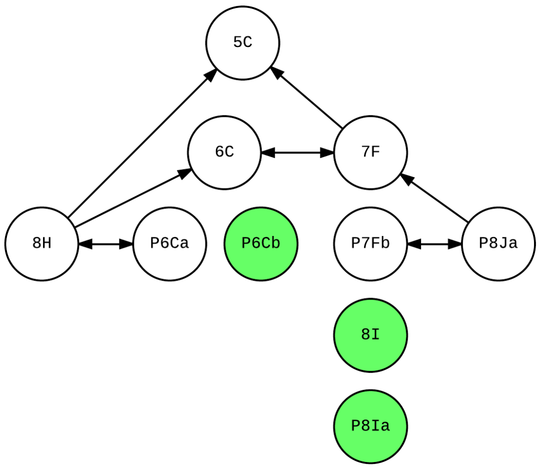
- Position (A) is sceptically justified if GC_{2A}, GC_{2B}, GC_{2C}, GC_{2AB}, GC_{2AC}, GC_{2BC}, and/or GC_{2ABC} is/are more preferred than BC_{2D}, BC_{2E}, BC_{2DE}, GC_{2F} and GC_{2H}.
  A second possibility is that GC_{2A}, GC_{2B}, GC_{2C}, GC_{2AB}, GC_{2AC}, GC_{2BC}, and/or GC_{2ABC} is/are more preferred than BC_{2D}, BC_{2E}, BC_{2DE} and GC_{2F} and that P_{6Ba} is more preferred than P_{5Ba}.
- Position (B) is sceptically justified if GC_{2F} is more preferred than GC_{2A}, GC_{2B}, GC_{2C}, GC_{2AB}, GC_{2AC}, GC_{2BC}, GC_{2ABC}, and GC_{2H}.
  A second possibility is that GC_{2F} is more preferred than GC_{2A}, GC_{2B}, GC_{2C}, GC_{2AB}, GC_{2AC}, GC_{2BC}, and GC_{2ABC} and that P_{6Ba} is more preferred than P_{5Ba}.



(a) Dung graph of sub-tree under argument BC_{2I}, without sub-trees under 5E and 5C.



(b) Dung graph of sub-tree under argument 5E



(c) Dung graph of sub-tree under argument 5C

Figure 8: Dung graph of sub-tree under argument BC_{2I}. Red nodes are arguments that are never justified under any semantics. Green nodes are arguments that are sceptically justified under any semantics.

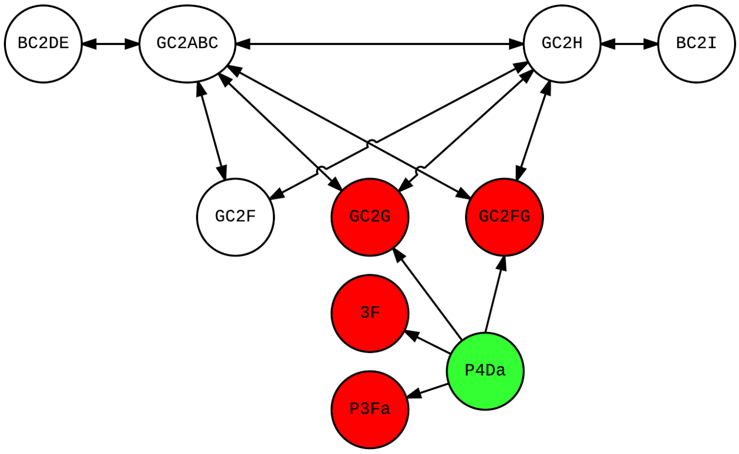


Figure 9: Dung graph of debate, without sub-tree under BC_{2I}. GC_{2ABC} is short for GC_{2A}, GC_{2B}, GC_{2C}, GC_{2AB}, GC_{2AC}, GC_{2BC}, and GC_{2ABC}, while BC_{2DE} is short for BC_{2D}, BC_{2E}, BC_{2DE}. The sub-trees under GC_{2A}, GC_{2B} and GC_{2C} are also omitted; only green nodes could be added for these statements.

- Position (C) is sceptically justified iff P_{5Ba} is more preferred than P_{6Ba} and GC_{2H} is more preferred than GC_{2A}, GC_{2B}, GC_{2C}, GC_{2AB}, GC_{2AC}, GC_{2BC}, GC_{2ABC}, and GC_{2F}.

ABA

The preferred and stable extensions are defined by the union of the following sets:

- S_1 where $S_1 = \mathtt{Sub}(GC_{2A}) \cup \mathtt{Sub}(GC_{2B}) \cup \mathtt{Sub}(GC_{2C}) \cup \mathtt{Sub}(GC_{2D}) \cup \mathtt{Sub}(GC_{2E}) \setminus \{GC_{2A}, GC_{2B}, GC_{2C}, GC_{2D}, GC_{2E}, A_{not\_GC_{2D}}, A_{not\_GC_{2E}}, A_{not\_GC_{2DE}}, A_{not\_GC_{2I}}\}$
- S_2 where $S_2 = \{P_{2Fa}, P_{2Fb}, P_{2Gb}, P_{2Ha}, P_{2Hb}, P_{2Ia}, P_{4Da}, P_{6Cb}, P_{6Fa}, P_{6Ga}, P_{6Ha}, P_{7Ga}, P_{8Ia}, 8I\}$
- S_3 where $S_3 \in \{\{8H, P_{8Ja}, A_{not\_\mathtt{6Ca}}, A_{not\_\mathtt{6Da}}, A_{not\_\mathtt{5Ca}}\}, \{8H, P_{7Fb}, 7F, A_{not\_\mathtt{6Ca}}, A_{not\_\mathtt{5Ca}}\}, \{P_{8Ja}, P_{6Ca}, 6C, A_{not\_\mathtt{6Da}}\}, \{P_{6Ca}, P_{7Fb}, 6C, A_{not\_\mathtt{6Da}}\}, \{P_{6Ca}, P_{7Fb}, 7F, A_{not\_\mathtt{5Ca}}\}\}$
- S_4 where $S_4 \in \{\{P_{5Da}\}, \{P_{6Ea}\}\}$
- S_5 where S_5 is one of the following three sets, and $T \in \{\{BC_{2I}, 3G, 4E, A_{r4E}, P_{6Ba}\}, \{P_{5Ba}\}\}$:
  1. $\{GC_{2A}, GC_{2B}, GC_{2C}, GC_{2AB}, GC_{2AC}, GC_{2BC}, GC_{2ABC}, A_{not\_\mathtt{1B-a}}, A_{not\_\mathtt{1C-a}}, A_{not\_GC_{2D}}, A_{not\_GC_{2E}}, A_{not\_GC_{2DE}}\} \cup T$
  2. $\{GC_{2F}, BC_{2D}, BC_{2E}, BC_{2DE}, A_{not\_\mathtt{1A-a}}, A_{not\_\mathtt{1C-a}}\} \cup T$
  3. $\{GC_{2H}, P_{5Ba}, BC_{2D}, BC_{2E}, BC_{2DE}, A_{not\_\mathtt{1A-a}}, A_{not\_\mathtt{1B-a}}, A_{not\_GC_{2I}}\}$

- S_6 where $S_6 = R \setminus \{\mathrm{A}_{\mathit{r4E}}, \mathrm{A}_{\mathit{r6H}}\}$ and R is the set of all arguments that express rule applicability

The grounded extension consists of $S_1 \cup S_2 \cup S_6$. The following arguments are neither sceptically nor credulously justified in any semantics:

- $\{\mathrm{GC}_{2\mathrm{G}}, \mathrm{GC}_{2\mathrm{FG}}, 3\mathrm{F}, \mathrm{P}_{3\mathrm{Fa}}, 5\mathrm{C}, 5\mathrm{D}, 5\mathrm{E}, 6\mathrm{H}, \mathrm{A}_{not\_\mathtt{3G-b}}, \mathrm{A}_{not\_\mathtt{6F-a}}, \mathrm{A}_{not\_\mathtt{6G-a}}, \mathrm{A}_{\mathit{r6H}}\}$

As well as in the *ASPIC*$^+$ formalisation, each of the three positions (A), (B) and (C) is credulously justified under preferred and stable semantics. Furthermore, the number of preferred and stable extensions is the same as with the *ASPIC*$^+$ formalisation. For each preferred and stable extension in the *ASPIC*$^+$ formalisation, there is a preferred and stable extension in the ABA formalisation, where the latter contains corresponding arguments for each argument in the former. The same holds for the grounded extension: each argument in the grounded extension of the *ASPIC*$^+$ formalisation has a corresponding argument in the grounded extension of the ABA formalisation.

Classical argumentation
Many additional arguments are in the classical-logic reconstruction, and the knowledge-base is much larger, as well. We have already shown that the extensions are different for this model. Comparing the grounded extension to that of the *ASPIC*$^+$ and ABA models, a number of arguments that were in the latter two are not in this one. This is caused by the additional attacks from arguments from contraposition. The number of preferred and stable extensions is much higher than for the previous two models. These extensions do not correspond as those of the *ASPIC*$^+$ and ABA models did.

3.4 A formalisation based on Assumption-Based Argumentation

Below, a formalisation of the debate is given following the decision making techniques of [9], based on assumption-based argumentation. First the *extended decision framework* is constructed, based on input from the debate. This *edf* will form the input for an ABA formalisation, which is constructed next.

3.4.1 Extended decision framework

The first step in the process of creating this framework is to determine decisions, attributes and goals in the debate. Possible decisions are equal to the possible actions and therefore the three points of view in the debate: act immediately (`imm`), prioritise (`later`), do not act at all (`never`).

The consequences of these actions form the set of attributes. The table DA represents knowledge about which attribute is a consequence of which action. Possible attributes in this case are: there is a 1 meter sea-level rise (*`sea1`*), there is a 2 meter sea-level rise (*sea2*), there is a 2 degree temperature rise (*tmp2*), there is a 6 degree temperature rise (*tmp6*), there are positive side-effects like decreased economic crisis (*sideEff*), Arctic ice melts faster (*fastMelt*), resources are deployed more effectively (*moreEff*), less money is spend (*cheaper*).

Of course, the actions are proposed with one or more goals in mind. In this case, the following goals can be recognised in the debate: saving the environment (`saveEnv`), improve quality of life on the planet (`lifeQual`), respond proportionately to the problem (`propResp`).

Together with a preference order on (sets of) goals, this leads to the following *edf*:

$\mathbf{D} = \{\texttt{imm}, \texttt{later}, \texttt{never}\}$

$\mathbf{A} = \{sea1, sea2, tmp2, tmp6, sideEff, fastMelt, moreEff, cheaper\}$

$\mathbf{G} = \{\texttt{saveEnv}, \texttt{lifeQual}, \texttt{propResp}\}$

$\mathbf{P} = \{\texttt{saveEnv} > \texttt{lifeQual} > \texttt{propResp}\}$

DA =

	sea1	*sea2*	*tmp2*	*tmp6*	*sideEff*	*fastMelt*	*moreEff*	*cheaper*
`imm`	1	0	0	0	1	0	0	0
`later`	0	1	1	0	0	1	1	1
`never`	0	1	0	1	0	1	0	1

GA =

	sea1	*sea2*	*tmp2*	*tmp6*	*sideEff*	*fastMelt*	*moreEff*	*cheaper*
`saveEnv`	1	0	1	0	0	0	1	0
`lifeQual`	1	0	1	0	0	0	1	0
`propResp`	0	0	0	0	0	0	1	1

3.4.2 ABA framework

An $ABAF = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C} \rangle$ is constructed. The preference relation is translated into a binary relation, expressing for each pair of (sets of) goals which one is preferred. Next, the decision-attribute and goal-attribute pairs that are true according to the tables DA and GA, are represented by rules of the form '`decision`.*attribute* ←' and '`goal`.*attribute* ←'. Because there are rules for every decision-goal pair of the kind '`decision.goal` ← `decision`.*attribute*, `goal`.*attribute*', the relation between a decision and a goal becomes explicit if there is an attribute of that decision that satisfies this goal: '`decision.goal`' becomes true[7].

By default, each decision is in $\mathcal{A}$, which means each decision is assumed to be a good decision. At the same time there are rules for each decision which lead to a contradiction of these assumptions. The information from the tables and the preference order can lead to the fact that some of these rules are not applied and therefore some decision remains true.

[7] This works the same for sets of goals.

$$
\begin{array}{l}
\mathcal{R} = \{ \\
\quad \begin{array}{lll}
P\{\texttt{saveEnv}\}\{\texttt{lifeQual}\} \leftarrow & P\{\texttt{saveEnv}\}\{\texttt{propResp}\} \leftarrow & P\{\texttt{lifeQual}\}\{\texttt{propResp}\} \leftarrow \\
\texttt{imm}.\textsf{\textit{sea1}} \leftarrow & \texttt{imm}.\textsf{\textit{sideEff}} \leftarrow & \texttt{later}.\textsf{\textit{sea2}} \leftarrow \\
\texttt{later}.\textsf{\textit{tmp2}} \leftarrow & \texttt{later}.\textsf{\textit{fastMelt}} \leftarrow & \texttt{later}.\textsf{\textit{moreEff}} \leftarrow \\
\texttt{later}.\textsf{\textit{cheaper}} \leftarrow & \texttt{never}.\textsf{\textit{sea2}} \leftarrow & \texttt{never}.\textsf{\textit{tmp6}} \leftarrow \\
\texttt{never}.\textsf{\textit{fastMelt}} \leftarrow & \texttt{never}.\textsf{\textit{cheaper}} \leftarrow & \\
\{\texttt{saveEnv}\}.\textsf{\textit{sea1}} \leftarrow & \{\texttt{saveEnv}\}.\textsf{\textit{sea2}} \leftarrow & \{\texttt{saveEnv}\}.\textsf{\textit{tmp2}} \leftarrow
\end{array} \\
\quad \ldots \\
\quad \texttt{imm}\{\texttt{saveEnv}\} \leftarrow \texttt{imm}.\textsf{\textit{sea1}}, \{\texttt{saveEnv}\}.\textsf{\textit{sea1}} \\
\quad Nd^{\texttt{imm}} \leftarrow \texttt{later}\{\texttt{saveEnv}\}, N\texttt{imm}\{\texttt{saveEnv}\}, NX^{\texttt{immLater}}_{\{\texttt{saveEnv}\}} \\
\quad Nd^{\texttt{imm}} \leftarrow \texttt{never}\{\texttt{saveEnv}\}, N\texttt{imm}\{\texttt{saveEnv}\}, NX^{\texttt{immNever}}_{\{\texttt{saveEnv}\}} \\
\quad Nd^{\texttt{later}} \leftarrow \texttt{imm}\{\texttt{saveEnv}\}, N\texttt{later}\{\texttt{saveEnv}\}, NX^{\texttt{laterImm}}_{\{\texttt{saveEnv}\}} \\
\quad Nd^{\texttt{later}} \leftarrow \texttt{never}\{\texttt{saveEnv}\}, N\texttt{later}\{\texttt{saveEnv}\}, NX^{\texttt{laterNever}}_{\{\texttt{saveEnv}\}} \\
\quad Nd^{\texttt{never}} \leftarrow \texttt{imm}\{\texttt{saveEnv}\}, N\texttt{never}\{\texttt{saveEnv}\}, NX^{\texttt{neverImm}}_{\{\texttt{saveEnv}\}} \\
\quad Nd^{\texttt{never}} \leftarrow \texttt{later}\{\texttt{saveEnv}\}, N\texttt{never}\{\texttt{saveEnv}\}, NX^{\texttt{neverLater}}_{\{\texttt{saveEnv}\}} \\
\quad Nd^{\texttt{imm}} \leftarrow \texttt{later}\{\texttt{lifeQual}\}, N\texttt{imm}\{\texttt{lifeQual}\}, NX^{\texttt{immLater}}_{\{\texttt{lifeQual}\}} \\
\quad \ldots \\
\quad X^{\texttt{immLater}}_{\{\texttt{saveEnv}\}} \leftarrow \texttt{later}\{\texttt{lifeQual}\}, N\texttt{imm}\{\texttt{lifeQual}\}, P\{\texttt{saveEnv}\}\{\texttt{lifeQual}\} \\
\quad X^{\texttt{immNever}}_{\{\texttt{saveEnv}\}} \leftarrow \texttt{never}\{\texttt{lifeQual}\}, N\texttt{imm}\{\texttt{lifeQual}\}, P\{\texttt{saveEnv}\}\{\texttt{lifeQual}\} \\
\quad \ldots \\
\quad \}
\end{array}
$$

$$
\begin{array}{l}
\mathcal{A} = \{ \\
\quad \begin{array}{lll}
\texttt{imm} & NX^{\texttt{immLater}}_{\{\texttt{saveEnv}\}} & N\texttt{imm}\{\texttt{saveEnv}\} \\
\texttt{later} & NX^{\texttt{immNever}}_{\{\texttt{saveEnv}\}} & N\texttt{later}\{\texttt{saveEnv}\} \\
\texttt{never} & NX^{\texttt{laterImm}}_{\{\texttt{saveEnv}\}} & N\texttt{never}\{\texttt{saveEnv}\} \\
 & NX^{\texttt{laterNever}}_{\{\texttt{saveEnv}\}} & N\texttt{imm}\{\texttt{lifeQual}\} \\
 & NX^{\texttt{neverImm}}_{\{\texttt{saveEnv}\}} & \ldots \\
 & NX^{\texttt{neverLater}}_{\{\texttt{saveEnv}\}} & N\texttt{imm}\{\texttt{saveEnv}, \texttt{lifeQual}\} \\
 & NX^{\texttt{immLater}}_{\{\texttt{lifeQual}\}} & \ldots \\
 & \ldots & \\
 & NX^{\texttt{immLater}}_{\{\texttt{saveEnv},\texttt{lifeQual}\}} & \\
 & \ldots &
\end{array} \\
\quad \}
\end{array}
$$

$$\mathcal{C} = \{$$

$$
\begin{array}{ll}
\mathcal{C}(\texttt{imm}) = \{N\texttt{imm}\} & \mathcal{C}(NX^{\texttt{immLater}}_{\{\texttt{saveEnv}\}}) = \{X^{\texttt{immLater}}_{\{\texttt{saveEnv}\}}\} \\
\mathcal{C}(\texttt{later}) = \{N\texttt{later}\} & \mathcal{C}(NX^{\texttt{immNever}}_{\{\texttt{saveEnv}\}}) = \{X^{\texttt{immNever}}_{\{\texttt{saveEnv}\}}\} \\
\mathcal{C}(\texttt{never}) = \{N\texttt{never}\} & \mathcal{C}(NX^{\texttt{laterImm}}_{\{\texttt{saveEnv}\}}) = \{X^{\texttt{laterImm}}_{\{\texttt{saveEnv}\}}\} \\
 & \mathcal{C}(NX^{\texttt{laterNever}}_{\{\texttt{saveEnv}\}}) = \{X^{\texttt{laterNever}}_{\{\texttt{saveEnv}\}}\} \\
 & \mathcal{C}(NX^{\texttt{neverImm}}_{\{\texttt{saveEnv}\}}) = \{X^{\texttt{neverImm}}_{\{\texttt{saveEnv}\}}\} \\
 & \mathcal{C}(NX^{\texttt{neverLater}}_{\{\texttt{saveEnv}\}}) = \{X^{\texttt{neverLater}}_{\{\texttt{saveEnv}\}}\} \\
 & \mathcal{C}(NX^{\texttt{immLater}}_{\{\texttt{lifeQual}\}}) = \{X^{\texttt{immLater}}_{\{\texttt{lifeQual}\}}\} \\
 & \ldots \\
 & \mathcal{C}(NX^{\texttt{immLater}}_{\{\texttt{saveEnv},\texttt{lifeQual}\}}) = \{X^{\texttt{immLater}}_{\{\texttt{saveEnv},\texttt{lifeQual}\}}\} \\
 & \ldots
\end{array}
$$

}

This way of formalising is not suited to merely describe or model an existing discussion. It is designed to model a situation in which all information is known, and where one or more agents with clear goals can argue about and take decisions based on that information. Included and fixed information in the model is:

- the set of all the possible decisions to make;
- the set of all the attributes, which are the connections between decisions and goals;
- the set of all the goals of the agents;
- which decisions have which attributes;
- which attributes satisfy which goals.

Although this way of modelling can be relevant in programming autonomous software agents, it is not the same as analysing (human) discussions. First of all, in analysing a discussion, it is often impossible to recognise goals of the participants (if there is information about participants at all).

Furthermore, both the preference ordering and the relation between decisions, attributes and goals are often part of the discussion. Many of the values that are put in the tables above, would not actually be consensus between debaters. Perhaps this framework could be extended in such a way that preferences are not assumed beforehand (but possible topics of debate) and goals are not necessarily known.

4 Discussion

In this section, the different techniques as applied in the previous section will be compared. The focus is on the research questions stated in section 1.

In section 3.4, we argued that while it is possible to formalise human argumentation about action proposals, the method from [9] is not the best method to use. As an alternative, we discussed the ABA reconstruction of the $ASPIC^+$ formalisation. In the following sections, the last method will be compared to $ASPIC^+$ and the reconstruction in classical argumentation, both discussed in section 3.3. In the following, [9] will not be discussed.

In some of the following discussions, it is relevant information which number of each type of attack is present in the debate. See the tables below:

Type	$ASPIC^+$	**ABA**	**classical argumentation**
rebuttal	69	0	0
undercut	2	0	0
premise-attack	22	93	93+
Total	**93**	**93**	**93+**

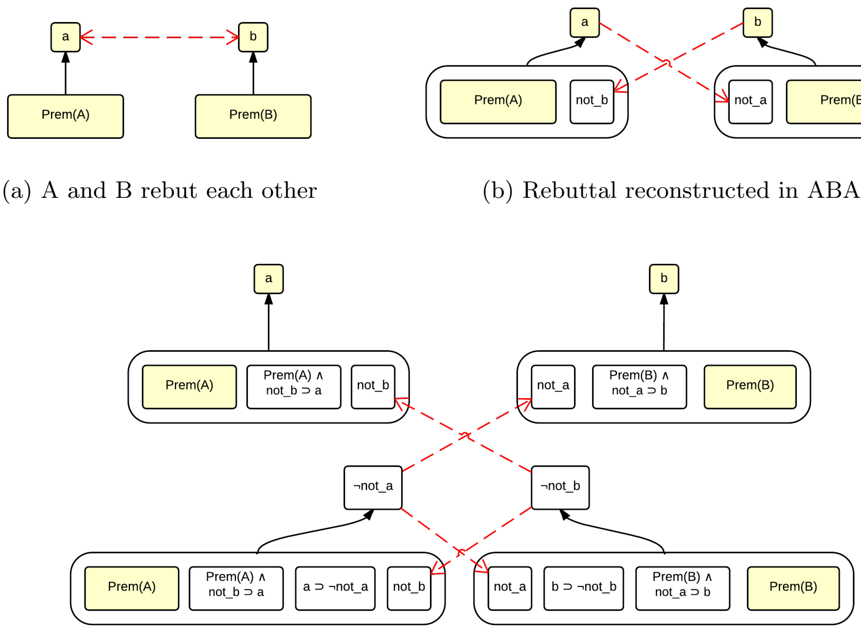
There are ninety-three attacks in total in the $ASPIC^+$ and ABA formalisations. In the classical-logic formalisation, a large number of extra attacks is generated in a manner we discussed in section 3.3.3. According to [15], a low number of attacks was found in other case-studies [16, 12, 14, 17, 5] that would be categorised as premise-attacks in $ASPIC^+$. Here, a substantial number of attacks is of that type. A cause for this could be that the debate on DebateGraph is not worked out entirely in all details. For example, sources of information are missing, which would turn present premises into supported statements. Attacks on these statements would then be rebuttals in $ASPIC^+$.

4.1 Defeasibility of conclusions

In this section, the question will be discussed whether a realistic model should allow rebuttals. First we will discuss which way of formalising such an attack is most natural. Then we will look at the number of attacks of this kind in the investigated debate, to find out how relevant our question is in this particular case.

4.1.1 Methods of formalising

Throughout this section, two arbitrary arguments A and B (with conflicting conclusions `a` and `b`, respectively) will continuously serve as an example.



(a) A and B rebut each other

(b) Rebuttal reconstructed in ABA

(c) Rebuttal reconstructed in classical argumentation

Figure 10: Different formalisations of an attack on a conclusion.

We have seen different possible ways to formalise arguments with conflicting conclusions, which will be discussed briefly below. We will then discuss how natural these different ways of formalising are, by comparing them to the structured debate as in section 3.

We saw that one way to formalise such a conflict is as a rebuttal [11]. If conclusion `b` of B is in conflict with conclusion `a` of A, then this is formalised as A and B rebutting each other.

If a conflict is formalised as a rebuttal, the point of conflict in the formalisation is between the conclusions of the conflicting arguments. Hence, if we formalise the conflict between A and B as a rebuttal, the point of conflict is the same as in its natural form.

In [8], an alternative way to formalise a conflict like this, by translating the $ASPIC^+$ formalisation into ABA, is described. Since ABA does not allow attacks directed at conclusions, it is not possible to formalise the conflict between A and B as a conflict between conclusions of arguments in ABA. B can only attack A if its conclusion disagrees with a premise of A. To simulate the possibility of an attack on `a`, we saw that a negation-as-failure assumption is added to A for each contrary or contradictory of `a`. In this case, since `b` is a contrary or contradictory of `a`, the assumption that there is no argument for `b` is added to A. B attacks A on this assumption, since B is such an argument for b.

In this formalisation, the conflict between b and a is reduced to a conflict between the conclusion of B and a premise of A. Hence the point of conflict differs from that of the arguments in their natural form.

A third possibility, which we discussed in section 3.3.3, is representing arguments by logical proofs in a classical argumentation system; attacks are arguments claiming the negation of some premise of another argument. Like in ABA, attacks on conclusions are not allowed in such a system. Negation-as-failure is simulated to create a similar way of formalising a conflict such as the one between A and B. The support of B, in combination with an additional rule, entails the negation of the 'negation-as-failure' premise in the support of A.

In the last two options, where rebuttals are reduced to premise-attacks, the attack is directed at the assumption that there is no acceptable argument for a conflicting claim. Let us review this in the context of the criteria from section 1.

Suppose attacks on conclusions are more naturally formalised when they are reduced to premise-attacks. This can only be the case if a model is more natural when it contains the negation-as-failure elements that were added in the ABA formalisation and simulated in the classical-logic formalisation. Recall the criteria we mentioned in section 1. We consider a formalisation more natural if (1) all implicit elements from a original debate are made explicit, and (2) if all elements that are made explicit are elements which were in the original debate. In other words, the negation-as-failure elements that were added would actually have to be implicit elements which were made explicit, or our initial supposition (that attacks on conclusions are more naturally formalised when they are reduced to premise-attacks) cannot be true. Thus, the question arises whether the negation-as-failure elements are indeed natural elements: do humans indeed, when they make a claim, implicitly base that claim on the assumption that the contrary of that claim is false? Constructing arguments of that kind feels unnatural: basing a claim c on the assumption that c is not false is a tautologous reasoning step which does not fit human reasoning. Supporting c *is not false* with the assumption that c is true seems just as valid as the other way around. In the same line of thought, we can add a premise expressing that c *is not not not false*. If c *is not false* is truly an assumption made by humans, then it is unclear why c *is not not not false* could not be. Hence, it is sensible to view negation-as-failure elements as necessary elements for certain models, but not as elements that correspond directly to human argumentation. Since we classify the elements that were made explicit as unnatural elements, the formalisations in which attacks on conclusions are reduced to premise-attacks contain more unnatural elements than the formalisation including rebuttals. Furthermore, there are no implicit elements from the original debate missing in any of the formalisations. Therefore, considering our criteria, the formalisation including rebuttals is more natural. Returning to our initial question: attacks on conclusions are less naturally formalised when they are reduced to premise-attacks.

To summarise, if arguments with conflicting conclusions, like A and B, are formalised in a model in which attacks on conclusions are allowed, the point of conflict is the same in the formalisation as that in the natural form. This leads to the conclusion that this kind of model is formalised in a more natural way.

4.1.2 Relevance

We will discuss the relevance of the distinction made above by considering the number of relevant attacks from the investigated debate. If we count symmetric rebuttals as two attacks, there are sixty-two rebuttals between arguments of good and bad consequences. In the epistemic part of the debate, there are twenty more, but thirteen of these are directed at premises. We will not take these into account, since these are not more naturally formalised when rebuttals are allowed. Therefore, the relevant attacks in this section add up to sixty-nine, out of ninety-three attacks in total. These sixty-nine attacks (74%) are formalised in a more natural way, by a formalisation technique that allows attacks on conclusions.

4.2 Defeasibility of rules

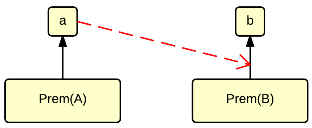
In this section, the topic of discussion will be whether attacks on rules should be reduced to premise-attacks to obtain a more realistic modelling technique. We will also discuss how relevant the difference is, by considering how many attacks would be more naturally formalised when choosing a specific representation of these attacks.

4.2.1 Methods of formalising

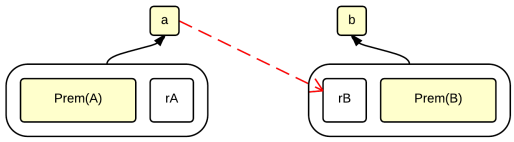
Throughout this section, A and B refer to two arguments, where conclusion `a` of A is in conflict with an inference of B, an inference for which rule *rB* is applied. In section 3 we have discussed several options to formalise this kind of conflict. These will be briefly mentioned below. Then we will discuss which way is more natural.

First, we discussed [11], in which the conflict between B and A would be formalised as an undercut. Informally, `a` conflicts with *rB*, and since A is an argument for `a`, A attacks B on *rB*. If the conflict between A and B is formalised as an undercut, then the point of conflict is on the inference of B.

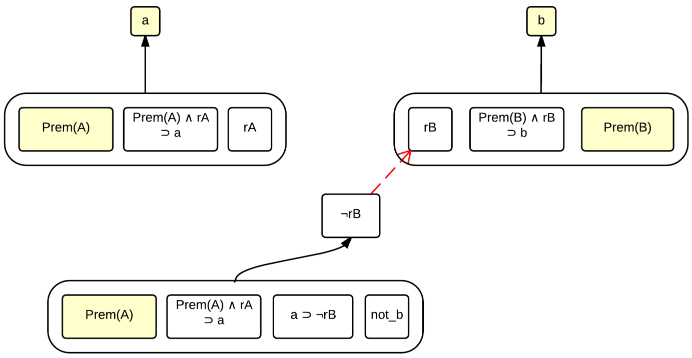
An alternative way to formalise A's attack on B is in a transition from $ASPIC^+$ to ABA. Attacks on inferences are not allowed in ABA, hence undercuts are formalised another way. An extra assumption is added to arguments in general: the assumption that a rule is applicable. Argument A no longer attacks *rB* itself, but the assumption that *rB* is applicable. In this formalisation, the attack of argument A is directed at an assumption.



(a) A undercuts B on *rB*



(b) Undercut reconstructed in ABA



(c) Undercut reconstructed in classical argumentation

Figure 11: Different formalisations of an attack on an inference.

Let us suppose that an attack on an inference is more naturally formalised as an attack on the assumptions that an inference rule is applicable, than by an undercut as in $ASPIC^+$, and let us view the consequences of this supposition. First of all, a formalisation in which attacks on inferences are reduced to premise-attacks contains more explicit elements in each argument, namely the assumption that the inference rule is applicable. Recall that we view a formalisation as less natural if it contains more explicit elements that are not a part of the original debate. Hence, our initial supposition (that an attack on an inference is more naturally formalised when it is reduced to a premise-attack) can only hold if the assumptions that were added in this formalisation are actually implicit elements that were made explicit. Is this the case? Do humans indeed base a claim `c` on the assumption that `c` can be inferred from the support of their claim? One difference between the two representations of the defeasibility of a rule, is the following: adding a premise to an argument that the applied rule is applicable, suggests that it would in theory be possible to construct an argument for this new premise. There is no syntactical difference between this statement and other statements, and all statements can be supported by others. When the applicability of a defeasible rule is formalised on a meta-level, as in $ASPIC^+$ supporting it is not possible. One could argue that having that option is more natural, that arguments supporting the applicability of a rule is a good feature. On the other hand, arguments supporting the applicability of a rule are not found in this debate, so this case does not provide the necessary information to classify one option as more natural than the other option. Hence it is not certain whether assumptions of this type are natural elements of arguments. Since we cannot, on the basis of our criteria, state whether those assumptions are natural elements of human argumentation, we are also unable to say whether a formalisation including those elements contains more unnatural elements or not. Furthermore, we are unable to say whether formalisations excluding those elements are excluding elements that should be included. In other words, it is not certain which of those two categories of formalisations is a more natural representation of a natural debate.

To conclude, there is no convincing argument for one of the two representations of attacks on inferences.

4.2.2 Relevance

One could argue that the above discussion is irrelevant, because this kind of conflict does not occur in real life debates. For this reason, we will list attacks on inferences from the investigated debate:

- 5B on the inference of 4E;
- 7G on the inference of 6H.

Two attacks out of ninety-three would possibly be more realistically, naturally formalised using a specific technique. Since this is a very low number of attacks, the difference between a formalisation of this debate including attacks on inferences and one excluding attacks on inferences is very small. The relevance of the above discussion is, in this case, minimal.

4.3 Use of Preferences

In this section, the use of preferences will be discussed as a possible feature for a realistic model. First, the different approaches to formalising preferences in a debate are discussed. Then, the relevance of formalising preferences is discussed by looking at the number of preferences in this debate.

4.3.1 Approaches to formalising preferences

$ASPIC^+$ has an explicit notion of preferences between arguments, hence the preference between A and B is formalised in a straightforward way. Arguments between statements can be lifted to preferences between arguments. A preference between statements `p` and `q` is important only if there are two conflicting arguments P and Q, such that `p` is a premise of P and `q` of Q. If P and Q are rebutting or undermining each other, the preference between `p` and `q` can determine which of the two attacks is successful.

In ABA, preferences are not allowed. In [18], the author prefers to encode preferences using other methods, and then make a mapping into ABA. An example is the method from [9]. This is done to keep the underlying framework as simple as possible. However, despite the fact that an ABA framework without preferences may be simpler than one with preferences, this does not imply that the process of formalising argumentation is simpler this way. If we need an additional framework and a mapping each time we want to formalise a debate in ABA, there is no benefit of having a simpler framework. Furthermore, the complexity of the underlying framework is a criterion which we did not value as highly as our criteria mentioned in section 1. There is no benefit of excluding preferences in the framework. To translate preferences in an ABA model, additional unnatural elements are necessary. Hence, considering our criteria, an ABA formalisation excluding preferences is less natural.

In [10], there is no claim about the use of preferences. In [4], however, the following is said about preferences over elements in the knowledge base:

> [W]e do not assume any meta-level information about formulae. In particular, we do not assume some preference ordering or "certainty ordering" over formulae. [...] Such orderings can be useful to resolve conflicts by, for example, selecting formulae from a more reliable source. However, this, in a sense, pushes the problem of dealing with conflicting information to one of finding and using orderings over formulae, and as such raises further questions such as: Where does the knowledge about reliability of the sources come from? How can it be assessed? How can it be validated? [...]

> This is not to say priorities [...] are not useful. Indeed it is important to use them in some situations when they are available, but we believe that to understand the elements of argumentation, [...] we need to have a comprehensive framework for argumentation that works without recourse to priorities over formulae.

It seems that Besnard and Hunter do not want to exclude preferences, but neither make them required input. They seem to argue that preferences might be valuable as optional input, although they do not describe nor refer to a way of handling preferences. However, in our classical-logic formalisation, we saw that without preferences it was impossible to create a correct formalisation. Hence, when formalising a debate in classical argumentation, including preferences provides a more natural result.

4.3.2 Relevance

In this particular debate, only one preference between consequences of actions has been found. There are no preferences in the epistemic part. The *precautionary principle* states a preference of 2I over 2H. Therefore, based on this case, we view the method of formalising preferences as irrelevant.

4.4 Summary

In section 4.1, we discussed whether attacks on conclusions should be reduced to premise-attacks when formalising argumentation about action proposals. We saw that the number of additional explicit elements is much higher in the ABA and classical-logic reconstructions, and argued that these additional elements were not natural elements from arguments in the original debate. We concluded that a model allowing rebuttals provides more natural formalisations. Next, we concluded that the number of attacks that were less naturally formalised in the ABA and classical-logic formalisations, was high enough to classify this discussion relevant.

In section 4.2, we discussed whether attacks on inference rules should be reduced to premise-attacks in an argumentation system. We first saw that an attack on an inference rule is not necessarily more naturally formalised by an undercut. When such attacks are reduced to premise-attacks, a premise has to be added to each defeasible argument to make attacks possible against the application of the defeasible rule. This means there is a higher number of explicit elements, but it remains debatable whether these elements are already implicitly present in a natural debate. We also concluded that this discussion was not relevant, since the number of attacks on inference rules is very low.

The joint conclusion of sections 4.1 and 4.2 is the following: since both rebuttals and undercuts can only be directed at defeasible inferences, the distinction between strict and defeasible rules is valuable. Undercuts do not make a difference based on this case, since there very few found, but the number of rebuttals is high enough to conclude that this debate is more naturally formalised when a distinction between strict and defeasible rules is made.

In section 4.3, we discussed whether preferences would be a valuable feature for argumentation systems. We saw that the ABA and classical-logic reconstructions lacked alternatives to formalise preferences. However, the very low number of preferences found in the debate suggests it is not relevant to include preferences.

5 Conclusion

In section 2, we have discussed several techniques with which we can formalise argumentation in general and argumentation about action proposals. In particular, we discussed the $ASPIC^+$ framework, assumption-based argumentation and classical argumentation. In section 3, we used these techniques to formalise our interpretation of a debate from DebateGraph about how to respond to climate change. A formalisation in ABA was made, as well as a $ASPIC^+$ formalisation which was reconstructed in ABA and classical argumentation. In section 4, we discussed the different formalisations of the debate, focussing on the questions from section 1. We concluded the following:

1. **Is the most realistic model created by a modelling technique in which *attacks on conclusions* are reduced to premise-attacks, or by one which does not, and why?**
   This category of attack is most naturally and realistically formalised when it is not reduced to a premise-attack. Since sixty-nine out of ninety-three attacks are directed at conclusions, a more realistic model is created by a modelling technique in which attacks on conclusions are not reduced to premise-attacks.
2. **Is the most realistic model created by a modelling technique in which *attacks on inferences* are reduced to premise-attacks, or by one which does not, and why?**
   This category of attack is most naturally and realistically formalised when it is not reduced to a premise-attack. However, only two out of ninety-three attacks are directed at inferences. Hence if we compare models that do and models that do not reduce attacks of this kind to premise-attacks, there is a minimal difference in how naturally these models formalise argumentation.
3. **Is the most realistic model created by a modelling technique which includes *preferences*, or by one which does not, and why?**
   We did not find an alternative to formalise preferences. However, the number of preferences in our case is very low. Hence if we compare models that do and models that do not include preferences, there is a minimal difference in how naturally these models formalise argumentation.

Case-studies like this one help us by providing insight in how theory works in practice. Here we provided insight in which features a model for argumentation should include, by comparing different theories applied to an actual debate. This case was specifically focussed on argumentation about action proposals. However, we saw that a large part of the debate of our choice is an epistemic part, in the sense that choices for actions are supported by epistemic reasoning. A consequence is that a natural formalisation includes an epistemic part, as well. It therefore inherits some behaviour from epistemic reasoning. In that context, results concerning defeasible reasoning are not surprising. In earlier publications it was already argued that defeasible reasoning is of value in epistemic argumentation. These results suggest that defeasible reasoning can be valuable in practical reasoning just as well. Furthermore, the same problems play a role in reconstructing an $ASPIC^+$ formalisation into classical-logic argumentation.

More case-studies can be valuable in multiple ways. First, more case-studies can provide insight in which categories of attack are most used in human argumentation. For example, only a few attacks on inferences were found in this case, but this could very well be incidental. Second, since our case-study only addresses some of the argumentation frameworks in the literature, other studies could be addressing other frameworks. One possibility would be to modify the method from [9]. This method would be more suitable to formalise human argumentation if the possibility was left open to include a formalisation of the epistemic part. Thus expanding the model with the possibility to add regular ABA assumptions to support the arguments generated from the *edf* would contribute to the ability to create natural formalisations. For example, goals could be chosen such that they correspond to consequences of actions. The question why something should or should not be a goal would then be the topic of regular, epistemic argumentation.

References

[1] K. Atkinson, T. Bench-Capon, and P. McBurney. Computational representation of practical argument. *Synthese*, 152(2):157–206, 2006.

[2] T. Bench-Capon, H. Prakken, and W. Visser. Argument schemes for two-phase democratic deliberation. In K. D. Ashley and T. M. van Engers, editors, *Proceedings of the 13th International Conference on Artificial Intelligence and Law*, ICAIL '11, pages 21–30, New York, NY, USA, 2011. ACM.

[3] P. Besnard, A. Garcia, A. Hunter, S. Modgil, H. Prakken, G. Simari, and F. Toni. Introduction to structured argumentation. *Argument & Computation*, 5(1):1–4, 2014.

[4] P. Besnard and A. Hunter. Argumentation based on classical logic. In I. Rahwan and G. R. Simari, editors, *Argumentation in Artificial Intelligence*, pages 133–152. Springer, 2009.

[5] F. Bex. Evidence for a good story. a hybrid theory of arguments, stories and criminal evidence. Doctoral dissertation Faculty of Law, University of Groningen, 2009.

[6] P. M. Dung. On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321 – 357, 1995.

[7] P. M. Dung, R. A. Kowalski, and F. Toni. Assumption-based argumentation. In *Argumentation in Artificial Intelligence*, pages 199–218. Springer, 2009.

[8] P. M. Dung and P. M. Thang. Closure and consistency in logic-associated argumentation. *Journal of Artificial Intelligence Research*, 49:79–109, 2014.

[9] X. Fan, R. Craven, R. Singer, F. Toni, and M. Williams. Assumption-based argumentation for decision-making with preferences: A medical case study. In J. a. Leite, T. Son, P. Torroni, L. van der Torre, and S. Woltran, editors, *Computational Logic in Multi-Agent Systems*, volume 8143 of *Lecture Notes in Computer Science*, pages 374–390. Springer, Berlin Heidelberg, 2013.

[10] N. Gorogiannis and A. Hunter. Instantiating abstract argumentation with classical logic arguments: Postulates and properties. *Artificial Intelligence*, 175(9-10):1479 – 1497, 2011.

[11] S. Modgil and H. Prakken. The ASPIC+ framework for structured argumentation: a tutorial. *Argument & Computation*, 5(1):31–62, 2014.

[12] H. Prakken. Formalising ordinary legal disputes: a case study. *Artificial Intelligence and Law*, 16:333–359, 2008.

[13] H. Prakken. Formalising a legal opinion on a legislative proposal in the ASPIC+ framework. In B. Schäfer, editor, *Legal Knowledge and Information Systems. JURIX*

2012: The Twenty-fifth Annual Conference, pages 119–128, Amsterdam, 2012. IOS Press.

[14] H. Prakken. Reconstructing Popov v. Hayashi in a framework for argumentation with structured arguments and dungean semantics. *Artificial Intelligence and Law*, 20:57–82, 2012.

[15] H. Prakken. Some reflections on two current trends in formal argumentation. In A. Artikis, R. Craven, N. Cicekli, B. Sadighi, and K. Stathis, editors, *Logic Programs, Norms and Action. Essays in Honour of Marek J. Sergot on the Occasion of his 60th Birthday*, pages 249–272. Springer, Berlin/Heidelberg, 2012.

[16] H. Prakken and S. Renooij. Reconstructing causal reasoning about evidence: a case study. In *Legal Knowledge and Information Systems. JURIX 2001: The Fourteenth Annual Conference*, pages 131–142, Amsterdam, 2001. IOS Press.

[17] J. Sombekke, T. van Engers, and H. Prakken. Argumentation structures in legal dossiers. In *Proceedings of the Eleventh International Conference on Artificial Intelligence and Law*, pages 277–281, New York, 2007. ACM Press.

[18] F. Toni. A tutorial on assumption-based argumentation. *Argument & Computation*, 5(1):89–117, 2014.

A A Text Version of the Structured Climate Debate

The debate is represented using the following notation.

[Position (X)] This is a position in the debate.
+[n1] Together with n2, this is an argument for X.
| <[m] This is a sub-argument of n1.
&[n2] Together with n1, this is an argument for X.
| A[l] ***This is an attacker for n2.***
| −[k] ***This argument and n2 attack each other.***
A[o1] ***Together with o2, this is an attacker for X.***
&[o2] ***Together with o1, this is an attacker for X.***

[Position (A)] We should act on climate change immediately.

+[2A-a] Acting immediately implies minimising the negative consequences.

&[2A-b] Minimising the negative consequences is good.

| <[3A-a] There are negative consequences.

| | <[4A-a] The expected climate change impact is much worse than people think.

| | | <[5A-a] What most people think is based on something which gives a too optimistic view on the expected climate change impact.

| | | | <[6A-a] What most people think is based on IPCC reports.

| | | | &[6A-b] The IPCC reports give a too optimistic view on the expected climate impact.

| | | | | <[7A-a] The IPCC report was a conservative, best-case analysis.

| | | | | | <[8A-a] Data in IPCC reports is generally stale by time of publication, because the IPCC requires information to be submitted years early.

| | | | | | <[8B-a] IPCC consensus process tends to eliminate more dramatic findings.

| | | | | | | <[expField] Professor Chris Field on BBC News.

| | | | | | <[8C-a] IPCC report tends to use linear models.

| | | | | | | <[9B-a] IPCC models do not consider ice sheet collapse.

| | | | | | | | <[IPCC'07] IPCC report 2007.

| | | | | | | <[9C-a] IPCC models do not include the amazon die-off feedback loop.

| | | | | | | | <[IPCC'07] IPCC report 2007.

| | | | | | | <[9D-a] IPCC models do not include the impact of methane clathrates collapse.

| | | | | | | | <[IPCC'07] IPCC report 2007.

| | | | | | | <[9E-a] IPCC models do not include the ocean CO_2 absorption feedback loop.

| | | | | | | | <[IPCC'07] IPCC report 2007.

| | | | | | | <[9F-a] IPCC models do not include the permafrost collapse feedback loop.

| | | | | | | | <[IPCC'07] IPCC report 2007.

| | | | | | &[8C-b] A report on climate change which uses tends to use linear models is a conservative, best-case analysis.

| | | | | | | <[9G-a] The world does not change in a linear fashion

| | | | | | | | <[10F-a] Tipping points are frequently observed

| | | | | | | | &[10F-b] If tipping points are frequently observed, then the world does not change in a linear fashion.

| | | | | | | | | <[11A-a] Tipping points lead to large-scale discontinuities.

| | | | | | | &[9G-b] Linear models give a more positive view than realistic models.

| | | | | | <[8D-a] IPCC predicts a global average temperature increase of 2°-5°C by 2100.

| | | | | | | <[IPCC'07] IPCC report 2007.

| | | | | | &[8D-b] An increase of 2°-5°C by 2100 is a conservative, best-case analysis.

| | | | | | | <[9I-a] The increase will be 6°C plus by 2100.

| | | | | | | | <[MIT] MIT probabilistic forecast for 21st century climate based on uncertainties in emissions (without policy) and climate parameters.

| | | | | | | &[9I-b] If the increase will be 6°C plus by 2100, then an increase of 2°-5°C is a conservative, best case analysis.

| | | | | <[7B-a] The actual worldwide CO_2 emissions have been higher than the worst case assumptions of the IPCC.

| | | | | | <[webMiller] Dan Miller's website: http://climateplace.org/

| | | | | <[7C-a] Arctic ice has been melting faster than IPCC expected worst case.

| | | | | | <[webMiller] Dan Miller's website: http://climateplace.org/

| | | | | <[7D-a] Global sea-level is likely to rise at least twice as much as in IPCC projections.

| | | | | | <[CD'09] The Copenhagen diagnosis (2009)

| &[3A-b] If there are negative consequences, then minimising the negative consequences is good.

+[2B-a] Acting immediately implies decarbonising the global economy.

&[2B-b] Decarbonising the global economy is good.

| <[3B-a] Decarbonisation of global economy is synergistic with other tasks. Decarbonisation also addresses the threats of Peak Oil, energy insecurity and economic recession, since it creates demand and supplies employment.

| &[3B-b] Addressing these additional threats is good.

+[2C-a] Acting immediately implies solving multiple problems at once.

| <[3C-a] If nothing is done, all human and social problems will be exacerbated by ACW.

&[2C-b] Solving multiple problems at once is good.

A[2D-a] Acting immediately may imply performing irrelevant actions.

| <[3D-a] Technological advances may make any current action irrelevant.

| | <[4B-a] Technological advances in the short and medium term may confound our current concerns about climate change.

| | <[4C-a] Technological advances in the short and medium term may enable the adoption of cheaper and more effective counter measures than are possible at present.

&[2D-b] Performing irrelevant actions is bad.

A[2E-a] *Acting immediately implies distracting governments from urgent needs where progress can be made.*

| <[3E-a] Grand narrative of climate change is a costly/hypocritical distraction. Governments' espoused enthusiasm (where it exists) for climate treaties is often not borne out in action - either by failing to meet the targets or via loopholes that render the targets meaningless. As such, the grand narrative is a costly distraction from urgent needs where progress can be made.

&[2E-b] *Distracting governments from urgent needs where progress can be made is bad.*

[Position (B)] We should prioritise our challenges before acting.

+[2F-a] Prioritise other challenges before acting implies responding proportionately to costs, benefits and needs.

&[2F-b] Responding proportionately to costs, benefits and needs is good.

+[2G-a] Prioritise other challenges before acting implies using resources earmarked for combatting climate change in a more effective way.

| <[3F-a] Resources earmarked for combatting climate change are more effectively and reliably directed at securing tangible improvements in the immediate quality of life on the planet.

| | A[4D-a] The response to climate change increases quality of life for all. It addresses issues of Peak Oil and energy insecurity which are going to make life very difficult in the next 1-2 decades, unless plentiful carbon energy is available.

&[2G-b] Using resources earmarked for combatting climate change in a more effective way is good.

[Position (C)] We should not act on climate change.

+[2H-a] Not acting on climate change implies saving money, time, and resources.

&[2H-b] Saving money, time, and resources is good.

A[2I-a] *Not acting on climate change implies continuing the current influence on climate change.*

&[2I-b] *Continuing the current influence is bad.*

| A[3G-a] *Continuing the current influence on climate causes a rise in temperature.*

| | <[4E-a] Basic textbook physics shows that the CO2 man has introduced into the atmosphere commits the planet to a 1°C rise in temperature. Positive feedback mechanisms will raise the temperature by further 2°C or more.

| | | A[5B-a] *Policy makers are placing absolute confidence in the scientific models and evidence. This glosses over the uncertainties inherent in the models and associated predictions.*

| | | | A[6B-a] *The entire climate change debate amongst scientists is expressed in terms of probabilities. Nothing is claimed to be certain.*

| &[3G-b] *A further temperature rise is bad.*

| | A[4F-a] *The risks of climate change are being exaggerated and dramatised for political ends.*

| | | <[5C-a] AGW is a government tax scam.

| | | | <[6C-a] Governments look for ways to make people pay.

| | | | &[6C-b] AGW is a good reason to raise taxes.

| | | | | A[7E-a] *Governments do not look for ways to tax people.*

| | | | | | <[com] Common belief

| | | | A[6D-a] *AGW can not be a government tax scam.*

| | | | | <[7F-a] An AGW related tax scam would not be beneficial for governments.

| | | | | | <[8I-a] The necessary cessation of industrial activity required in order to counter AGW would require the global economy to shrink dramatically.

| | | | | &[7F-b] Something that is not beneficial for governments cannot be a government scam.

| | | | | | A[8J-a] *Governments could be greedily pursuing the short-term monetary and political gains from anti-global-warming technology, regardless of the long-term cost their economies.*

| | | <[5D-a] Climate change is being used as lever to achieve global government.

| | | | A[6E-a] *Climate change as a lever to achieve global government is an unsubstantiated conspiracy theory.*

| | | <[5E-a] Environmentalism is an ideological threat to humanity in the 21st Century akin to communism and fascism in the 20th Century.

| | | | A[6F-a] *Unlike fascism and communism, environmental movements are all democratic.*

| | | | <[6H-a] Opponents of the climate change agenda are being intimidated.

| | | | | A[7G-a] *The level of intimidation is relatively small.*

| | | A[5F-a] *The motivation argument cuts both ways: climate change risks are also being minimised for political ends. The only way to progress this debate is to become familiar with the science.*

[Meta-argument] Precautionary principle: acting against the risks of climate change will bring significant spin-off benefits even if the risk prove to have been exaggerated; not acting will be catastrophic if the risks are borne out.

B $\boldsymbol{ASPIC^+}$ Formalisation and Reconstructions

B.1 Formalisation in $\boldsymbol{ASPIC^+}$

$\mathcal{L}$

1A-a	2G-a	3G-b	6C-a	8A-a	9I-a
1B-a	2G-b	4A-a	6C-b	8B-a	9I-b
1C-a	2H-a	4B-a	6D-a	8C-a	10F-a
2A-a	2H-b	4C-a	6E-a	8C-b	10F-a
2A-b	2I-a	4D-a	6F-a	8D-a	11A-a
2B-a	2I-b	4E-a	6G-a	8D-b	cBelief
2B-b	3A-a	4F-a	6H-a	8I-a	cd'09
2C-a	3A-b	5A-a	7A-a	8J-a	expField
2C-b	3B-a	5B-a	7B-a	9B-a	ipcc'07
2D-a	3B-b	5C-a	7C-a	9C-a	mit
2D-b	3C-a	5D-a	7D-a	9D-a	webMiller
2E-a	3D-a	5E-a	7E-a	9E-a	
2E-b	3E-a	6A-a	7F-a	9F-a	
2F-a	3F-a	6A-b	7F-b	9G-a	
2F-b	3G-a	6B-a	7G-a	9G-b	

$\mathcal{R}_s$

r6A: 6A-a, 6A-b → 5A-a
r8C: 8C-a, 8C-b → 7A-a

$\mathcal{R}_d$

r2A: 2A-a, 2A-b ⇒ 1A-a
r2B: 2B-a, 2B-b ⇒ 1A-a
r2C: 2C-a, 2C-b ⇒ 1A-a
r2AB: 2A-a, 2A-b, 2B-a, 2B-b ⇒ 1A-a
r2AC: 2A-a, 2A-b, 2C-a, 2C-b ⇒ 1A-a
r2BC: 2B-a, 2B-b, 2C-a, 2C-b ⇒ 1A-a
r2ABC: 2A-a, 2A-b, 2B-a, 2B-b, 2C-a, 2C-b ⇒ 1A-a
r2D: 2D-a, 2D-b ⇒ 2D
r2E: 2E-a, 2E-b ⇒ 2E
r2DE: 2D-a, 2D-b, 2E-a, 2E-b ⇒ 2DE
r2F: 2F-a, 2F-b ⇒ 1B-a
r2G: 2G-a, 2G-b ⇒ 1B-a
r2FG: 2F-a, 2F-b, 2G-a, 2G-b ⇒ 1B-a
r2H: 2H-a, 2H-b ⇒ 1C-a
r2I: 2I-a, 2I-b ⇒ 2I
r3A: 3A-a, 3A-b ⇒ 2A-b
r3B: 3B-a, 3B-b ⇒ 2B-b
r3C: 3C-a ⇒ 2C-a
r3D: 3D-a ⇒ 2D-a
r3E: 3E-a ⇒ 2E-a
r3F: 3F-a ⇒ 2G-a
r3G: 3G-a, 3G-b ⇒ 2I-b
r4A: 4A-a ⇒ 3A-a
r4B: 4B-a ⇒ 3D-a
r4C: 4C-a ⇒ 3D-a

r4E: 4E-a ⇒ 3G-a
r5A: 5A-a ⇒ 4A-a
r5C: 5C-a ⇒ 4F-a
r5D: 5D-a ⇒ 4F-a
r5E: 5E-a ⇒ 4F-a
r6C: 6C-a, 6C-b ⇒ 5C-a
r6H: 6H-a ⇒ 5E-a
r7A: 7A-a ⇒ 6A-b
r7B: 7B-a ⇒ 6A-b
r7C: 7C-a ⇒ 6A-b
r7D: 7D-a ⇒ 6A-b
r7F: 7F-a, 7F-b ⇒ 6D-a
r8A: 8A-a ⇒ 7A-a
r8B: 8B-a ⇒ 7A-a
r8D: 8D-a, 8D-b ⇒ 7A-a
r8E: webMiller ⇒ 7B-a
r8F: webMiller ⇒ 7C-a
r8G: cd'09 ⇒ 7D-a
r8H: cBelief ⇒ 7E-a
r8I: 8I-a ⇒ 7F-a
r9A: expField ⇒ 8B-a
r9B: 9B-a ⇒ 8C-a
r9C: 9C-a ⇒ 8C-a
r9D: 9D-a ⇒ 8C-a
r9E: 9E-a ⇒ 8C-a
r9F: 9F-a ⇒ 8C-a
r9G: 9G-a, 9G-b ⇒ 8C-b
r9H: ipcc'07 ⇒ 8D-a
r9I: 9I-a ⇒ 8D-b
r10A: ipcc'07 ⇒ 9B-a
r10B: ipcc'07 ⇒ 9C-a
r10C: ipcc'07 ⇒ 9D-a
r10D: ipcc'07 ⇒ 9E-a
r10E: ipcc'07 ⇒ 9F-a
r10F: 10F-a, 10F-b ⇒ 9G-a
r10G: mit ⇒ 9I-a
r11A: 11A-a ⇒ 10F-b

$\mathcal{K}_n$

cBelief	cd'09	expField	ipcc'07	mit	webMiller

$\mathcal{K}_p$

2A-a	2G-b	3C-a	4E-a	6E-a	8I-a
2B-a	2H-a	3E-a	5B-a	6F-a	8J-a
2C-b	2H-b	3F-a	5D-a	6G-a	9G-a
2D-b	2I-a	3G-b	6A-a	6H-a	9I-b
2E-b	3A-b	4B-a	6B-a	7F-b	10F-a
2F-a	3B-a	4C-a	6C-a	7G-a	11A-a
2F-b	3B-b	4D-a	6C-b	8A-a	

–

1A-a = −1B-a
1A-a = −1C-a
1B-a = −1C-a
2D = −1A-a
2E = −1A-a
2DE = −1A-a
2I = −1C-a
4D-a ∈ $\overline{\text{3F-a}}$
4F-a = −3G-b
5B-a ∈ $\overline{r4E}$
6B-a = −5B-a
6D-a = −5C-a

6E-a = –5D-a
6F-a ∈ $\overline{\texttt{5E-a}}$
6G-a = –5E-a
7E-a = –6C-a
7G-a ∈ $\overline{\textit{r6H}}$
8J-a = –7F-b

Arguments

P_{2Aa}: 2A-a
P_{2Ba}: 2B-a
P_{2Cb}: 2C-b
P_{2Db}: 2D-b
P_{2Eb}: 2E-b
P_{2Fa}: 2F-a
P_{2Fb}: 2F-b
P_{2Gb}: 2G-b
P_{2Ha}: 2H-a
P_{2Hb}: 2H-b
P_{2Ia}: 2I-a
P_{3Ab}: 3A-b
P_{3Ba}: 3B-a
P_{3Bb}: 3B-b
P_{3Ca}: 3C-a
P_{3Ea}: 3E-a
P_{3Fa}: 3F-a
P_{3Gb}: 3G-b
P_{4Ba}: 4B-a
P_{4Ca}: 4C-a
P_{4Da}: 4D-a
P_{4Ea}: 4E-a
P_{5Ba}: 5B-a
P_{5Da}: 5D-a
P_{6Aa}: 6A-a
P_{6Ba}: 6B-a
P_{6Ca}: 6C-a
P_{6Cb}: 6C-b
P_{6Ea}: 6E-a
P_{6Fa}: 6F-a
P_{6Ga}: 6G-a
P_{6Ha}: 6H-a
P_{7Fb}: 7F-b
P_{7Ga}: 7G-a
P_{8Aa}: 8A-a
P_{8Ia}: 8I-a
P_{8Ja}: 8J-a
P_{9Ga}: 9G-a
P_{9Ib}: 9I-b
P_{10Fa}: 10F-a
P_{11Aa}: 11A-a
CBelief: cBelief
CD'09: cd'09
ExpField: expField
WebMiller: webMiller
IPCC'07: ipcc'07
MIT: mit
11A: $P_{11Aa} \Rightarrow$ 10F-b
10G: MIT $\Rightarrow$ 9I-a
DMP_{10F}: P_{10Fa}, 11A $\Rightarrow$ 9G-a
10E: IPCC'07 $\Rightarrow$ 9F-a
10D: IPCC'07 $\Rightarrow$ 9E-a
10C: IPCC'07 $\Rightarrow$ 9D-a
10B: IPCC'07 $\Rightarrow$ 9C-a
10A: IPCC'07 $\Rightarrow$ 9B-a
DMP_{9I}: 10G, $P_{9Ib} \Rightarrow$ 8D-b
9H: IPCC'07 $\Rightarrow$ 8D-a
9G: 10F, $P_{9Gb} \Rightarrow$ 8C-b
9F: 10E $\Rightarrow$ 8C-a
9E: 10D $\Rightarrow$ 8C-a
9D: 10C $\Rightarrow$ 8C-a
9C: 10B $\Rightarrow$ 8C-a
9B: 10A $\Rightarrow$ 8C-a
9A: ExpField $\Rightarrow$ 8B-a
8I: $P_{8Ia} \Rightarrow$ 7F-a
8H: CBelief $\Rightarrow$ 7E-a
8G: CD'09 $\Rightarrow$ 7D-a
8F: webMiller $\Rightarrow$ 7C-a
8E: webMiller $\Rightarrow$ 7B-a
8D: $P_{8D\text{-}a}$, $DMP_{9I} \Rightarrow$ 7A-a
$8C^{i}$: 9F, 9G $\Rightarrow$ 7A-a
$8C^{ii}$: 9E, 9G $\Rightarrow$ 7A-a
$8C^{iii}$: 9D, 9G $\Rightarrow$ 7A-a
$8C^{iv}$: 9C, 9G $\Rightarrow$ 7A-a
$8C^{v}$: 9B, 9G $\Rightarrow$ 7A-a

8B: 9A $\Rightarrow$ `7A-a`
8A: $P_{8Aa} \Rightarrow$ `7A-a`
7F: $8I, P_{7Fb} \Rightarrow$ `6D-a`
7D: 8G $\Rightarrow$ `6A-b`
7C: 8F $\Rightarrow$ `6A-b`
7B: 8E $\Rightarrow$ `6A-b`
$7A^{i}$: 8D $\Rightarrow$ `6A-b`
$7A^{ii}$: $8C^{i} \Rightarrow$ `6A-b`
$7A^{iii}$: $8C^{ii} \Rightarrow$ `6A-b`
$7A^{iv}$: $8C^{iii} \Rightarrow$ `6A-b`
$7A^{v}$: $8C^{iv} \Rightarrow$ `6A-b`
$7A^{vi}$: $8C^{v} \Rightarrow$ `6A-b`
$7A^{vii}$: 8B $\Rightarrow$ `6A-b`
$7A^{viii}$: 8A $\Rightarrow$ `6A-b`
6H: $P_{6Ha} \Rightarrow$ `5E-a`
6C: $P_{6Ca}, P_{6Cb} \Rightarrow$ `5C-a`
$6A^{i}$: $P_{6Aa}, 7D \rightarrow$ `5A-a`
$6A^{ii}$: $P_{6Aa}, 7C \rightarrow$ `5A-a`
$6A^{iii}$: $P_{6Aa}, 7B \rightarrow$ `5A-a`
$6A^{iv}$: $P_{6Aa}, 7A^{i} \rightarrow$ `5A-a`
$6A^{v}$: $P_{6Aa}, 7A^{ii} \rightarrow$ `5A-a`
$6A^{vi}$: $P_{6Aa}, 7A^{iii} \rightarrow$ `5A-a`
$6A^{vii}$: $P_{6Aa}, 7A^{iv} \rightarrow$ `5A-a`
$6A^{viii}$: $P_{6Aa}, 7A^{v} \rightarrow$ `5A-a`
$6A^{ix}$: $P_{6Aa}, 7A^{vi} \rightarrow$ `5A-a`
$6A^{x}$: $P_{6Aa}, 7A^{vii} \rightarrow$ `5A-a`
$6A^{xi}$: $P_{6Aa}, 7A^{viii} \rightarrow$ `5A-a`
5E: 6H $\Rightarrow$ `4F-a`
5D: $P_{5Da} \Rightarrow$ `4F-a`
5C: 6C $\Rightarrow$ `4F-a`
5A: $\Rightarrow$ `4A-a`
$5A^{i}$: $6A^{i} \Rightarrow$ `4A-a`
$5A^{ii}$: $6A^{ii} \Rightarrow$ `4A-a`
$5A^{iii}$: $6A^{iii} \Rightarrow$ `4A-a`
$5A^{iv}$: $6A^{iv} \Rightarrow$ `4A-a`
$5A^{v}$: $6A^{v} \Rightarrow$ `4A-a`
$5A^{vi}$: $6A^{vi} \Rightarrow$ `4A-a`
$5A^{vii}$: $6A^{vii} \Rightarrow$ `4A-a`
$5A^{viii}$: $6A^{viii} \Rightarrow$ `4A-a`
$5A^{ix}$: $6A^{ix} \Rightarrow$ `4A-a`
$5A^{x}$: $6A^{x} \Rightarrow$ `4A-a`
$5A^{xi}$: $6A^{xi} \Rightarrow$ `4A-a`
4E: $P_{4Ea} \Rightarrow$ `3G-a`
4C: $P_{4Ca} \Rightarrow$ `3D-a`
4B: $P_{4Ba} \Rightarrow$ `3D-a`
$4A^{i}$: $5A^{i} \Rightarrow$ `3A-a`
$4A^{ii}$: $5A^{ii} \Rightarrow$ `3A-a`
$4A^{iii}$: $5A^{iii} \Rightarrow$ `3A-a`
$4A^{iv}$: $5A^{iv} \Rightarrow$ `3A-a`
$4A^{v}$: $5A^{v} \Rightarrow$ `3A-a`
$4A^{vi}$: $5A^{vi} \Rightarrow$ `3A-a`
$4A^{vii}$: $5A^{vii} \Rightarrow$ `3A-a`
$4A^{viii}$: $5A^{viii} \Rightarrow$ `3A-a`
$4A^{ix}$: $5A^{ix} \Rightarrow$ `3A-a`
$4A^{x}$: $5A^{x} \Rightarrow$ `3A-a`
$4A^{xi}$: $5A^{xi} \Rightarrow$ `3A-a`
3G: $4E, P_{3Gb} \Rightarrow$ `2I-b`
3F: $P_{3Fa} \Rightarrow$ `2G-a`
3E: $P_{3Ea} \Rightarrow$ `2E-a`
3D: $P_{3Da} \Rightarrow$ `2D-a`
3C: $P_{3Ca} \Rightarrow$ `2C-a`
3B: $P_{3Ba}, P_{3Bb} \Rightarrow$ `2B-b`
$DMP_{3A}{}^{i}$: $4A^{i}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{ii}$: $4A^{ii}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{iii}$: $4A^{iii}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{iv}$: $4A^{iv}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{v}$: $4A^{v}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{vi}$: $4A^{vi}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{vii}$: $4A^{vii}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{viii}$: $4A^{viii}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{ix}$: $4A^{ix}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{x}$: $4A^{x}, P_{3Ab} \Rightarrow$ `2A-b`
$DMP_{3A}{}^{xi}$: $4A^{xi}, P_{3Ab} \Rightarrow$ `2A-b`
BC_{2I}: $P_{2Ia}, 3G \Rightarrow$ `2I`
GC_{2H}: $P_{2Ha}, P_{2Hb} \Rightarrow$ `1C-a`
GC_{2FG}: $P_{2Fa}, P_{2Fb}, 3F, P_{2Gb} \Rightarrow$ `1B-a`
GC_{2G}: $3F, P_{2Gb} \Rightarrow$ `1B-a`
GC_{2F}: $P_{2Fa}, P_{2Fb} \Rightarrow$ `1B-a`
BC_{2DE}: $3D, P_{2Db}, 3E, P_{3Eb} \Rightarrow$ `2DE`
BC_{2E}: $3E, P_{2Db} \Rightarrow$ `2E`
BC_{2D}: $3D, P_{2Db} \Rightarrow$ `2D`
GC_{2ABC}: $P_{2Aa}, 3A^{n}, P_{2Ba}, 3B, 3C, P_{2Cb} \Rightarrow$ `1B-a`
GC_{2BC}: $P_{2Ba}, 3B, 3C, P_{2Cb} \Rightarrow$ `1A-a`
GC_{2AC}: $P_{2Aa}, 3A^{n}, 3C, P_{2Cb} \Rightarrow$ `1A-a`
GC_{2AB}: $P_{2Aa}, 3A^{n}, P_{2Ba}, 3B \Rightarrow$ `1A-a`

GC_{2C}: $3C, P_{2Cb} \Rightarrow$ `1A-a`
GC_{2B}: $P_{2Ba}, 3B \Rightarrow$ `1A-a`
GC_{2A}: $P_{2Aa}, 3A^{n} \Rightarrow$ `1A-a`

Attacks

GC_{2ABC}, GC_{2BC}, GC_{2AC}, GC_{2AB}, GC_{2C}, GC_{2B}, and GC_{2A} rebut GC_{2H} on `1C-a`, and GC_{2FG}, GC_{2G}, and GC_{2F} on `1B-a`.

GC_{2FG}, GC_{2G}, and GC_{2F} rebut GC_{2H} on `1C-a`, and GC_{2ABC}, GC_{2BC}, GC_{2AC}, GC_{2AB}, GC_{2C}, GC_{2B}, and GC_{2A} on `1A-a`.

GC_{2H} rebuts GC_{2FG}, GC_{2G}, and GC_{2F} on `1B-a`, and GC_{2ABC}, GC_{2BC}, GC_{2AC}, GC_{2AB}, GC_{2C}, GC_{2B}, and GC_{2A} on `1A-a`.

BC_{2D} rebuts GC_{2ABC}, GC_{2BC}, GC_{2AC}, GC_{2AB}, GC_{2C}, GC_{2B}, and GC_{2A} on `1A-a`, and vice versa on `2D`.

BC_{2E} rebuts GC_{2ABC}, GC_{2BC}, GC_{2AC}, GC_{2AB}, GC_{2C}, GC_{2B}, and GC_{2A} on `1A-a`, and vice versa on `2E`.

BC_{2DE} rebuts GC_{2ABC}, GC_{2BC}, GC_{2AC}, GC_{2AB}, GC_{2C}, GC_{2B}, and GC_{2A} on `1A-a`, and vice versa on `2DE`.

BC_{2I} and GC_{2H} rebut each other.

P_{4Da} rebuts P_{3Fa} and undermines 3F, GC_{2G}, and GC_{2FG} on `3F-a`.

P_{5Ba} undercuts 4E on *r4E*.

P_{6Ba} and P_{5Ba} rebut each other.

5C rebuts P_{3Gb} and undermines 3G on `3G-b`.

5D rebuts P_{3Gb} and undermines 3G on `3G-b`.

5E rebuts P_{3Gb} and undermines 3G on `3G-b`.

P_{6Ea} and P_{5Da} rebut each other.

P_{6Ea} undermines 5D on `5D-a`.

P_{6Fa} rebuts 6H on 6H.

P_{6Fa} rebuts 5E on 6H.

P_{6Ga} and 6H rebut each other.

P_{6Ga} rebuts 5E on 6H.

7F and 6C rebut each other.

7F rebuts 5C on 6C.

P_{7Ga} undercuts 6H on *r6H*.

P_{8Ja} and P_{7Fb} rebut each other.

P_{8Ja} undermines 7F on `7F-b`.

8H and P_{6Ca} rebut each other.

8H undermines 6C and 5C on `6C-a`.

B.2 Reconstruction in ABA

$\mathcal{A}$

r2A
r2B
r2C
r2AB
r2AC
r2BC
r2ABC
r2D
r2E
r2DE
r2F
r2G
r2FG
r2H
r2I
r3A
r3B
r3C
r3D
r3E
r3F
r3G
r4A
r4B
r4C
r4E
r5A
r5C
r5D
r5E
r6C
r6H
r6I
r6J
r6L
r6M
r7A
r7B
r7C
r7D
r7F
r8A
r8B
r8D
r8E
r8F
r8G
r8H
r8I
r9A
r9B
r9C
r9D
r9E
r9F
r9G
r9H
r9I
r10A
r10B
r10C
r10D
r10E
r10F
r10G
r11A
*not*_1A-a
*not*_1B-a
*not*_1C-a
*not*_2D
*not*_2E
*not*_2DE
*not*_2I
*not*_5C-a
*not*_6D-a
*not*_6F-a
*not*_6G-a
*not*_8J-a

$\mathcal{R}$

From $\mathcal{R}_d$:
r2A: *r2A*, *not*_1B-a, *not*_1C-a, *not*_2D, *not*_2E, *not*_2DE, 2A-a, 2A-b $\rightarrow$ 1A-a
r2B: *r2B*, *not*_1B-a, *not*_1C-a, *not*_2D, *not*_2E, *not*_2DE, 2B-a, 2B-b $\rightarrow$ 1A-a
r2C: *r2C*, *not*_1B-a, *not*_1C-a, *not*_2D, *not*_2E, *not*_2DE, 2C-a, 2C-b $\rightarrow$ 1A-a
r2AB: *r2AB*, *not*_1B-a, *not*_1C-a, *not*_2D, *not*_2E, *not*_2DE, 2A-a, 2A-b, 2B-a, 2B-b $\rightarrow$ 1A-a
r2AC: *r2AC*, *not*_1B-a, *not*_1C-a, *not*_2D, *not*_2E, *not*_2DE, 2A-a, 2A-b, 2C-a, 2C-b $\rightarrow$ 1A-a
r2BC: *r2BC*, *not*_1B-a, *not*_1C-a, *not*_2D, *not*_2E, *not*_2DE, 2B-a, 2B-b, 2C-a, 2C-b $\rightarrow$ 1A-a
r2ABC: *r2ABC*, *not*_1B-a, *not*_1C-a, *not*_2D, *not*_2E, *not*_2DE, 2A-a, 2A-b, 2B-a, 2B-b, 2C-a, 2C-b $\rightarrow$ 1A-a

r2D: *r2D*, *not*_1A-a, 2D-a, 2D-b → 2D
r2E: *r2E*, *not*_1A-a, 2E-a, 2E-b → 2E
r2DE: *r2DE*, *not*_1A-a, *not*_1C-a, 2D-a, 2D-b, 2E-a, 2E-b → 2DE
r2F: *r2F*, *not*_1A-a, *not*_1C-a, 2F-a, 2F-b → 1B-a
r2G: *r2G*, *not*_1A-a, *not*_1C-a, 2G-a, 2G-b → 1B-a
r2FG: *r2FG*, *not*_1A-a, *not*_1C-a, 2F-a, 2F-b, 2G-a, 2G-b → 1B-a
r2H: *r2H*, *not*_1A-a, *not*_1B-a, *not*_2I, 2H-a, 2H-b → 2D
r2I: *r2I*, *not*_1C-a, 2I-a, 2I-b → 2I

r3A: *r3A*, 3A-a, 3A-b → 2A-b
r3B: *r3B*, 3B-a, 3B-b → 2B-b
r3C: *r3C*, 3C-a → 2C-a
r3D: *r3D*, 3D-a, 3D-b → 2C-b
r3E: *r3E*, 3E-a → 2E-a
r3F: *r3F*, 3F-a → 2G-a
r3G: *r3G*, 3G-a, 3G-b → 2I-b
r4A: *r4A*, 4A-a → 3A-a
r4B: *r4B*, 4B-a → 3D-a
r4C: *r4C*, 4C-a → 3D-a
r4E: *r4E*, 4E-a → 3G-a
r5A: *r5A*, 5A-a → 4A-a
r5C: *r5C*, 5C-a → 4F-a
r5D: *r5D*, 5D-a → 4F-a
r5E: *r5E*, 5E-a → 4F-a
r6C: *r6C*, *not*_6D-a, 6C-a, 6C-b → 5C-a
r6H: *r6H*, *not*_6F-a, *not*_6G-a, 6H-a → 5E-a
r7A: *r7A*, 7A-a → 6A-b
r7B: *r7B*, 7B-a → 6A-b
r7C: *r7C*, 7C-a → 6A-b
r7D: *r7D*, 7D-a → 6A-b
r7F: *r7F*, *not*_5C-a, 7F-a → 6D-a
r8A: *r8A*, 8A-a → 7A-a
r8B: *r8B*, 8B-a → 7A-a
r8D: *r8D*, 8D-a → 7A-a
r8E: *r8E*, webMiller → 7B-a
r8F: *r8F*, webMiller → 7C-a
r8G: *r8G*, cd'09 → 7D-a
r8H: *r8H*, *not*_6C-a, cBelief → 7E-a
r8I: *r8I*, *not*_8J-a, 8I-a → 7F-a
r9A: *r9A*, expField → 8B-a
r9B: *r9B*, 9B-a → 8C-a
r9C: *r9C*, 9C-a → 8C-a
r9D: *r9D*, 9D-a → 8C-a
r9E: *r9E*, 9E-a → 8C-a
r9F: *r9F*, 9F-a → 8C-a
r9G: *r9G*, 9G-a, 9G-b → 8C-b
r9H: *r9H*, ipcc'07 → 8D-a
r9I: *r9I*, 9I-a → 8D-b
r10A: *r10A*, ipcc'07 → 9B-a
r10B: *r10B*, ipcc'07 → 9C-a
r10C: *r10C*, ipcc'07 → 9D-a
r10D: *r10D*, ipcc'07 → 9E-a
r10E: *r10E*, ipcc'07 → 9F-a
r10F: *r10F*, 10F-a, 10F-b → 9G-a
r10G: *r10G*, mit → 9I-a
r11A: *r11A*, 11A-a → 10F-b

From $\mathcal{R}_s$:
r6A: 6A-a, 6A-b → 5A-a
r8C: 8C-a, 8C-b → 7A-a

From $\mathcal{K}_n$:
rCBelief: → cBelief
rCD'09: → cd'09
rExpField: → expField
rWebMiller: → webMiller

rIPCC'07: $\rightarrow$ ipcc'07 *rMIT*: $\rightarrow$ mit

$\mathcal{C}$

1A-a $\in \mathcal{C}(\mathit{not}$ _1A-a)
1B-a $\in \mathcal{C}(\mathit{not}$ _1B-a)
1C-a $\in \mathcal{C}(\mathit{not}$ _1C-a)
2D $\in \mathcal{C}(\mathit{not}$ _2D)
2E $\in \mathcal{C}(\mathit{not}$ _)
2DE $\in \mathcal{C}(\mathit{not}$ _2DE)
2I $\in \mathcal{C}(\mathit{not}$ _2I)
5C-a $\in \mathcal{C}(\mathit{not}$ _5C-a)
6D-a $\in \mathcal{C}(\mathit{not}$ _6D-a)
6F-a $\in \mathcal{C}(\mathit{not}$ _6F-a)
6G-a $\in \mathcal{C}(\mathit{not}$ _6G-a)
8J-a $\in \mathcal{C}(\mathit{not}$ _8J-a)
1A-a $= -$1B-a
1A-a $= -$1C-a
1B-a $= -$1C-a
2D $= -$1A-a
2E $= -$1A-a
2DE $= -$1A-a
2I $= -$1C-a
4D-a $\in \mathcal{C}($3F-a)
4F-a $\in \mathcal{C}($3G-b)
5B-a $\in \mathcal{C}(\mathit{r4E})$
6B-a $= -$5B-a
6D-a $= -$5C-a
6E-a $= -$5D-a
6F-a $\in \mathcal{C}($5E-a)
6G-a $= -$5E-a
7E-a $= -$6C-a
7G-a $\in \mathcal{C}(\mathit{r6H})$
8J-a $= -$7F-b