Argument Revision

MARK SNAITH and CHRIS REED, *Centre for Argument Technology, University of Dundee, Dundee DD1 4HN, Scotland.*
E-mail: m.snaith@dundee.ac.uk; c.a.reed@dundee.ac.uk

Abstract

Understanding the dynamics of argumentation systems is a crucial component in the development of computational models of argument that are used as representations of belief. To that end, in this article, we introduce a model of *Argument Revision*, presented in terms of the contraction and revision of a system of structured argumentation. Argument Revision is influenced by the AGM model of belief revision, but with certain key differences. First, Argument Revision involves modifying the underlying model (system of argumentation) from which beliefs are derived, allowing for a finer-grained approach to modifying beliefs. Secondly, the richer structure provided by a system of argumentation permits a determination of minimal change based on quantifiable effects on the system as opposed to qualitative criteria such as entrenchment orderings. Argument Revision does, however, retain a close link to the AGM approach to belief revision. A basic set of postulates for rational revisions and contractions in Argument Revision is proposed; these postulates are influenced by, and capture the spirit of, those found in AGM belief revision. After specifying a determination of minimal change, based on measurable effects on the system, we conclude the article by going on to show how Argument Revision can be used as a strategic tool by a participant in a multi-agent dialogue, assisting with commitment retraction and dishonesty. In systems of argumentation that contain even small knowledge bases, it is difficult for a dialogue participant to fully assess the impact of seemingly trivial changes to that knowledge base, or other parts of the system; we demonstrate, by means of an example, that Argument Revision solves this problem through a determination of minimal change that is justifiable and intuitive.

Keywords: Argumentation, argument revision, argument dynamics, belief revision, dishonesty.

1 Introduction

Connections between argumentation and belief revision have recently found new momentum [23, 24, 26] with the emergence of two broad approaches—the use of argumentation to support belief revision (e.g. [25, 31]) and the use of belief revision operators to model change in argumentation systems (e.g. [4, 14, 39]).

The dynamic argumentation framework defined in Argument Theory Change (ATC) [35, 39] is based on Dung-style abstract frameworks [22] but incorporates a model of dynamics. The exploration of change in abstract argumentation in [14] assesses the impact of change by using existing properties of Dung's work without defining a new version of it. Similarly, the work of [4] explores the dynamics of an argument-based decision system based on the addition of a new argument. Decision systems use arguments to support and decide between different options, with the aim of the work in [4] being to investigate the impact that a new argument has on the ordering of the options, again without defining a new version of the system. Both approaches have their advantages: defining a new argumentation framework that incorporates a model of dynamics results in a means of assessing the effects of a change being more readily available. Additionally, through the use of a sub-argument relation, the framework defined in ATC is also structured, providing a better model of dynamics and a clear link to natural arguments, which are inherently structured. On the other hand, developing a model of dynamics for an existing framework decouples the model from the framework. Such a decoupling has the advantage of being able to assess the dynamics of an existing framework, with established properties, without needing to extend the framework and hence prove its properties still

Vol. 27 No. 7, © The Author, 2016. Published by Oxford University Press.

Published online September 13, 2016 doi:10.1093/logcom/exw028

hold. Furthermore, both ATC and the works of [4, 14] consider only the addition and, in the case of ATC, removal of arguments (and associated interactions). Neither consider the effects of, for instance, making an argument unacceptable (w.r.t. some Dung-style [22] semantics). Both models are based on Dung's abstract approach, and thus incorporate a mechanism for determining argument acceptability. Our approach is to combine the advantages of these foundations in developing a new model of Argument Revision.

Our new model is influenced by the AGM model of belief revision. We propose sets of postulates (Section 6) that describe rational revisions and contractions of a system of argumentation in a similar vein to the AGM postulates that describe rational revisions and contractions of belief sets. Furthermore, and again similar to belief revision, we incorporate into Argument Revision a determination of minimal change. This determination is quantitative, based on measurable effects that a change has on the system.

In classic belief revision, modifications are carried out on flat belief sets or (logically closed) belief bases. In Argument Revision, we modify the underlying representation from which belief is derived, specifically the ASPIC$^+$ system of argumentation. This system provides rich structure and incorporates Dung's theory of abstraction argumentation [22]. This allows revisions to be performed by either modifying arguments themselves or merely their *acceptability*. In both cases, the effect on the derived beliefs can be the same (e.g. no longer believing a certain statement) but the actual modifications in the underlying representation are significantly different.

Argumentation and belief revision are both used for managing inconsistency. The aim of Argument Revision is to intersect the two fields, providing a mechanism that allows the dynamics of a system of structured argumentation to be modelled. This mechanism finds a use in two specific scenarios in multi-agent dialogue—commitment retraction and dishonesty. Both scenarios can require a dialogue participant to choose between several different actions. Techniques specified by Argument Revision assist them in making their choice.

The work in this article is built on our previous work in [42, 43] in four ways: (i) by extending Argument Revision to allow rules, preferences and contraries to be revised as well as the knowledge base. This is achieved through a specification of modular argumentation theories; (ii) refining the types of Argument Revision into those that more closely reflect those found in AGM belief revision—*revision* and *contraction*; (iii) proposing a set of postulates for rational revisions and contractions in Argument Revision; and (iv) showing how dishonesty can be used as an alternative to retracting existing commitments in dialogue.

The article proceeds as follows: in Section 2 we provide a brief introduction to related work upon which this paper builds; in Section 3 we define a modular Argumentation Theory in the ASPIC$^+$ framework, which assists in simplifying the Argument Revision process; in Sections 4 and 5 we specify a model for Argument Revision, describing the process and properties; in Section 6 we propose a set of basic postulates for rational revisions and contractions in Argument Revision; in Section 7 measures of minimal change in Argument Revision are defined; in Section 8 we show how Argument Revision can be used in a dialogical context both for commitment retraction and dishonesty. Finally, in Section 10 we conclude the article and outline directions for future work.

2 Background

2.1 Argumentation

One of the main bodies of work in computational argumentation is abstract argumentation, building on the influential work of Dung [22] in which arguments are reduced to purely abstract entities, over which an attack relation can be defined. The main concept is that of an argumentation framework.

DEFINITION 2.1 (Argumentation Framework)
An argumentation framework is a pair $\mathcal{AF} = \langle \mathcal{A}rgs, \mathcal{A}tts \rangle$ where $\mathcal{A}rgs$ is a set of arguments and $\mathcal{A}tts \subseteq \mathcal{A}rgs \times \mathcal{A}rgs$, a binary attack relation.

In this article we will use the ASPIC$^+$ framework [38], a structured instantiation of Dung's work. We choose ASPIC$^+$ because its combination of structure and acceptability provides a rich model for specifying revision operators. Furthermore, it has an established connection to the Argument Interchange Format (AIF) [8, 15] which underpins the Argument Web [7], allowing the present work to be applied, in principle, to natural language arguments and real online debates.

The ASPIC$^+$ framework builds on the work of [3]. ASPIC$^+$ instantiates Dung's abstract approach by combining the work of [37] on defeasible reasoning with the work of [44] on the structure of arguments. Definitions 2.2 through 2.8 in the remainder of this subsection characterize the basic principles of ASPIC$^+$.

The fundamental notions of ASPIC$^+$ are *argumentation systems* and a *knowledge base*:

DEFINITION 2.2 (Argumentation System)
An argumentation system is a tuple $\mathcal{AS} = \langle \mathcal{L}, cf, \mathcal{R}, \leq \rangle$ where $\mathcal{L}$ is a logical language; *cf* is a contrariness function from $\mathcal{L}$ to $2^{\mathcal{L}}$; $\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_d$ is a set of strict ($\mathcal{R}_s$) and defeasible ($\mathcal{R}_d$) inference rules such that $\mathcal{R}_s \cap \mathcal{R}_d = \emptyset$; and $\leq$ is a partial pre-order on $\mathcal{R}_d$.

REMARK 2.3
Defeasible rules take the form $\varphi_1, \ldots, \varphi_n \Rightarrow \varphi$ while strict rules take the form $\varphi_1 \ldots, \varphi_n \rightarrow \varphi$. In some cases, we label rules using a subscript on the consequent operator, thus: $\Rightarrow_{r_1}$ or $\rightarrow_{r_2}$ (for some defeasible rule r_1 and some strict rule r_2).

As in [38], we do not specify the nature of the logical language because one of the cornerstones of the ASPIC$^+$ framework is that it can, in principle, use any logical language. We do, however, place a single constraint on the language, specifically that it contains modular representations of components of an argumentation system. This is explained in more detail in Section 3.

For some $\phi_1 \in \mathcal{L}$, the notation $\overline{\phi_1}$ is the set of all contraries of ϕ_1 (i.e. $cf(\phi_1) = \overline{\phi_1}$). Given some other formula $\phi_2 \in \mathcal{L}$, $\phi_1 \in \overline{\phi}_2$ means, informally, 'ϕ_1 is a contrary of ϕ_2'.

DEFINITION 2.4 (Knowledge base)
A knowledge base in an argumentation system $\langle \mathcal{L}, cf, \mathcal{R}, \leq \rangle$ is a pair $\langle \mathcal{K}, \leq' \rangle$ where $\mathcal{K} \subseteq \mathcal{L}$ and $\mathcal{K} = \mathcal{K}_n \cup \mathcal{K}_p \cup \mathcal{K}_a$, with $\mathcal{K}_n \cap \mathcal{K}_p \cap \mathcal{K}_a = \emptyset$ and:

- $\mathcal{K}_n$ being a set of (necessary) axioms. Arguments cannot be attacked on their axiom premises;
- $\mathcal{K}_p$ being a set of (ordinary) premises. Arguments can be attacked on their ordinary premises, with the success of the attack depending on preferences between the arguments;
- $\mathcal{K}_a$ being a set of assumptions. Arguments can be attacked on their assumption premises, with the attack always succeeding.

$\leq'$ is a partial pre-order on $\mathcal{K} \backslash \mathcal{K}_n$

In Section 5, we make use of the distinction between ordinary premises and assumptions in specifying the process of Argument Revision.

From the knowledge base ($\mathcal{K}$) and rules ($\mathcal{R}$) of an argumentation system, arguments are constructed. For an argument $\mathcal{A}$, *Prem*($\mathcal{A}$) is a function that returns all the premises in $\mathcal{A}$; *Conc*($\mathcal{A}$) is a function that returns the conclusion of $\mathcal{A}$; *Sub*($\mathcal{A}$) is a function that returns all the sub-arguments of $\mathcal{A}$; *DefRules*($\mathcal{A}$) is a function that returns all defeasible rules in $\mathcal{A}$; and *TopRule*($\mathcal{A}$) is a function that returns the last inference rule used in $\mathcal{A}$.

We also overload two of these functions for sets of arguments. Given a set of arguments $\mathcal{S}$, $Prem(\mathcal{S}) = \bigcup_{\mathcal{A}\in\mathcal{S}} Prem(\mathcal{A})$; and $Sub(\mathcal{S}) = \bigcup_{\mathcal{A}\in\mathcal{S}} Sub(\mathcal{A})$.

On the basis of these functions, $\mathcal{A}$ is:

(1) p if $p\in\mathcal{K}$ with: $Prem(\mathcal{A})=\{p\}$; $Conc(\mathcal{A})=p$; $Sub(\mathcal{A})=p$; $DefRules(\mathcal{A})=\emptyset$; $TopRule(\mathcal{A})=$ undefined.

(2) $\mathcal{A}_1,\ldots,\mathcal{A}_n\rightarrow\psi$ if $\mathcal{A}_1,\ldots,\mathcal{A}_n$ are arguments such that there exists a strict rule $Conc(\mathcal{A}_1),\ldots Conc(\mathcal{A}_n)\rightarrow\psi$ in $\mathcal{R}_s$ with: $Prem(\mathcal{A})=Prem(\mathcal{A}_1)\cup\ldots\cup Prem(\mathcal{A}_2)$; $Conc(\mathcal{A})=\psi$; $Sub(\mathcal{A})=Sub(\mathcal{A}_1)\cup\ldots\cup Sub(\mathcal{A}_n)\cup\{\mathcal{A}\}$;
$DefRules(\mathcal{A})=DefRules(\mathcal{A}_1)\cup\ldots\cup DefRules(\mathcal{A}_n)$;
$TopRule(\mathcal{A})=Conc(\mathcal{A}_1),\ldots,Conc(\mathcal{A}_n)\rightarrow\psi$.

(3) $\mathcal{A}_1,\ldots,\mathcal{A}_n\Rightarrow\psi$ if $\mathcal{A}_1,\ldots,\mathcal{A}_n$ are arguments such that there exists a defeasible rule $Conc(\mathcal{A}_1),\ldots Conc(\mathcal{A}_n)\Rightarrow\psi$ in $\mathcal{R}_d$ with: $Prem(\mathcal{A})=Prem(\mathcal{A}_1)\cup\ldots\cup Prem(\mathcal{A}_2)$; $Conc(\mathcal{A})=\psi$;
$Sub(\mathcal{A})=Sub(\mathcal{A}_1)\cup\ldots\cup Sub(\mathcal{A}_n)\cup\{\mathcal{A}\}$; $DefRules(\mathcal{A})=DefRules(\mathcal{A}_1)\cup\ldots\cup DefRules(\mathcal{A}_n)$;
$TopRule(\mathcal{A})=Conc(\mathcal{A}_1),\ldots,Conc(\mathcal{A}_n)\Rightarrow\psi$.

An argument can be attacked in three ways: on a (non-axiom) premise (undermine), on a defeasible inference rule (undercut) or on a conclusion (rebut). Formally, undermining and rebutting are defined thus:

Definition 2.5 (Undermining)
An argument $\mathcal{A}$ **undermines** an argument $\mathcal{B}$ if $\exists\phi\in Prem(\mathcal{B})$ s.t. $Conc(\mathcal{A})\in\overline{\phi}$

Definition 2.6 (Rebutting)
An argument $\mathcal{A}$ **rebuts** an argument $\mathcal{B}$ (on $\mathcal{B}'$) if $\exists\mathcal{B}'\in Sub(\mathcal{B})$ s.t. $Conc(\mathcal{A})\in\overline{\mathcal{B}'}$

A definition of undercutting is provided in Section 3.1 (Definition 3.3), since it relies on aspects of the model we introduce later in the article.

Given an argumentation system $\mathcal{AS}$ and a knowledge base $\mathcal{KB}$, an *Argumentation Theory* is defined:

Definition 2.7 (Argumentation Theories)
An *Argumentation Theory* is a triple $\mathcal{AT}=\langle\mathcal{AS},\mathcal{KB},\preceq\rangle$ where $\preceq$ is an argument ordering on the set of all arguments that can be constructed from *KB* in $\mathcal{AS}$.

The argument ordering in an Argumentation Theory is arrived at by considering the orderings on rules and knowledge base premises. Two such orderings are specified in [38], *last-link* and *weakest-link*. Formal definitions of these orderings are provided in [38], but rely on additional formalisms that are not required by the present work. Informally, the definitions are as follows:

> 'The **last-link principle** prefers an argument $\mathcal{A}$ over another argument $\mathcal{B}$ if the last defeasible rules used in $\mathcal{B}$ are less preferred than the last defeasible rules in $\mathcal{A}$ or, in case both arguments are strict, if the premises of $\mathcal{B}$ are less preferred than the premises of $\mathcal{A}$.'

> 'The **weakest-link principle** considers not the last but all uncertain elements in an argument. It prefers an argument $\mathcal{A}$ over an argument $\mathcal{B}$ if $\mathcal{A}$ is preferred to $\mathcal{B}$ on both their premises and their defeasible rules.'

To ensure that Argumentation Theories respect the intended use of contrariness and subsequently satisfy certain rationality postulates, a *well-formedness* condition is imposed.

DEFINITION 2.8 (Well-formedness)
An Argumentation Theory is *well-formed* iff, if $\phi \in \overline{\psi}$:

- $\psi \notin \mathcal{K}_n$; and
- $\nexists r \in \mathcal{R}_s$ s.t. $Cons(r) = \psi$

Informally, Definition 2.8 stipulates that neither axioms nor consequents of strict rules can be contradicted. The second condition is required for ASPIC$^+$ argumentation theories to satisfy Caminada and Amgoud's 'indirect consistency' rationality postulate [13]. This rationality postulate requires that the closure of the set of conclusions of all arguments in an extension under strict-rule application is consistent. Permitting contraries to the consequents of strict rules risks violating this condition; consider in the simplest case the arguments $\mathcal{A}_1$: p, $\mathcal{A}_2$: $\mathcal{A}_1 \rightarrow q$ and $\mathcal{A}_3$: r, and that $r \in \overline{q}$. The set of acceptable conclusions is $\{p, r\}$; the closure of this set under strict rule application is $\{p, q, r\}$, which is inconsistent.

Given an Argumentation Theory $\mathcal{AT} = \langle \mathcal{AS}, \mathcal{KB}, \preceq \rangle$, in this article: $\mathcal{K}(\mathcal{AT})$ is the knowledge base in $\mathcal{AS}$; $\mathcal{R}(\mathcal{AT})$ is the rule set in $\mathcal{AS}$; and $\mathcal{A}rgs(\mathcal{AT})$ is the set of all arguments in $\mathcal{AT}$ (on the basis of $\mathcal{K}(\mathcal{AT})$ and $\mathcal{R}(\mathcal{AT})$). Henceforth, unless otherwise stated, if any examples do not refer to any of these sets, they are assumed to be empty. Similarly, where no preference information is given in an example, none exists.

The abstract framework derived from $\mathcal{AT}$ will be denoted as $\mathcal{AF}_{\mathcal{AT}}$. When considering the acceptability of arguments we will, unless otherwise stated, use the grounded semantics, with $E(\mathcal{AT})$ denoting the grounded extension.

We will also use two functions over rules. For a rule $r \in \mathcal{R}(\mathcal{AT})$, $Ant(r)$ returns the antecedents of the rule and $Cons(r)$ returns the consequent. We also assume that a rule's antecedents and consequent are expressed in conjunctive normal form.

2.2 *Belief revision*

Belief revision is the study of how a knowledge base can accommodate new or conflicting information. The field has been shaped by [1] in which postulates are proposed for three types of change that can be made to a knowledge base K:

- **Expansion**, denoted $K + \Phi$, where a new sentence Φ is added to K, together with its logical consequences;
- **Revision**, denoted $K \dot{+} \Phi$, where a new sentence that is inconsistent with K is added, but consistency is maintained by removing some sentences from K; and
- **Contraction**, denoted $K \dot{-} \Psi$, where a sentence Φ in K is retracted without adding any new sentences. To maintain closure under logical consequences, some other sentences may need to be given up.

There are eight postulates each for revision and contraction, with the first six known as the *basic postulates* [28], and the seventh and eighth describing composite operations. Our focus in this article will be on the six basic postulate, which are summarized in Tables 1 and 2 respectively. The notation $K_\perp$, used in $K\dot{+}5$, represents an inconsistent belief set.

TABLE 1. AGM postulates for revision

$(K\dot{+}1)$	The outputs of the revision function are belief sets.	$K\dot{+}\Phi$ is a belief set
$(K\dot{+}2)$	The input sentence Φ is accepted in $K\dot{+}\Phi$.	$\Phi \in K\dot{+}\Phi$
$(K\dot{+}3)$	Defined together and cover	$K\dot{+}\Phi \subseteq K+\Phi$
$(K\dot{+}4)$	the case $\neg\Phi \notin K$.	If $\neg\Phi \notin K$, then $K+\Phi \subseteq K\dot{+}\Phi$
$(K\dot{+}5)$	$K\dot{+}\Phi$ is consistent unless Φ is logically impossible.	$K\dot{+}\Phi = K_{\perp}$ iff $\vdash \neg\Phi$
$(K\dot{+}6)$	Revisions are based on the *knowledge* level and not the *analytic*.	If $\vdash \Phi \leftrightarrow \Psi$, then $K\dot{+}\Phi = K\dot{+}\Psi$

TABLE 2. AGM postulates for contraction

$(K\dot{-}1)$	The outputs of the revision function are belief sets.	$K\dot{-}\Phi$ is a belief set
$(K\dot{-}2)$	No new beliefs occur in $K\dot{-}\Phi$.	$K\dot{-}\Phi \subseteq K$
$(K\dot{-}3)$	When $\Phi \notin K$, nothing is retracted.	If $\Phi \notin K$, $K\dot{-}\Phi = K$
$(K\dot{-}4)$	The belief set $K\dot{-}\Phi$ does not entail Φ, unless Φ is logically valid.	If not $\vdash \Phi$, then $\Phi \notin K\dot{-}\Phi$
$(K\dot{-}5)$	The *recovery postulate*; contractions can be 'undone'.	If $\Phi \in K$, then $K \subseteq (K\dot{-}\Phi)+\Phi$
$(K\dot{-}6)$	Contractions are based on the *knowledge* level and not the *analytic*.	If $\vdash \Phi \leftrightarrow \Psi$, then $K\dot{-}\Phi = K\dot{-}\Psi$

A main concern of the AGM theory is deciding exactly how a belief set should be revised or contracted.[1] Several different principles have been proposed, some based on measurable effects and others using qualitative criteria.

The *principle of economy* seeks to keep loss to a minimum, where the loss to be minimized needs to be specified [5]. One such specification, using measurable effects, is in terms of set-theoretic inclusion, proposed in the seminal AGM paper [1]: given a belief set K and a choice of possible revisions, the revision that leaves the largest subset of K should be chosen. This is known as the *principle of conservatism* [29]. This specification, however, has a significant drawback in that it can lead to undesirable results [2, 41]. Nevertheless, it forms the basis of *partial-meet contractions*, which select the 'best' elements in the set of all maximal subsets of K that do not imply the belief(s) to be contracted [1].

Principles based on qualitative criteria include the *principle of entrenchment*, where an entrenchment ordering is placed over beliefs: given a belief set K and two formulae $\phi, \psi \in K$, $\phi < \psi$ means 'ϕ is less entrenched than ψ' and thus, given a choice, ϕ is more likely to be given up [27].

The principle of entrenchment is similar to the principles of indifference, strict preference and weak indifference proposed in [41]. In the *principle of indifference* objects held in equal regard should be treated equally; in the *principle of strict preference* objects held in higher regard should be afforded a more favourable treatment. Together, these lead to the *principle of weak indifference*, where if an object Φ is held in equal or higher regard than another object Ψ, Φ should be treated no worse than Ψ.

Finally, inspired by the idea of minimizing the loss of informational value, the *principle of cognitive economy* combines the quantitative principle of economy with the qualitative principle of entrenchment [5]. This is guided by the *principle of weak monotony*: let V be an index on theories; if $X \subset Y$, $V(X) \leq V(Y)$. In other words, Y might carry more information than X, but the overall value of that information may be the same.

[1] No such consideration is required for expansion, because an expansion does not consider the logical consequences of adding a new belief.

These various principles share a common goal—that any losses (or gains) in a belief set should in some way be minimized, either through measurable effects or through the use of an entrenchment (preference) ordering over beliefs.

Another consideration in classic theories of belief revision, and specifically belief contraction, is how to frame the problem: if the intention is to contract some formula ϕ from a belief set K, should this be framed as *'what should be retained in K to ensure* $K \nvdash \phi$*?'* or *'what should be removed from K to ensure* $K \nvdash \phi$*?'*. The former case (what should be retained) is *partial-meet* contraction; the latter (what should be given up) is *safe contraction* [40]. Both approaches differ in other respects as well, with a full list of these differences provided in [40].

2.3 Belief revision and argumentation

Having described our chosen system of structured argumentation (ASPIC$^+$) and briefly reviewed various approaches and techniques to Belief Revision, we now outline how we intend to combine the two in the remainder of the article.

ASPIC$^+$ provides a preference ordering over arguments and the components from which they are comprised (premises and rules), which suggests a connection to qualitative, entrenchment-based approaches to minimal change. Using ASPIC$^+$'s preference ordering in this way is, however, too simplistic because it is designed as a mechanism for determining defeat; it is not necessarily an indicator of informational value. The importance of information in a system of argumentation lies in the role that information plays in the structure of arguments and the interactions (attacks) between them. Removing a single premise or attack can have a considerable effect on the remaining arguments, both in terms of structure and acceptability. As such, in this article we will focus on a specification of minimal change in structured argumentation that is influenced by the principle of conservatism. This is described in more detail in Section 7.

In framing the problem of contraction (safe contraction vs. partial-meet contraction; what to give up vs. what to retain), we find that both belief revision approaches are insufficient for argument contraction, which is defined in terms of the acceptability of arguments. This means arguments can be contracted by adding new, defeating arguments, something that neither safe nor partial-meet contraction accounts for.

While in principle both approaches could be extended to account for additions (i.e. 'what to give up **or add**' and 'what to retain **or add**' respectively), this is still insufficient. For safe-contraction, an ordering over elements is required to determine what should be given up but. A simple relation over formulae would not work, however, for we would also need to express a preference between addition and removal—consider, for instance, giving up some formula ϕ vs. adding another formula ψ.

Partial-meet contraction cannot be extended to account for additions because taking the intersection of all theories would ignore anything that was added to any theories; in other words, the final contracted theory will only ever contain elements present in the original theory.

Our operator for argument contraction is therefore instantiated using a new approach. In Section 5, we define a structure called a *change graph* that models all possible ways in which an Argumentation Theory can be revised or contracted with respect to a set of input arguments. A path cost function based on four measures of change (defined in Section 7) is then used to select the 'best' (i.e. lowest cost) contraction. Our operator for argument revision is also instantiated using change graphs, meaning they also offer a unified solution to both types of change in to Argumentation Theory.

Finally, while classic belief revision is carried out on belief sets (a set of sentences closed under logical consequence) or belief bases (which are not necessarily closed), Argument Revision is carried out with respect to the arguments in a system of argumentation. While this might involve modifying the system's knowledge base, the effects of the revision and its success are examined in terms of the arguments and/or their acceptability. Arguments are constructed by closing a knowledge base with respect to the strict and defeasible rules in the system, and so our approach to argument revision bears similarities to the revision of belief sets.

3 Modular Argumentation Theories

In this section, we specify a modularity constraint for ASPIC$^+$ Argumentation Theories which requires that rules, preferences and contraries have explicit representations in the object language, and that the Argumentation Theory contains arguments whose conclusions are those representations. This provides two key advantages for Argument Revision; first, every component of an argumentation system has an argument supporting it. This allows for a single method of revising arguments to be specified, which can then be applied to the arguments for components (i.e. rules, preferences and contraries) thus providing new ways of revising arguments (e.g. by removing a rule).

Secondly, it allows for a principled characterization of the distinction between strict and defeasible rules. A rule is defeasible iff there exists some exceptional circumstance(s) in which it cannot be applied; otherwise, it is strict. Note that it does not need to be the case that the source of the exception is currently true (or, in the context of an Argumentation Theory, be supported by an acceptable argument)—all that needs to exist is the declaration that *if* the exceptional formula is true, then the rule does not apply. We define this formally by stipulating that a rule is strict iff the formula that represents it has no contraries; otherwise, it is defeasible. In a broader context, this characterization of strict and defeasible rules allows for a more precise definition of undercutting than is provide in [38]. We introduce this definition (Definition 3.3) following the formal specification of modular argumentation theories.

To show the difference between a component (rule, preference and contrary) and the formulae representing it, we use the notation $\lceil\rceil$. For instance, if r is a rule, $\lceil r \rceil$ is the formula representing r. This representation can also be nested; for instance, $\lceil x \in \overline{\lceil r \rceil} \rceil$ is the formula representing that x is a contrary of $\lceil r \rceil$, and thus that r is defeasible with a potential exception x (see Subsection 3.1 below).

Definition 3.1 (Modular Argumentation Theory)
An Argumentation Theory $\mathcal{AT} = \langle \mathcal{AS}, KB, \preceq \rangle$, where $\mathcal{AS} = \langle \mathcal{L}, cf, \mathcal{R}, \leq \rangle$ is **modular** iff:

- $r \in \mathcal{R}(\mathcal{AT}) \leftrightarrow \exists \mathcal{A} \in Args(\mathcal{AT})$ s.t. $Conc(\mathcal{A}) = \lceil r \rceil$
- $\phi \in \overline{\psi} \leftrightarrow \exists \mathcal{A} \in Args(\mathcal{AT})$ s.t. $Conc(\mathcal{A}) = \lceil \phi \in \overline{\psi} \rceil$
- $\phi \leq' \psi \leftrightarrow \exists \mathcal{A} \in Args(\mathcal{AT})$ s.t. $Conc(\mathcal{A}) = \lceil \phi \leq' \psi \rceil$
- $r_1 \leq r_2 \leftrightarrow \exists \mathcal{A} \in Args(\mathcal{AT})$ s.t. $Conc(\mathcal{A}) = \lceil r_1 \leq r_2 \rceil$

In broad terms, given a component (rule, preference or contrary) Φ in a modular Argumentation Theory, there is an argument (that is possibly atomic, i.e. having no supporting premises) for $\lceil \Phi \rceil$ in the Argumentation Theory; conversely, if there is an argument with conclusion $\lceil \Psi \rceil$ in an Argumentation Theory $\mathcal{AT}$, Ψ is a component (of whichever type) in the argumentation system in $\mathcal{AT}$.

Note that modularity does not extend ASPIC$^+$ but instead simply puts a constraint on the object language (it must contain representations of rules, preferences and contraries) and the construction of the theory (it must contain arguments that conclude the representations of all rules, preferences and contraries).

A modular Argumentation Theory allows arguments about the components of arguments to be constructed. This is similar to meta-argumentation, in particular the characterization in [46]. The difference, however, is that the meta-Argumentation Theory defined in [46] only requires that rules, preferences and contraries be represented in a meta-language with no guarantee that arguments for them exists. While in principle it would be possible to impose such a constraint, this introduces extra complexity not required for Argument Revision.

It is also possible to model preferences as contraries of existing attacks in a modular Argumentation Theory. Formally, this is specified as: if $\exists \mathcal{A}_1, \mathcal{A}_2 \in \mathcal{A}rgs$ s.t. $Conc(\mathcal{A}_1) \in \overline{Conc(\mathcal{A}_2)}$ and $\mathcal{A}_1 \preceq \mathcal{A}_2$, then $\lceil \mathcal{A}_1 \preceq \mathcal{A}_2 \rceil \in \overline{\lceil Conc(\mathcal{A}_1) \in \overline{Conc(\mathcal{A}_2)} \rceil}$. In other words, $\mathcal{A}_2$ being preferred to $\mathcal{A}_1$ is an attack on $\mathcal{A}_1$ attacking $\mathcal{A}_2$. This characterization proves advantageous when deriving an extended argumentation framework (EAF) from a modular Argumentation Theory (see Section 3.2).

To illustrate our specification of modular argumentation theories we use the familiar example from Pollock [37]:

> *The object appears red therefore the object is red.*

Assuming no other information, this can be encoded in ASPIC$^+$ with the following strict rule:

$$\mathcal{R}(\mathcal{AT}) = \{object\_appears\_red \rightarrow_{r1} object\_is\_red\}$$

For the Argumentation Theory containing $\mathcal{AS}$ to be modular, a corresponding argument for $\lceil r1 \rceil$ is required. This can be achieved by a premise in the knowledge base:

$$\mathcal{K}(\mathcal{AT}) = \{\lceil object\_appears\_red \rightarrow_{r1} object\_is\_red \rceil\}$$

If $\mathcal{K}(\mathcal{AT})$ also contained *object_appears_red*, we could strictly infer that *object_is_red*. The inference is strict because the Argumentation Theory does not contain any information to suggest otherwise. This, however, ignores the point made by Pollock in [37]—an object may appear red if it is not, if it is being illuminated by a red light; in other words, r_1 should be a defeasible inference rule, not strict. This will be addressed in the next sub-section.

3.1 Strict and defeasible inference rules

Modular argumentation theories allow for a principled characterisation of strict and defeasible rules:

DEFINITION 3.2 (Strict and defeasible rules)
$r \in \mathcal{R}_s(\mathcal{AT}) \leftrightarrow \overline{\lceil r \rceil} = \emptyset$
$r \in \mathcal{R}_d(\mathcal{AT}) \leftrightarrow \overline{\lceil r \rceil} \neq \emptyset$

In Argument Revision, this characterization allows strict rules to be revised into defeasible rules by introducing a potential contrary with respect to the rule (again, without necessarily introducing the formula that is the contrary). Conversely, a defeasible rule can be revised into a strict rule by removing its contraries.

Returning to our example, we provide some new information:

> *The object appears red therefore the object is red* ***unless the object is illuminated by a red light.***

This new information refines the ASPIC$^+$ representation by introducing the following contrary:

$$illum\_by\_red\_light \in \overline{\ulcorner object\_appears\_red \Rightarrow_{r1} object\_is\_red \urcorner}$$

This leads to the following knowledge base and rule set:

$$\mathcal{K}(\mathcal{AT}) = \left\{ \begin{array}{l} \ulcorner object\_appears\_red \Rightarrow_{r1} object\_is\_red \urcorner, \\ \ulcorner illum\_by\_red\_light \in \overline{\ulcorner object\_appears\_red \Rightarrow_{r1} object\_is\_red \urcorner} \urcorner \end{array} \right\}$$

$$\mathcal{R}(\mathcal{AT}) = \{object\_appears\_red \Rightarrow_{r1} object\_is\_red\}$$

Notice that r_1 is now defeasible because the formula that represents it has an identified contrary. The contrary itself (*illum_by_red_light*) cannot be inferred but this is not required. Simply identifying that there *could* be a contrary argument is sufficient to make a rule defeasible.

Where a rule's contrary can be inferred, the exception is activated and the rule is undercut. In [38], there is an assumption that rules can be represented in the object language of an argumentation system in order to represent undercutting. Using modular argumentation, we provide a more precise definition of undercutting that uses the formulae representing rules as opposed to the rules themselves. For a rule to be defeasible, the formula representing it will have at least one potential contrary (per Definition 3.2). If there then exists an argument for a contrary in the ground system, the rule is undercut by that argument. This is formally defined thus:

DEFINITION 3.3 (Undercutting)
An argument $\mathcal{A} \in \mathcal{A}rgs(\mathcal{AT})$ *undercuts* an argument $\mathcal{B} \in \mathcal{A}rgs(\mathcal{AT})$ (on $\mathcal{B}' \in Sub(\mathcal{B})$) iff $Conc(\mathcal{A}) \in cf(\ulcorner TopRule(\mathcal{B}') \urcorner)$.

In other words, an argument $\mathcal{A}$ undercuts an argument $\mathcal{B}$ (on a sub-argument $\mathcal{B}'$) if and only if the conclusion of $\mathcal{A}$ is defined as being contrary to the formula representing the topmost rule in $\mathcal{B}'$.

To complete the example, we expand the knowledge base:

$$\mathcal{K}(\mathcal{AT}) = \left\{ \begin{array}{l} object\_appears\_red, \\ illum\_by\_red\_light \\ \ulcorner object\_appears\_red \Rightarrow_{r1} object\_is\_red \urcorner, \\ \ulcorner illum\_by\_red\_light \in \overline{\ulcorner object\_appears\_red \Rightarrow_{r1} object\_is\_red \urcorner} \urcorner \end{array} \right\}$$

$$\mathcal{R}(\mathcal{AT}) = \{object\_appears\_red \Rightarrow_{r1} object\_is\_red\}$$

The leads to the arguments:

$\mathcal{A}_1$: *object_appears_red* $\qquad$ $\mathcal{A}_2$: *illum_by_red_light*

$\mathcal{A}_3$: $\ulcorner object\_appears\_red \Rightarrow_{r1} object\_is\_red \urcorner$

$\mathcal{A}_4$: $\ulcorner illum\_by\_red\_light \in \overline{\ulcorner object\_appears\_red \Rightarrow_{r1} object\_is\_red \urcorner} \urcorner$

$$\mathcal{A}_5: \frac{\mathcal{A}_1(object\_appears\_red)}{object\_is\_red} r_1$$

$\mathcal{A}_2$ is an argument for a contrary to r_1; thus $\mathcal{A}_2$ undercuts $\mathcal{A}_5$ on r_1.

Henceforth, unless made otherwise explicit, the phrase 'Argumentation Theory' and the notation $\mathcal{AT}$ will refer to modular Argumentation Theory. The symbol $\Pi_{\mathcal{L}}$ will be used to represent the set of all possible argumentation theories based on the language $\mathcal{L}$, while $\mathcal{A}rgs(\Pi_{\mathcal{L}}) = \bigcup_{\mathcal{AT} \in \Pi_{\mathcal{L}}} \mathcal{A}rgs(\mathcal{AT})$ is the set of all possible arguments in all possible theories based on $\mathcal{L}$.

3.2 Extended Argumentation Frameworks

A core feature of ASPIC$^+$ is the ability to derive a Dung-style abstract framework from an Argumentation Theory. This allows structured arguments and the relations between them to be evaluated using established acceptability semantics for abstract argumentation. An abstract framework derived from an Argumentation Theory uses the arguments in the theory as the set of arguments, and successful attacks (defeat) in place of the attack relation.

When using modular argumentation theories, an ordinary Dung-style framework cannot be derived because of the extra levels of attack afforded—for instance, attacks on attacks (through an attack on a preference or contrariness relation). Attacks on attacks in abstract argumentation are introduced in [33] through the use of Extended Argumentation Frameworks (EAFs); this, however, is still insufficient for our purposes because EAFs allow only attacks between arguments, and attacks on attacks between arguments. In modular argumentation theories, there is no limit on the depth to which attacks can go (e.g. given four arguments $\mathcal{A}_1,\ldots,\mathcal{A}_4$, $\mathcal{A}_1$ attacks that $\mathcal{A}_2$ attacks that $\mathcal{A}_3$ attacks $\mathcal{A}_4$). We therefore modify the definition of an EAF to allow for, in principle, unlimited depth of nesting of attack. With this modification, we also fold together the set of ordinary attacks (i.e. attacks between arguments) into the set of extended attacks (i.e. attacks on attacks).

DEFINITION 3.4 (Extended Argumentation Framework)
An *Extended Argumentation Framework* (EAF) is a tuple $\langle \mathcal{A}rgs, \mathcal{D} \rangle$ such that $\mathcal{A}rgs$ is a set of arguments, and:

- $\mathcal{D} = \mathcal{D}_0 \cup \mathcal{D}_1 \cup \ldots \cup \mathcal{D}_n$ with:
  - $D_0 \subseteq \mathcal{A}rgs \times \mathcal{A}rgs$
  - for $0 \leq i < n$, $\mathcal{D}_i \subseteq \mathcal{A}rgs \times \mathcal{D}_{i-1}$
- if $(X,(Y,Z)),(X',(Z,Y)) \in \mathcal{D}_1$ then $(X,X'),(X',X) \in \mathcal{D}$

Using this modified definition of EAFs, we now show how a structured EAF [34] is obtainable from a modular Argumentation Theory. Again, we require an updated definition to account for the extra levels of attack afforded in our modular argumentation theories.

Notice that preferences are not explicitly mentioned in the definition; this is because modular argumentation allows preferences to be expressed in terms of contraries (Section 3) and are thus accounted for within the attack definitions.

DEFINITION 3.5 (Structured EAF)
$EAF = \langle \mathcal{A}rgs, \mathcal{D} \rangle$ is a structured EAF, where $\mathcal{A}rgs = \mathcal{A}rgs(\mathcal{AT})$ and:

- If $\exists \mathcal{A}_1, \mathcal{A}_2 \in \mathcal{A}rgs$ s.t. $Conc(\mathcal{A}_1) \in cf(Conc(\mathcal{A}_2))$, $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{D}_0$ (rebutting and undermining)
- if $\exists \mathcal{A}_1, \mathcal{A}_2 \in \mathcal{A}rgs$ s.t $Conc(\mathcal{A}_2) = \lceil \varphi_1, \ldots, \varphi_n \Rightarrow_r \varphi \rceil$ and $Conc(\mathcal{A}_1) \in cf(Conc(\mathcal{A}_2))$, then $(\mathcal{A}_1, \mathcal{A}_2) \in \mathcal{D}_0$ and $\forall \mathcal{A}_3 \in \mathcal{A}rgs$ s.t. $r \in DefRules(\mathcal{A}_3)$, $(\mathcal{A}_1, \mathcal{A}_3) \in \mathcal{D}_0$ (undercutting)

- $\forall(\mathcal{A}_1,\mathcal{A}_2)\in\mathcal{D}_0$, $\forall\mathcal{A}_3\in\mathcal{A}rgs$ s.t. $\mathcal{A}_2\in Sub(\mathcal{A}_3)$, $(\mathcal{A}_1,\mathcal{A}_3)\in\mathcal{D}_0$ (sub-argument rebutting)
- $\forall\mathcal{A}_1,\mathcal{A}_2,\mathcal{A}_3\in\mathcal{A}rgs$, if $Conc(\mathcal{A}_1)\in\overline{\ulcorner Conc(\mathcal{A}_2)\in\overline{Conc(\mathcal{A}_3)}\urcorner}$, then given the function f:

$$f(W,X)=\begin{cases}(W,f(Y,Z)) & \text{if } Conc(X)=\ulcorner Conc(Y)\in\overline{Conc(Z)}\urcorner\\ (W,X) & \text{otherwise}\end{cases}$$

  - $(\mathcal{A}_1,(f(\mathcal{A}_2,\mathcal{A}_3)))\in\mathcal{D}$
  - if $Conc(\mathcal{A}_3)=\ulcorner\varphi_1,\ldots,\varphi_n\Rightarrow_r\varphi\urcorner$, then $\forall\mathcal{A}_5\in\mathcal{A}rgs$ s.t. $r\in DefRules(\mathcal{A}_5)$, $(\mathcal{A}_1,f(\mathcal{A}_2,\mathcal{A}_5))\in\mathcal{D}$
  - $\forall\mathcal{A}_4\in\mathcal{A}rgs$ s.t. $\mathcal{A}_3\in Sub(\mathcal{A}_4)$, then $(\mathcal{A}_1,(f(\mathcal{A}_2,\mathcal{A}_4)))\in\mathcal{D}$

  (Attacks on attacks)

It should be noted that when obtaining an Extended Argumentation Framework from a modular Argumentation Theory, then for every attack on an attack in D_i, $i>0$, there is a corresponding attack on an argument in D_0. This is because the modularity constraints require that every attack be represented by an argument. For instance, if there exist three arguments $\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{X}\in\mathcal{A}rgs$ such that $Conc(\mathcal{A})\in\overline{Conc(\mathcal{X})}$ and $Conc(\mathcal{X})=\ulcorner Conc(\mathcal{B})\in\overline{Conc(\mathcal{C}}\urcorner$ then, from Definition 3.5, $(\mathcal{A},(\mathcal{B},\mathcal{C}))\in\mathcal{D}_1$ and $(\mathcal{A},\mathcal{X})\in\mathcal{D}_0$.

In the ASPIC$^+$ framework, when an abstract framework is derived, it uses the notion of *defeat* in place of attack (with a defeat being an attack which is successful w.r.t. preferences). However, when using modular argumentation, preferences and attacks become arguments which can themselves be attacked. Our extension to EAFs, permitting attacks on attacks to any level, requires revised definitions of defeat and acceptability to that provided in [34]. We still use the notion of defeat being parameterized, but as in Definition 3.5 we need to account for the multiple layers of attack now afforded.

DEFINITION 3.6 (Defeat in EAFs)
An argument X defeats an argument Y with respect to a set of arguments S, expressed $X\rightarrow^S Y$, iff, if $\mathcal{A}tt=(X,Y)$:

- $\mathcal{A}tt\in\mathcal{D}_0$, and
- $\nexists Z\in S$ s.t. $g(Z,\mathcal{A}tt)\neq null$, where g is defined as:

$$g(Z,\mathcal{A}tt)=\begin{cases}(Z,\mathcal{A}tt) & \text{if } (Z,\mathcal{A}tt)\in\mathcal{D}\\ (Z,(V,g(W,\mathcal{A}tt))) & \text{if } \exists V,W\in\mathcal{A}rgs \text{ s.t. } g(W,\mathcal{A}tt)\neq null\\ & \ \text{and } (V,g(W,\mathcal{A}tt))\in\mathcal{D}\\ null & otherwise\end{cases}$$

The function g in Definition 3.6 determines whether or not the argument Z indirectly attacks the attack (X,Y). For example, in the attack relation $(\mathcal{A},(\mathcal{B},(\mathcal{C},(\mathcal{D},\mathcal{E}))))$, $\mathcal{A}$ indirectly attacks $(\mathcal{D},\mathcal{E})$ because it attacks $(\mathcal{B},(\mathcal{C},(\mathcal{D},\mathcal{E})))$. This is illustrated in Figure 1, where a double-headed arrow is used to distinguish an attack on an attack from an attack on an argument, a dashed arrow represents an unsuccessful (i.e. defeated) attack, and a solid arrow represents a successful (i.e. undefeated) attack.

Acceptability in EAFs differs from that found in a standard Dung AF. This definition is also adapted from [34] but again accounts for multiple layers of attack by using the function g to determine indirect attacks on attacks:

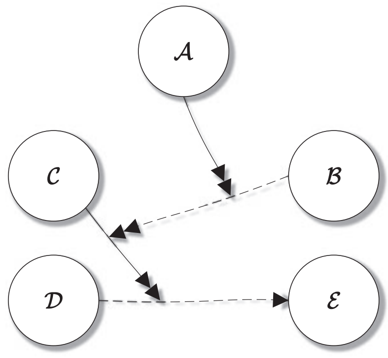

FIGURE 1. EAF showing an indirect attack by $\mathcal{A}$ on the attack $(\mathcal{D},\mathcal{E})$.

DEFINITION 3.7 (EAF acceptability)
Given an Extended Argumentation Framework $EAF=\langle \mathcal{A},\mathcal{C},\mathcal{D}\rangle$, let $S\subseteq\mathcal{A}$. Let $\mathcal{N}_S=\{X_1\rightarrow^S Y_1,\ldots,X_n\rightarrow^S Y_n\}$ where for $i=1\ldots n$, $X_i\in S$. Then $\mathcal{N}_S$ is a reinstatement set for $A\rightarrow^S B$ iff $A\rightarrow^S B\in\mathcal{N}_S$, and $\forall X\rightarrow^S Y\in\mathcal{N}_S, \forall Y'$ s.t. $g(Y',(X,Y))\in\mathcal{D}$, $\exists X'\rightarrow^S Y'\in\mathcal{N}_S$

X is acceptable w.r.t. $S\subseteq\mathcal{A}$ iff $\forall Y$ s.t. $Y\rightarrow^S X$, there is a reinstatement set for some $Z\rightarrow^S Y$.

Given definition 3.7, extensions are defined in the same way as for a standard Dung AF [22] (except, as noted in [34], '$X\rightarrow^S Y$' replaces $(X,Y)\in\mathcal{D}_0$ in the definition of stable extensions).

Henceforth, when using the notation $\mathcal{AF}_{\mathcal{AT}}$ we refer to an Extended Argumentation Framework derived from an Argumentation Theory $\mathcal{AT}$.

As well as updating the definitions of extended argumentation frameworks to account for modularity in Argumentation Theories, we also need to redefine what constitutes a conflict-free set of arguments.

DEFINITION 3.8 (Conflict-free set)
A set of arguments $\mathcal{S}$ is *conflict-free* iff $\nexists\mathcal{A},\mathcal{B},\mathcal{C}\in\mathcal{S}$ s.t. $Conc(\mathcal{A})=\ulcorner Conc(\mathcal{B})\in\overline{Conc(\mathcal{C})}\urcorner$

Definition 3.8 updates Dung's original definition in [22] such that a set of arguments is not conflict-free if the conclusion of any argument in the set introduces conflict between two other arguments in the set. Specifying this condition on a single level of attack is sufficient because all attacks are represented as arguments and so attacks on attacks (on attacks etc.) eventually take the form $Conc(\mathcal{A})\in\overline{Conc(\mathcal{B})}$, where potentially $Conc(\mathcal{B})=\ulcorner Conc(\mathcal{C})\in\overline{Conc(\mathcal{D})}\urcorner$, and potentially $Conc(\mathcal{D})=\ulcorner Conc(\mathcal{E})\in\overline{Conc(\mathcal{F})}\urcorner$ and so on.

4 A model for argument revision

4.1 Principles of Argument Revision

To translate AGM belief revision operators (contraction, expansion and revision) into ones suitable for Argument Revision, we need to consider the basic differences between the two representation formats to which the operators are applied (belief sets vs. a system of structured argumentation, ASPIC$^+$).

First, it is necessary to establish what constitutes belief in the ASPIC$^+$ framework. In [19], the arguments in the grounded extension of a Dung-style framework are used as the basis of agent beliefs.

In the present article, we extend this idea to structured arguments. Agent beliefs are the conclusions of all arguments in the grounded extension of the framework derived from an ASPIC$^+$ Argumentation Theory. Thus, in order to revise beliefs, it is possible to modify only the acceptability of arguments, without necessarily adding or removing the arguments themselves. Note that this does not preclude the addition or removal of arguments, but it does provide an additional method of revision, which may result in fewer changes (see Section 7 for further exploration of minimal change).

Secondly, consistency in the ASPIC$^+$ framework is different from that found in the belief sets used in the AGM theory. ASPIC$^+$ satisfies the rationality postulates of [12], provided certain conditions are met. It is through two of these rationality postulates that consistency is specified:

- **Direct consistency**—the set of conclusions of all arguments in an extension is consistent; and
- **Indirect consistency**—the closure of the set of conclusions of all arguments in an extension under strict-rule application is consistent

The conditions required for these postulates to hold are that (i) the preference ordering over arguments is reasonable [38, p. 15]; and (ii) that the Argumentation Theory is well-formed (Section 2, Definition 2.8).[2]

Well-formedness, however, depends entirely on the composition of the Argumentation Theory and does not have a pre-defined algorithm that guarantees it for all theories. During an Argument Revision process, it is possible that a theory which is not well-formed is yielded—consider in the simplest case adding to the knowledge base a formulae that is a contrary of the consequent of a strict rule. Thus, a second overall principle in Argument Revision is that the resultant Argumentation Theory be well-formed. This ensures that the set of conclusions in an extension remains consistent, and thus that the agent's beliefs do likewise.

4.2 Argument contraction and revision

Argument Revision is sub-divided into to main processes—*argument contraction* and *argument revision*. The aim of argument contraction is to ensure certain arguments are no longer acceptable (possibly through their removal). Similarly, the aim of argument revision is to ensure that certain arguments are acceptable (by first introducing them, if required).

Our aim in this article is to show how the AGM theory of belief revision can be adapted and applied to a system of structured argumentation when that system is being used as the basis for beliefs. To maintain this close connection to the AGM theory, we restrict our definitions such that acceptability is a binary distinction—an argument is acceptable or it is not (a formula is believed or it is not). In other words, we do not consider credulous vs. sceptical acceptability, nor do we apply labelings in the style of [11]. The operators we define can in principle, however, be applied to any set of arguments deemed to be 'acceptable' under some unspecified definition (e.g. taking one of the extensions, or the intersection of all extensions in a multi-extension semantics). For simplicity in presentation, we use 'extension' to refer to these sets of arguments, regardless if they are actually extensions, or instead some derivations thereof. In our examples, we assume that beliefs are derived from the acceptable arguments yielded by a single-extension semantics; henceforth the notation $\mathcal{A} \in E(\mathcal{AT})$ means there is an acceptable argument in the abstract framework derived from the Argumentation Theory $\mathcal{AT}$ under some unspecified, unique-extension semantics, subsumed by complete semantics.

[2]In the present work, we will always use either the last-link or weakest-link principles for establishing a preference ordering, both of which [38] proves to be reasonable, and thus that condition will always be satisfied.

As is the case with belief revision, we assume it is not possible to uniquely specify Argument Revision functions, because there may be multiple possible methods of achieving the goal [27]. This will be further explained in the definitions of the operators.

4.2.1 Argument contraction

Argument contraction is the process of modifying an Argumentation Theory to yield a new, well-formed and modular Argumentation Theory such that certain arguments that were acceptable in the original theory are not acceptable in the new theory.

DEFINITION 4.1 (Argument contraction)
Given an Argumentation Theory $\mathcal{AT}$ and a set of arguments $\mathcal{S}$, $\mathcal{AT}^-$ is a **contraction** of $\mathcal{AT}$ by $\mathcal{S}$ iff:

- $\mathcal{AT}^-$ is well-formed and modular
- $\forall \mathcal{A} \in \mathcal{S}$, $\mathcal{A} \notin E(\mathcal{AT}^-)$

Multiple methods of performing a contraction may exist, each yielding a different Argumentation Theory that satisfies the properties of the contraction. To capture all these theories, we define the *argument contraction set*:

DEFINITION 4.2 (Argument contraction set)
Given an Argumentation Theory $\mathcal{AT}$ and a set of arguments $\mathcal{S}$, $\mathcal{AT} \dot{-} \mathcal{S}$ is an **argument contraction set**, a maximal (w.r.t. set inclusion) set of argumentation theories where $\forall \mathcal{AT}^- \in \mathcal{AT} \dot{-} \mathcal{S}$, $\mathcal{AT}^-$ is a contraction of $\mathcal{AT}$ by $\mathcal{S}$ (per Definition 4.1).

Per Definition 4.1, argument contraction does not preclude the complete removal from $\mathcal{AT}$ of any $\mathcal{A} \in \mathcal{S}$ but it does not require it. In some situations, however, we do want to require the complete removal of arguments. For this, we define *simple argument contraction*, so-called because it it simpler to completely remove arguments (through removing premises and/or rules) than examining all possible ways of making them unacceptable.

DEFINITION 4.3 (Simple argument contraction)
Given an Argumentation Theory $\mathcal{AT}$ and a set of arguments $\mathcal{S}$, $\mathcal{AT}^-$ is a **simple contraction** of $\mathcal{AT}$ by $\mathcal{S}$ iff:

- $\mathcal{AT}^-$ is well-formed and modular
- $\forall \mathcal{A} \in \mathcal{S}$:
  - if $\mathcal{A} \in E(\mathcal{AT})$, $\mathcal{A} \notin \mathcal{A}rgs(\mathcal{AT}^-)$; or
  - if $\mathcal{A} \notin E(\mathcal{AT})$ and $\mathcal{A} \in \mathcal{A}rgs(\mathcal{AT})$, $\mathcal{A} \in \mathcal{A}rgs(\mathcal{AT}^-)$

The purpose of simple argument contraction is to completely remove *acceptable* arguments from the theory. Arguments that are already unacceptable should remain in the new theory.

It can be seen from Definitions 4.1 and 4.3 that simple argument contraction satisfies the definition of argument contraction in general, because if a set of arguments is not present in the theory, they also cannot be present in the extension; formally, if $\mathcal{S} \cap \mathcal{A}rgs(\mathcal{AT}) = \emptyset$ then $\mathcal{S} \cap E(\mathcal{AT}) = \emptyset$.

As with argument contraction, multiple methods of performing a simple argument contraction may exist. We therefore define the *simple argument contraction set*:

DEFINITION 4.4 (Simple argument contraction set)
Given an Argumentation Theory $\mathcal{AT}$ and a set of arguments $\mathcal{S}$, $\mathcal{AT} - \mathcal{S}$ is a **simple argument contraction set**, a maximal (w.r.t. set inclusion) set of argumentation theories where $\forall \mathcal{AT}^- \in \mathcal{AT} \dot{-} \mathcal{S}$, $\mathcal{AT}^-$ is a simple contraction of $\mathcal{AT}$ by $\mathcal{S}$ (per Definition 4.3).

The argumentation theories in an argument contraction set and simple argument contraction set satisfy the goal of the c

Each argumentation theory in an argument contraction set satisfies the goal of the contraction, i.e. the input arguments to the contraction function are not acceptable; similarly, each argumentation theory in a simple argument contraction set satisfies the goal of the simple contraction, i.e. acceptable input arguments to the simple contraction are not present. Both of these types of contraction have similarities to partial-meet contraction in belief revision, where the aim is to select sentences to be retained [40]. The difference, however, is that in partial-meet contraction, a new belief set is obtained by taking the intersection of all contracted sets. The theories in our argument contraction sets remain distinct, with one ultimately being chosen based on measures of minimal change (Section 7).

Different methods will be used to reach each Argumentation Theory in $\mathcal{AT} \dot{-} \mathcal{S}$; these possible methods will be described in further detail in Section 5.

4.2.2 Argument revision

Argument revision is the process of modifying an Argumentation Theory into a new, well formed and modular Argumentation Theory, such that a given set of arguments are both present and acceptable in the new theory.

DEFINITION 4.5 (Argument revision)
Given an Argumentation Theory $\mathcal{AT}$ and a set of arguments $\mathcal{S}$, $\mathcal{AT}^+$ is a **revision** of $\mathcal{AT}$ by $\mathcal{S}$ iff:

- $\mathcal{AT}^+$ is well-formed and modular
- $\forall \mathcal{A} \in \mathcal{S}, \mathcal{A} \in E(\mathcal{AT}^+)$

As with argument contraction, multiple methods of performing a revision may exist, each yielding a different Argumentation Theory that satisfies the properties of revision (i.e. the specified arguments are present and acceptable in it). To capture all these theories, we define an *argument revision set*.

DEFINITION 4.6 (Argument revision set)
Given an Argumentation Theory $\mathcal{AT}$ and a set of arguments $\mathcal{S}$, $\mathcal{AT} \dot{+} \mathcal{S}$ is an **argument revision set**, a maximal (w.r.t. set inclusion) set of argumentation theories where $\forall \mathcal{AT}^+ \in \mathcal{AT} \dot{+} \mathcal{S}$, $\mathcal{AT}^+$ is a revision of $\mathcal{AT}$ by $\mathcal{S}$ (per Definition 4.5).

For the sake of completeness, we provide also a definition of *argument expansion*, where the goal is to add arguments to a theory with no consideration for acceptability, with an associated *argument expansion set*.

DEFINITION 4.7 (Argument expansion)
Given an Argumentation Theory $\mathcal{AT}$ and a set of arguments $\mathcal{S}$, $\mathcal{AT}^+$ is an **expansion** of $\mathcal{AT}$ by $\mathcal{S}$ iff:

- $\mathcal{AT}^+$ is well-formed and modular

- $\forall \mathcal{A} \in \mathcal{S}$, $\mathcal{A} \in \mathcal{A}rgs(\mathcal{AT}^+)$

DEFINITION 4.8 (Argument expansion set)
Given an Argumentation Theory $\mathcal{AT}$ and a set of arguments $\mathcal{S}$, $\mathcal{AT}+\mathcal{S}$ is an **argument expansion set**, a maximal (w.r.t. set inclusion) set of argumentation theories where $\forall \mathcal{AT}^+ \in \mathcal{AT}+\mathcal{S}$, $\mathcal{AT}^+$ is an expansion of $\mathcal{AT}$ by $\mathcal{S}$ (per Definition 4.7).

5 Process and properties of argument revision

At the simplest level, revising an Argumentation Theory, whether it be through expansion, revision or contraction, involves making a change to that theory which in turn results in arguments being added or removed, or having their acceptability altered. For example, a contraction with respect to a set of arguments $\mathcal{S}$ may be carried out by adding new arguments that render all arguments in $\mathcal{S}$ unacceptable; similarly, a revision with respect to a set of arguments $\mathcal{S}$ may be carried out by removing arguments that defeat arguments in $\mathcal{S}$.

Recall from Section 2 that every argument in an ASPIC$^+$ Argumentation Theory is built on premises from the theory's knowledge base, and that those premises are themselves arguments. Recall also that our specification of modular argumentation (Section 3) requires that every component (rule, preference and contrary) in a modular argumentation theory be represented by an argument. Thus, every part of an argumentation theory rests upon premises taken from the knowledge base.

This allows us to express the process of argument revision as knowledge base modifications, achieved through expansion and simple contraction with respect to atomic arguments. This applies whether we are attempting to add new arguments, completely remove existing arguments, or change existing arguments' acceptability.

As an example, consider the following knowledge base and rules in an argumentation theory:

$$\mathcal{K}(\mathcal{AT}) = \{a, b, \lceil a, b \Rightarrow_{r_1} c \rceil, \lceil d \Rightarrow_{r_2} e \rceil, \lceil e \in \overline{c} \rceil\}$$
$$\mathcal{R}(\mathcal{AT}) = \{a, b \Rightarrow_{r_1} c, d \Rightarrow_{r_2} e\}$$

The arguments in $\mathcal{AT}$ are:

$\mathcal{A}_1$: a $\quad$ $\mathcal{A}_2$: b $\quad$ $\mathcal{A}_3$: $\lceil a, b \Rightarrow_{r_1} c \rceil$

$\mathcal{A}_4$: $\lceil d \Rightarrow_{r_2} e \rceil$ $\quad$ $\mathcal{A}_5$: $\lceil e \in \overline{c} \rceil$ $\quad$ $\mathcal{A}_6$: $\dfrac{\mathcal{A}_1(a), \mathcal{A}_2(b)}{c} r_1$

In the contraction $\mathcal{AT} \dot{-} \{\mathcal{A}_6\}$, we can either remove $\mathcal{A}_6$ completely, or render it unacceptable. For it to be removed completely, at least one of its premises and/or rules must no longer be in the theory. Performing one of the following simple contractions achieves this:

- $\mathcal{AT} - \{\mathcal{A}_1\}$ (removes the premise a from the theory)
- $\mathcal{AT} - \{\mathcal{A}_2\}$ (removes the premise b from the theory)
- $\mathcal{AT} - \{\mathcal{A}_3\}$ (removes the premise $\lceil a, b \Rightarrow_{r_1} c \rceil$ and the rule $a, b \Rightarrow_{r_1} c$ from the theory)

To change the acceptability of $\mathcal{A}_6$, we require an acceptable argument that successfully attacks it. Since e is a contrary of c, introducing an argument for e would achieve this change in acceptability. By introducing the premise d, an argument for e would be constructed using the rule r_2. Assume, therefore, there exists an atomic argument $\mathcal{A}_7$: d; we then perform the expansion:

- $\mathcal{AT}+\{\mathcal{A}_7\}$ (adds the premise d to the theory)

When performing an expansion or simple contraction with respect to an atomic argument, the result is a single theory whose knowledge base reflects the required change, along with associated effects required to maintain modularity. With this in mind, we simplify the notation when referring to such expansions and simple contractions. Given an atomic argument $\mathcal{A}$ with $Conc(\mathcal{A})=\phi$, then $\mathcal{AT}+\{\phi\}$ and $\mathcal{AT}-\{\phi\}$ are single Argumentation Theories such that $\mathcal{AT}+\{\mathcal{A}\}=\{\mathcal{AT}+\{\phi\}\}$ and $\mathcal{AT}-\{\mathcal{A}\}=\{\mathcal{AT}-\{\phi\}\}$.

When modifying the knowledge base through expansions and simple contractions, more than one modification might be required to achieve the goal of the change process as a whole, with each stage presenting a choice of possible modifications. To model these choices, we define a structure called a *change graph*.

5.1 Change graphs

When revising or contracting an Argumentation Theory, there may be multiple ways in which the goal can be achieved (i.e., arriving at an Argumentation Theory that satisfies the definition of revision or contraction). Furthermore, some ways of achieving the goal may have more than one step—for instance, if adding a formula brings about an Argumentation Theory that is not well-formed, further modifications will be required to restore well-formedness. There might then be a choice of modifications, each of which result in Argumentation Theories that require further modification, and so on.

This raises first the question of how to identify all possible ways of revising or contracting an Argumentation Theory; then, there is the question of which way should be chosen in a given case. In this section, the issue of identifying the possible ways will be addressed; choosing between them is then addressed in Section 7.

Recall that the revision or contraction of an Argumentation Theory is performed through modifying the knowledge bases in the argumentation systems within the theory by expanding and/or contracting with respect to atomic arguments for knowledge base formulae. Modifications then continue to take place until the properties of the chosen change (i.e. revision or contraction) are satisfied. One such sequence of modifications can be seen as a route, or path, from the original Argumentation Theory to an Argumentation Theory that has been revised or contracted with respect to the input arguments.

To model the possible ways in which an Argumentation Theory can be either revised or contracted (with respect to a set of input arguments), we define a structure called a *change graph*. A change graph can be used for either revision or contraction. For convenience, we use the notation $\mathcal{AT}^{\pm}$ to refer to an Argumentation Theory that has been either revised or contracted, and Θ to represent a set of argumentation theories $\{\mathcal{AT}^{\pm}_1,\ldots,\mathcal{AT}^{\pm}_n\}$ that are modified versions of $\mathcal{AT}$, but do not necessarily satisfy the criteria of well-formedness and modularity for a specific revision or contraction.

Definition 5.1

A change graph $CG(\mathcal{AT},\Theta)$ for the modification of $\mathcal{AT}$ to a set of argumentation theories $\Theta=\{\mathcal{AT}^{\pm}_1,\ldots,\mathcal{AT}^{\pm}_n\}$ is a directed acyclic graph $\langle\Upsilon,\Omega\rangle$ where:

(1) $\Upsilon\subseteq\Pi_{\mathcal{L}}$ (for $\Pi_{\mathcal{L}}$, the set of all possible Argumentation Theories);
(2) $\Theta\subseteq\Upsilon$ is minimal (w.r.t. set inclusion) in that $\forall\mathcal{AT}'\in\Theta$, $\neg\exists\mathcal{AT}''\in\Theta$, $\mathcal{K}(\mathcal{AT})\ominus\mathcal{K}(\mathcal{AT}'')\subset\mathcal{K}(\mathcal{AT})\ominus\mathcal{K}(\mathcal{AT}')$;

(3) **(Atomic change)** $\Omega \subseteq \Upsilon \times \Upsilon$ where $\forall \omega \in \Omega$ such that $\omega = (\mathcal{AT}', \mathcal{AT}'')$, we have that $|\mathcal{K}(\mathcal{AT}') \ominus \mathcal{K}(\mathcal{AT}'')| = 1$

where $\ominus$ represents symmetric difference.[3]

The minimality constraint on Θ in Definition 5.1(2) is a pruning exercise that discards any argumentation theories arrived at by making unnecessary changes by considering the set of net changes between $\mathcal{AT}$ and some $\mathcal{AT}' \in \Theta$: there can exist no Argumentation Theory $\mathcal{AT}'' \in \Theta$ where the set of net changes in $\mathcal{AT}''$ is a proper subset of the set of net changes in $\mathcal{AT}'$. The atomic change condition mandates that on each edge of a change graph, we expand or simple-contract with respect to only a single atomic argument.

If any $\mathcal{AT}' \in \Theta$ satisfies the definition of $\mathcal{AT} \dot{-} \mathcal{S}$, then $\mathcal{AT}'$ is a contraction of $\mathcal{AT}$ by $\mathcal{S}$ and hence $\mathcal{AT}' \in \mathcal{AT} \dot{-} \mathcal{S}$. Similarly, if any $\mathcal{AT}' \in \Theta$ satisfies the definition of $\mathcal{AT} \dot{+} \mathcal{S}$ then $\mathcal{AT}'$ is a revision of $\mathcal{AT}$ by $\mathcal{S}$ and hence $\mathcal{AT}' \in \mathcal{AT} \dot{+} \mathcal{S}$.

5.2 Constructing change graphs

A change graph is constructed recursively. From the original theory and set of arguments, a new set of theories is arrived at which are expansions or simple contractions of the original theory. Each of these new theories is then expanded or simple-contracted, each yielding a new set of theories, and so on until each sub-graph arrives at a set of theories that satisfy the original revision or contraction.

The goal of a contraction is achieved by either completely removing the input arguments or making them unacceptable (or a combination thereof). To make them unacceptable, it is necessary to introduce or make acceptable defeaters of those arguments. Similarly, to achieve the goal of a revision all the input arguments must be acceptable, which means all defeaters of those arguments must be removed or made unacceptable.

To make an argument unacceptable we need to introduce at least one source of conflict with that argument. Multiple sources of conflict may exist (i.e. an argument may have more than one defeater) but only one is required. Conversely, to make an argument acceptable we need to remove all sources of conflict with that argument. We do not, however, need to remove all aspects of the conflict (i.e. we remove only an argument or an attack but not both).

To illustrate, consider the top argumentation framework in Figure 2 (where, for simplicity, attacks have been labelled in lieu of explicitly showing the arguments that represent them). The argument $\mathcal{A}_3$ is currently unacceptable because it is defeated by $\mathcal{A}_1$ (via $\mathcal{A}tt_1$) and $\mathcal{A}_2$ (via $\mathcal{A}tt_2$). If we were to remove either $\mathcal{A}_1$ or $\mathcal{A}tt_1$, $\mathcal{A}_3$ would still be defeated by $\mathcal{A}_2$ ($\mathcal{A}tt_2$)—in other words, a single source of conflict renders it unacceptable. This is shown to the bottom-left of Figure 2, where $\mathcal{A}tt_1$ has been removed. If we were to then remove either $\mathcal{A}_2$ or $\mathcal{A}tt_2$, $\mathcal{A}_3$ would be acceptable—in other words, removing all sources of conflict renders it acceptable. This is shown to the bottom-right of Figure 2 where $\mathcal{A}tt_2$ has been removed as well as $\mathcal{A}tt_1$.

To make a set of arguments unacceptable, we need introduce one source of conflict for every argument; where multiple sources of conflict exist for each arguments, we need to account for every possible combination. To do this, we define the *contraction disruption sets*. A contraction disruption set for a set of arguments $\mathcal{S}$ is a set of arguments $\mathcal{S}'$ such that if all arguments in $\mathcal{S}'$ are acceptable, no arguments in $\mathcal{S}$ are acceptable.

[3]Given two sets, symmetric difference is the set of elements that are in either set, but not in both; formally, $\mathcal{S}_1 \ominus \mathcal{S}_2 = (\mathcal{S}_1 \cup \mathcal{S}_2) \setminus (\mathcal{S}_1 \cap \mathcal{S}_2)$.

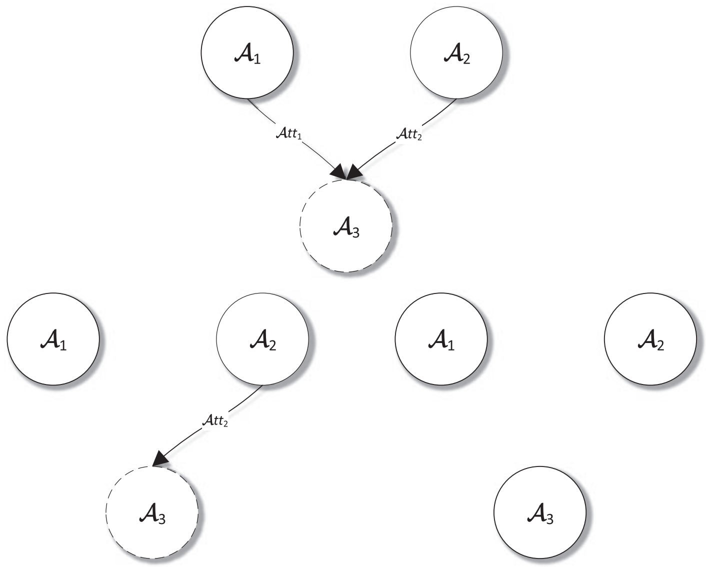

FIGURE 2. Example frameworks to illustrate making $\mathcal{A}_3$ acceptable.

DEFINITION 5.2 (Contraction Disruption Sets)
For a set of arguments $\mathcal{S}$, $DS^{\dot{-}}(\mathcal{S})$ are the *contraction disruption sets* for $\mathcal{S}$. If $\exists \mathcal{AT}$ s.t. $\mathcal{A}rgs(\mathcal{AT}) = \mathcal{S} \cup \mathcal{S}'$, $E(\mathcal{AT}) = \mathcal{S}'$ and $\nexists \mathcal{S}'' \subset \mathcal{S}'$ s.t. $\mathcal{S}'' \in DS^{\dot{-}}(\mathcal{S})$, then $\mathcal{S}' \in DS^{\dot{-}}(\mathcal{S})$.

The third constraint in Definition 5.2 ensures that every element of a contraction disruption set is relevant in causing the disruption.

Note that because $\mathcal{AT}$ is modular, the contraction disruption sets must explicitly contain representations of the attacks that cause the disruption. For instance, given a set of arguments $\mathcal{S} = \{\mathcal{A}, \mathcal{B}\}$, the set $\{\mathcal{C}, \lceil Conc(\mathcal{C}) \in \overline{Conc(\mathcal{A})}\rceil, \mathcal{D}, \lceil Conc(\mathcal{D}) \in \overline{Conc(\mathcal{B})}\rceil\}$ is a contraction disruption set for $\mathcal{S}$, but $\{\mathcal{C}, \mathcal{D}\}$ is not.

To make a set of arguments acceptable, we need to eliminate all sources of conflict with all arguments in the set. Similar to contraction, when each argument in the set has multiple ways of removing the sources of conflict we need to account for every possible combination.

Removals for a revision can be identified using the contraction disruption sets; for a single argument $\mathcal{A}$, removing one element from every set in $DS^{\dot{-}}(\{A\})$ will render $\mathcal{A}$ acceptable. Consider again the example in Figure 2; the contraction disruption sets for $\mathcal{A}_3$ are $DS^{\dot{-}}(\{A_3\}) = \{\{\mathcal{A}_1, \mathcal{A}tt_1\}, \{\mathcal{A}_2, \mathcal{A}tt_3\}\}$. Removing $\mathcal{A}_1$ and $\mathcal{A}_2$; $\mathcal{A}_1$ and $\mathcal{A}tt_2$; $\mathcal{A}tt_1$ and $\mathcal{A}_2$ or $\mathcal{A}tt_1$ and $\mathcal{A}tt_2$ (i.e. each pair with one element from each contraction disruption set) will render $\mathcal{A}_3$ acceptable.

To capture this formally, we define the *revision disruption sets*. Revision disruption sets are defined in two parts, first for a single argument then for a set of arguments. A revision disruption set for a single argument $\mathcal{A}$ contains one and only one argument from each contraction disruption set for $\{\mathcal{A}\}$. A revision disruption set for a set of arguments $\mathcal{S}$ is a superset of one disruption set for each argument in $\mathcal{S}$.

DEFINITION 5.3 (Revision Disruption Sets)

(1) For an argument $\mathcal{A}$ $DS^{\dot{+}}(\mathcal{A})$ are the revision disruption sets for $\mathcal{A}$.

Let $DS^{\dot{-}}(\{\mathcal{A}\}) = \{\pi_1, \ldots, \pi_n\}$ and $\pi = \bigcup_{i=1}^{n} \pi_i$.

$DS^{\dot{+}}(\mathcal{A}) = \{P : P \subseteq \pi,\ |P \cap \pi_i| = 1 \text{ for } 1 \leq i \leq n\}$

(2) For a set of arguments $\mathcal{S} = \{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$, $DS^{\dot{+}}(\mathcal{S})$ are the revision disruption sets for $\mathcal{S}$.

$DS^{\dot{+}}(\mathcal{S}) = \{\bigcup_{i=1}^{t} \pi_i : (\pi_1, \ldots, \pi_t) \in DS^{\dot{+}}(\mathcal{A}_1) \times \ldots \times DS^{\dot{+}}(\mathcal{A}_n)\}$

Using the contraction and revision disruption sets, we determine the vertices in a change graph using slightly different functions for contraction and revision. For a contraction $\mathcal{AT} \dot{-} \mathcal{S}$, the function $F^{\dot{-}}$ is:

$$F^{\dot{-}}(\mathcal{AT}, \mathcal{S}) =$$

$$\begin{cases} \emptyset & \text{if } \mathcal{S} \cap E(\mathcal{AT}) = \emptyset \\ \displaystyle\bigcup_{\varphi \in Prem(\mathcal{S})} F^{\dot{-}}(\mathcal{AT} - \{\varphi\}, \mathcal{S} \cap Args(\mathcal{AT} - \{\varphi\})) \cup \bigcup_{\mathcal{S}' \in DS^{\dot{-}}(\mathcal{S})} F^{\dot{+}}(\mathcal{AT}, \mathcal{S}') & \textit{otherwise} \end{cases}$$

The first part of $F^{\dot{-}}$ takes each premise (φ) from $\mathcal{S}$ and simple-contracts with respect to that premise ($\mathcal{AT} - \{\varphi\}$). The resultant theory is then used in a recursive call to $F^{\dot{-}}$ to simple-contract with respect to the premises of arguments in $\mathcal{S}$ that remain in the theory. When this recursion terminates, we arrive at all theories that are simple-contractions of $\mathcal{AT}$ by $\mathcal{S}$. The second part of the function takes into account the revision of $\mathcal{AT}$ by every contraction disruption set for $\mathcal{S}$. This step arrives at all theories that are non-simple contractions of $\mathcal{AT}$ by $\mathcal{S}$ (i.e. theories in which some arguments in $\mathcal{S}$ remain but are unacceptable).

The vertices in a change graph for a contraction $\mathcal{AT} \dot{-} \mathcal{S}$ are $\Upsilon = F^{\dot{-}}(\mathcal{AT}, \mathcal{S}) \cup \{\mathcal{AT}\}$.

For a revision, the function $F^{\dot{+}}$ is:

$$F^{\dot{+}}(\mathcal{AT}, \mathcal{S}) =$$

$$\begin{cases} \emptyset & \text{if } \mathcal{S} \subseteq E(\mathcal{AT}) \\ \displaystyle\bigcup_{\varphi \in Prem(\mathcal{S})} F^{\dot{+}}(\mathcal{AT} + \{\varphi\}, \mathcal{S} \setminus Args(\mathcal{AT} + \{\varphi\})) \cup \bigcup_{\mathcal{S}' \in DS^{\dot{+}}(\mathcal{S})} F^{\dot{-}}(\mathcal{AT}, \mathcal{S}') & \textit{otherwise} \end{cases}$$

The first part of $F^{\dot{+}}$ takes each premise (φ) from $\mathcal{S}$ and expands with respect to that premise ($\mathcal{AT} + \{\varphi\}$). The resultant theory is then used in a recursive call to $F^{\dot{+}}$ to expand with respect to the premises of arguments in $\mathcal{S}$ that have not yet been added to the theory. When this recursion terminates, we arrive at all theories that are expansions of $\mathcal{AT}$ by $\mathcal{S}$. The second part of the function attempts to make the arguments in $\mathcal{S}$ acceptable by contracting with respect to each revision disruption set for $\mathcal{S}$. This step arrives at all theories that are revisions of $\mathcal{AT}$ by $\mathcal{S}$.

The vertices in a change graph for a revision $\mathcal{AT} \dot{+} \mathcal{S}$ are $\Upsilon = F^{\dot{+}}(\mathcal{AT}, \mathcal{S}) \cup \{\mathcal{AT}\}$.

Each edge in a change graph connects an Argumentation Theory to a new theory that is an expansion or simple contraction of that theory with respect to a single atomic argument. This allows

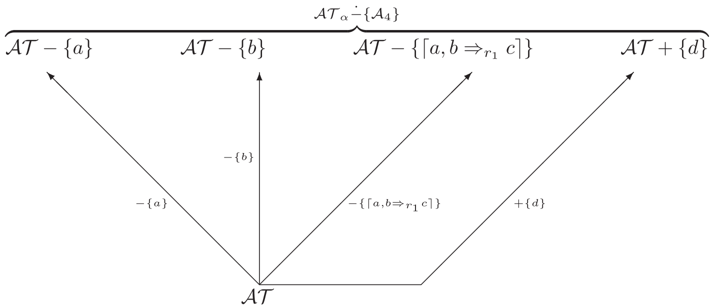

FIGURE 3. Change graph for $\mathcal{AT}_\alpha - \{\mathcal{A}_4\}$.

us to easily determine the edges (Ω) based on the vertices, thus:

$$\forall \mathcal{AT}' \in \Upsilon, \text{ if } (\mathcal{AT}' \pm \{\phi\}) \in \Upsilon \text{ (for any } \phi \in \mathcal{L}\text{), then } (\mathcal{AT}', \mathcal{AT}' \pm \{\phi\}) \in \Omega$$

where $\pm$ denotes either simple contraction or expansion.

To illustrate the how the functions $F^{\dot{-}}$ and $F^{\dot{+}}$ construct change graphs, we provide the following example. Consider the arguments in an Argumentation Theory $\mathcal{AT}$:

$\mathcal{A}_1$: a $\mathcal{A}_2$: b $\mathcal{A}_3$: $\lceil a,b \Rightarrow_{r_1} c\rceil$

$\mathcal{A}_4$: $\dfrac{\mathcal{A}_1(a), \mathcal{A}_2(b)}{c} r_1$ $\mathcal{A}_5$: $\lceil d \in \overline{c}\rceil$

And that there exists an argument $\mathcal{A}_6$ not in $\mathcal{AT}$:

$\mathcal{A}_6$: d

The change graph for the contraction $\mathcal{AT} \dot{-} \{\mathcal{A}_4\}$ is constructed as follows:

$$\begin{aligned}
F^{\dot{-}}(\mathcal{AT}, \{\mathcal{A}_4\}) \quad &= F^{\dot{-}}(\mathcal{AT}-\{a\}, \emptyset) \cup F^{\dot{-}}(\mathcal{AT}-\{b\}, \emptyset) \cup F^{\dot{-}}(\mathcal{AT}-\{\lceil a \Rightarrow_{r_1} b\rceil\}, \emptyset) \\
&\qquad \cup F^{\dot{+}}(\mathcal{AT}, \{\mathcal{A}_6\}) \\
&= \{\mathcal{AT}-\{a\}\} \cup \{\mathcal{AT}-\{b\}\} \cup \{\mathcal{AT}-\{\lceil a \Rightarrow_{r_1} b\rceil\}\} \\
&\qquad \cup \{\mathcal{AT}+\{d\}\} \\
&= \{\mathcal{AT}-\{a\}, \mathcal{AT}-\{b\}, \mathcal{AT}-\{\lceil a \Rightarrow_{r_1} b\rceil\}, \mathcal{AT}+\{d\}\}
\end{aligned}$$

The change graph is therefore $CG(\mathcal{AT}, \Theta) = \langle \Upsilon, \Omega \rangle$ where:

- $\Theta = \{\mathcal{AT}-\{a\}, \mathcal{AT}-\{b\}, \mathcal{AT}-\{\lceil a \Rightarrow_{r_1} b\rceil, \mathcal{AT}+\{d\}\}\}$
- $\Upsilon = \{\mathcal{AT}, \mathcal{AT}-\{a\}, \mathcal{AT}-\{b\}, \mathcal{AT}-\{\lceil a \Rightarrow_{r_1} b\rceil\}, \mathcal{AT}+\{d\}\}$
- $\Omega = \{(\mathcal{AT}, \mathcal{AT}-\{a\}), (\mathcal{AT}, \mathcal{AT}-\{b\}), (\mathcal{AT}, \mathcal{AT}-\{\lceil a,b \Rightarrow_{r_1} c\rceil\}), (\mathcal{AT}, \mathcal{AT}+\{d\})\}$

This change graph is rendered visually in Figure 3.

TABLE 3. Basic AGM postulates for revision

$(K\dot{+}1)$	The outputs of the revision function are belief sets.	$K\dot{+}\Phi$ is a belief set
$(K\dot{+}2)$	The input sentence Φ is accepted in $K\dot{+}\Phi$.	$\Phi \in K\dot{+}\Phi$
$(K\dot{+}3)$	Defined together and	$K\dot{+}\Phi \subseteq K+\Phi$
$(K\dot{+}4)$	cover the case $\neg\Phi \notin K$.	If$\neg\Phi \notin K$, then $K+\Phi \subseteq K\dot{+}\Phi$
$(K\dot{+}5)$	$K\dot{+}\Phi$ is consistent.	$K\dot{+}\Phi = K_\perp$ iff $\vdash \neg\Phi$
$(K\dot{+}6)$	Revisions are based on the *knowledge* level and not the *analytic*.	If $\vdash \Phi \leftrightarrow \Psi$, then $K\dot{+}\Phi = K\dot{+}\Psi$

TABLE 4. Basic AGM postulates for contraction

$(K\dot{-}1)$	The outputs of the revision function are belief sets.	$K\dot{-}\Phi$ is a belief set
$(K\dot{-}2)$	No new beliefs occur in $K\dot{-}\Phi$.	$K\dot{-}\Phi \subseteq K$
$(K\dot{-}3)$	When $\Phi \notin K$, nothing is retracted.	If $\Phi \notin K$, $K\dot{-}\Phi = K$
$(K\dot{-}4)$	The belief set $K\dot{-}\Phi$ does not entail Φ, unless Φ is logically valid.	If not $\vdash \Phi$, then $\Phi \notin K\dot{-}\Phi$
$(K\dot{-}5)$	The *recovery postulate*; contractions can be 'undone'.	If $\Phi \in K$, then $K \subseteq (K\dot{-}\Phi)+\Phi$
$(K\dot{-}6)$	Contractions are based on the *knowledge* level and not the *analytic*.	If $\vdash \Phi \leftrightarrow \Psi$, then $K\dot{-}\Phi = K\dot{-}\Psi$

6 Postulates for Argument Revision

Central to the AGM theory of belief revision is a set of postulates for rational revision and expansion functions defined over belief sets. The primary motivation behind the AGM postulates is that when changing beliefs, we want to retain as much information as possible from our old beliefs, i.e. a minimal change should be made [28].

In this section, we first summarize the AGM postulates then for five of the six basic postulates for each belief revision operation, go on to propose similar postulates for Argument Revision. We do not examine the sixth postulate because specifying a bijection between arguments that defines equivalence without equality requires a level of analysis beyond the scope of the present work. We also propose additional postulates for rational argument revisions and contractions that account for new phenomena intrinsic to Argument Revision.

There are several key difference between the AGM belief revision operators and our Argument Revision operators. The first is the underlying representation to which the operators are applied: in belief revision, it is flat (closed) belief sets; in Argument Revision, it is a system of structured argumentation. Second, the belief revision operators yield a single revised belief set that satisfies the goal; in Argument Revision, the operators yield multiple revised Argumentation Theories that each satisfy the goal. Nevertheless, it is possible to capture the spirit of most of the AGM postulates for Belief Revision in analogous versions for Argument Revision.

6.1 AGM postulates for belief revision

There are six basic postulates each for revision and contraction. These were provided in Section 2.2, and are rehearsed in Tables 3 and 4 respectively:

6.2 Postulates for Argument Revision

6.2.1 $K\dot{+}1$

In the AGM model of belief revision, this postulate states that the output of the revision function is a belief set. This is easily adapted for Argument Revision: every member of the set yielded from a revision is an Argumentation Theory:

$(\mathcal{AT}\dot{+}1)$ $\forall \mathcal{AT}^+ \in (\mathcal{AT}\dot{+}\mathcal{S})$, $\mathcal{AT}^+$ is an Argumentation Theory.

6.2.2 $K\dot{+}2$

This AGM postulate guarantees that the input sentence is accepted in the revised belief set. In argument revision, we guarantee that the input arguments are acceptable:

$(\mathcal{AT}\dot{+}2)$ $\forall \mathcal{AT}^+ \in (\mathcal{AT}\dot{+}\mathcal{S}), \mathcal{S} \subseteq E(\mathcal{AT}^+)$

Since for any Argumentation Theory $\mathcal{AT}$, $E(\mathcal{AT}) \subseteq \mathcal{A}rgs(\mathcal{AT})$, it follows that:

$$\forall \mathcal{AT}^+ \in (\mathcal{AT}\dot{+}\mathcal{S}),\ \mathcal{S} \subseteq \mathcal{A}rgs(\mathcal{AT}^+) \tag{1}$$

6.2.3 $K\dot{+}3\backslash 4$

In Belief Revision, the third and fourth postulates are used to describe the case $K+\Phi$ when $\neg\Phi \notin K$.

The third postulate states that a revision of K by Φ is a subset of the corresponding expansion. No analogue exists for Argument Revision because it is possible that a revision can be performed by introducing new arguments that defend the arguments to be revised.

The fourth postulate states that an expansion of K by Φ is a subset of the corresponding revision when $\neg\Phi \notin K$. When combined with the third postulate, this leads to mutual subset and thus equality.

In Argument Revision, we capture the spirit of these two postulates in a single postulate expressed solely in terms of equality. If the union of the extension in $\mathcal{AT}$ with all sub-arguments in $\mathcal{S}$ is conflict-free, then the result of a revision with respect to $\mathcal{S}$ is identical to the result of an expansion with respect to $\mathcal{S}$:

$(\mathcal{AT}\dot{+}3\backslash 4)$ If $E(\mathcal{AT}) \cup Sub(\mathcal{S})$ is conflict-free, then $(\mathcal{AT}\dot{+}\mathcal{S}) = (\mathcal{AT}+\mathcal{S})$

6.2.4 $K\dot{+}5$

The fifth postulate for revision in the AGM model of belief revision handles consistency. In a system of argumentation such as ASPIC$^+$, consistency is handled by the internal argument machinery. Nevertheless, we specify an analogous postulate in terms of well-formedness; an Argumentation Theory that is well-formed satisfies the rationality postulates of [13] which, among other things (see Section 2), ensures the set of conclusions in an extension is consistent (both directly, and when closed under strict rule application). To ensure that revised theories can themselves be revised, we also require that they be modular.

$(\mathcal{AT}\dot{+}5)$ $\forall \mathcal{AT}^+ \in (\mathcal{AT}\dot{+}\mathcal{S})$, $\mathcal{AT}^+$ is well-formed and modular

6.3 Postulates for argument contraction

6.3.1 $K\dot{-}1$

Similar to the first postulate for revision in the AGM model, the first postulate for expansion states that the output of a contraction is a belief set. We adapt this postulate in a similar way as for

Argument Expansion:

$(\mathcal{AT}\dot{-}1) \qquad \forall \mathcal{AT}^- \in (\mathcal{AT}\dot{-}\mathcal{S}),\ \mathcal{AT}^-$ is an Argumentation Theory

6.3.2 $K\dot{-}2$

The second postulate for Contraction in the AGM model stipulates that, because a contracted belief set is formed by giving up some beliefs, no new beliefs should occur in that contracted belief set. It is not possible to specify an analogous postulate for argument contraction because an argument can be contracted by making it unacceptable, which in turn can be achieved by adding new arguments to the theory.

We can, however, specify the equivalent postulate for simple argument contraction, where the goal is to completely remove arguments from the theory:

$(\mathcal{AT}-2) \qquad \forall \mathcal{AT}^- \in (\mathcal{AT}-\mathcal{S}),\ \mathcal{A}rgs(\mathcal{AT}^-) \subseteq \mathcal{A}rgs(\mathcal{AT})$

Imposing this condition on simple contraction further enforces the principle of minimal change. The only way to completely remove an argument from an Argumentation Theory is to remove premises and/or rules; introducing new arguments as part of this process does not achieve anything towards that goal.

From $(\mathcal{AT}-2)$ it follows that:

$$\forall \mathcal{AT}^+ \in (\mathcal{AT}-\mathcal{S}),\ E(\mathcal{AT}^-) \subseteq E(\mathcal{AT}) \tag{2}$$

Note that although (2) is specified in terms of the extension, it is a consequence of simple contraction and therefore does not apply to contraction in general.

6.3.3 $K\dot{-}3$

The third postulate for contraction in the AGM model requires that nothing be retracted from K when the input formula is not present (i.e. $K\dot{-}\Phi = K$). This is easily repurposed for Argument Contraction:

$(\mathcal{AT}\dot{-}3) \qquad$ If $\mathcal{S} \cap E(\mathcal{AT}) = \emptyset,\ (\mathcal{AT}\dot{-}\mathcal{S}) = \{\mathcal{AT}\}$

From $(\mathcal{AT}\dot{-}3)$, it follows that:

$$\text{If } \mathcal{S} \cap \mathcal{A}rgs(\mathcal{AT}) = \emptyset,\ (\mathcal{AT}\dot{-}\mathcal{S}) = \{\mathcal{AT}\} \tag{3}$$

6.3.4 $K\dot{-}4$

The fourth postulate for contraction in the AGM model ensures that the sentence to be contracted is not a logical consequence of the beliefs retained in the contracted belief set. This too is easily adapted for argument contraction:

$(\mathcal{AT}\dot{-}4) \qquad \forall \mathcal{AT}^- \in (\mathcal{AT}\dot{-}\mathcal{S}),\ \mathcal{S} \cap E(\mathcal{AT}^+) = \emptyset$

From ($\mathcal{AT}\dot{-}4$) it follows that:

$$\forall \mathcal{AT}^- \in (\mathcal{AT} - \mathcal{S}),\ \mathcal{S} \cap E(\mathcal{AT}^+) = \emptyset \tag{4}$$

6.3.5 $K\dot{-}5$

The fifth postulate for contraction in the AGM model allows contractions to be 'undone' (also called the 'recovery postulate'). In specifying analogous postulates for Argument Revision, we weaken the constraint slightly such that if a contracted Argumentation Theory is subsequently expanded by the same set of arguments, the original theory is one of the results of that expansion:

($\mathcal{AT}\dot{-}5$) $\quad \forall \mathcal{AT}^- \in (\mathcal{AT}\dot{-}\mathcal{S}), \mathcal{AT} \in (\mathcal{AT}^- \dot{+} \mathcal{S})$

Note that $\mathcal{AT}$ is not the only potential theory, but could be one of many. This is thanks in part to the modularity of the Argumentation Theories, whereby an argument could be reintroduced by using a different argument for one of its rules.

6.4 The Levi Identity

In the AGM model of belief revision, the *Levi Identity* (termed (*Def* $\dot{+}$) in [28]) expresses revision in terms of contraction and expansion, thus: $K\dot{+}\Phi = (K\dot{-}\neg\Phi) + \Phi$. In other words, a revision of K by Φ can be achieved first by contracting K with respect to $\neg\Phi$, then expanding the resulting belief set by Φ.

We formulate a similar identity in terms of argument revision, with some technical adjustments to account for (i) our revision and contraction operators yielding sets of theories; and (ii) structured argumentation's increased sophistication over belief sets.

THEOREM 6.1
(*Def* $\dot{+}$)

$$\mathcal{AT}\dot{+}\mathcal{S} = \bigcup_{\mathcal{S}' \in DS^{\dot{+}}(\mathcal{S})} \bigcup_{\mathcal{AT}^- \in (\mathcal{AT}\dot{-}\mathcal{S}')} (\mathcal{AT}^- + \mathcal{S})$$

We take each revision disruption set $\mathcal{S}'$ for $\mathcal{S}$ and make the arguments in $\mathcal{S}'$ unacceptable by contracting $\mathcal{AT}$. For each contraction, we then expand with respect to $\mathcal{S}$.

6.5 The Harper Identity

In the AGM model of belief revision, the *Harper Identity* (termed (*Def* $\dot{-}$) in [28]) expresses contraction in terms of revision, thus $K\dot{-}\Phi = K \cap K\dot{+}\neg\Phi$. In other words, the contraction of K by Φ is the same as the intersection of K with a revision of K by $\neg\Phi$. This identity holds because in revising K by $\neg\Phi$, Φ (and any beliefs from which it is a logical consequence) are removed (i.e. contracted) in order to maintain consistency. The intersect with K is required to satisfy the postulate ($K\dot{-}2$) (that no new beliefs occur in a contracted belief set—i.e. we need to ignore $\neg\Phi$).

Formulating a directly analogous identity for argument contraction (in terms of argument revision) is more complicated than for the Levi identity. First, a contraction can be achieved by either making arguments unacceptable or completely removing them from the theory. This means that, unlike in the AGM model of belief revision, we can have new beliefs (arguments) following a contraction in order to bring about acceptability change. Second, it is possible to use both strategies in a single contraction—consider, for instance, the contraction $\mathcal{AT}\dot{-}\{\mathcal{A},\mathcal{B}\}$. Completely removing $\mathcal{A}$ and making $\mathcal{B}$ unacceptable (or vice versa) satisfies the contraction.

THEOREM 6.2
(*Def* $\dot{-}$)

$$\mathcal{AT}\dot{-}\mathcal{S} = \bigcup_{\mathcal{S}'\in 2^{\mathcal{S}}} \bigcup_{\mathcal{S}_D\in DS^{\dot{-}}(\mathcal{S}')} \bigcup_{\mathcal{AT}^{+}\in(\mathcal{AT}\dot{+}\mathcal{S}_D)} (\mathcal{AT}^{+}-\mathcal{S})$$

We take each subset $\mathcal{S}'$ of $\mathcal{S}$ and attempt to make each argument in $\mathcal{S}'$ unacceptable by revising $\mathcal{AT}$ with respect to a disruption set. Each theory resulting from that revision is then simple-contracted by $\mathcal{S}$ to completely remove the remaining arguments in $\mathcal{S}$ (i.e. $\mathcal{S}\backslash\mathcal{S}'$). The Definition of simple contraction (Definition 4.3) means that any arguments in $\mathcal{S}'$ already rendered unacceptable (in the revision with respect to a disruption set) will not be removed.

Using subsets of $\mathcal{S}$ means that all combinations of (defeat-based) contraction and simple contraction are accounted for. This includes all arguments being simple-contracted, because $\emptyset\in 2^{\mathcal{S}}$, $DS^{\dot{-}}(\emptyset)=\{\emptyset\}$ and $\mathcal{AT}\dot{+}\emptyset=\{\mathcal{AT}\}$. We are therefore left with with the simple contraction $\mathcal{AT}-\mathcal{S}$.

While (*Def* $\dot{-}$) has little notational similarity to the Harper Identity, it expresses the same concept and complements our analogue for the Levi Identity (*Def* $\dot{+}$). Where (*Def* $\dot{+}$) expresses revision in terms of contraction and expansion, (*Def* $\dot{-}$) expresses contraction in terms of revision and simple contraction. This is possible because the distinction between completely removing arguments and rendering them unacceptable means that simple contraction is defined in Argument Revision but not belief revision.

7 Measuring minimal change

One of the foundational principles in the AGM theory of belief revision is that of *minimal change*. When beliefs are changed, the process is carried out with as little impact as possible on the beliefs that remain. Several different criteria are used to determine minimality, summarized in Section 2.2.

In determining minimal change in Argument Revision, we select a method that is influenced by the principle of conservatism, which although defined in terms of (belief) contraction is easily extended to account for both argument contraction and revision. When contracting or revising an Argumentation Theory, select the specific contraction or revision that has the fewest measurable effects on the theory. We choose the principle of conservatism as our influence because ASPIC^{+} provides a rich structure where even a seemingly small change (e.g. adding or removing a single knowledge base premise) can have a significant and disproportionate impact on the rest of the theory; similarly, a seemingly big change (e.g. adding or removing multiple knowledge base premises) can have a lesser impact on the rest of the theory while still achieving the same contraction or revision. Measuring this impact allows for the selection of a specific contraction or revision that keeps to a minimum the number of overall changes to the theory.

In this section, four measurable effects of a contraction or revision on an ASPIC^{+} argumentation theory are provided. We do not necessarily regard this as an exhaustive list, but do consider it to be fundamental: these four effects should always be taken into account when determining the minimal way of revising or contracting an Argumentation Theory.

First, the structure of arguments themselves must be considered: an Argumentation Theory is revised through modifying its knowledge base in order to add or remove arguments. Any element in a knowledge base may be a premise (or represent a rule) in more than one arguments resulting in not only an input argument being eliminated from the system and the theory, but also other arguments that are completely unrelated. Consider the following knowledge base and rules in an Argumentation Theory $\mathcal{AT}$:

$$\mathcal{K}_p = \{a, \lceil a \Rightarrow_{r_1} b \rceil, \lceil a \Rightarrow_{r_2} c \rceil\} \qquad \mathcal{R}_d = \{a \Rightarrow_{r_1} b, a \Rightarrow_{r_2} c\}$$

The following arguments can be constructed from $\mathcal{AT}$:

$\mathcal{A}_1$: a $\qquad$ $\mathcal{A}_2$: $\lceil a \Rightarrow_{r_1} b \rceil$ $\qquad$ $\mathcal{A}_3$: $\lceil a \Rightarrow_{r_2} c \rceil$

$\mathcal{A}_4$: $\dfrac{\mathcal{A}_1(a)}{b} r_1$ $\qquad$ $\mathcal{A}_5$: $\dfrac{\mathcal{A}_1(a)}{c} r_2$

Assume that we have no further information (i.e. relating to contrariness, preferences, etc.) and wish to contract $\mathcal{AT}$ with respect to $\mathcal{A}_4$; i.e. $\mathcal{AT} \dot{-} \{\mathcal{A}_4\}$. The only possible methods are either to simple-contract with respect to $\mathcal{A}_1$ (a) or $\mathcal{A}_2$ ($a \Rightarrow b$). The former would result in not only $\mathcal{A}_1$ being removed from the theory, but also $\mathcal{A}_5$; the latter a lesser effect, because it impacts on fewer arguments overall than a being removed from the theory.

A further consideration is the acceptability of arguments. If a previously acceptable argument becomes unacceptable or completely removed following a revision process, any arguments of which it was the sole defeater would become acceptable; conversely, an unacceptable argument that becomes acceptable would render unacceptable any arguments of which it is a defeater.

Consider the following knowledge base, rule set and contrariness relations:

$$\mathcal{K}_p = \{a, c, d, f\} \qquad \mathcal{R}_d = \{a \Rightarrow_{r_1} b, d \Rightarrow_{r_2} e\} \qquad b \in \overline{d},\ d \in \overline{f},\ c \in \overline{\lceil b \in \overline{d} \rceil}$$

This yields the arguments:

$\mathcal{A}_1$: a $\qquad$ $\mathcal{A}_2$: c $\qquad$ $\mathcal{A}_3$: d

$\mathcal{A}_4$: f $\qquad$ $\mathcal{A}_5$: $\lceil a \Rightarrow_{r_1} b \rceil$ $\qquad$ $\mathcal{A}_6$: $\lceil d \Rightarrow_{r_2} e \rceil$

$\mathcal{A}_7$: $\lceil b \in \overline{d} \rceil$ $\qquad$ $\mathcal{A}_8$: $\lceil d \in \overline{f} \rceil$ $\qquad$ $\mathcal{A}_9$: $\lceil c \in \overline{\lceil b \in \overline{d} \rceil} \rceil$

$\mathcal{A}_{10}$: $\dfrac{\mathcal{A}_1}{b} r_1$ $\qquad$ $\mathcal{A}_{11}$: $\dfrac{\mathcal{A}_4}{e} r_2$

The labelled extended abstract framework derived from $\mathcal{AT}$, evaluated under grounded semantics, is shown on the left-hand side of Figure 4. A solid border means the argument is 'in' and a dotted border means the argument is 'out'; similarly, a solid attack means the attack succeeds and a dotted attack means it does not. A double-headed arrow is used to distinguish attacks on attacks from attacks on arguments. For clarity, arguments that neither attack nor are attacked (commonly referred to as 'islands') are omitted.

If we were to contract $\mathcal{AT}$ with respect to $\mathcal{A}_2$ the framework would change such that the attacks $(\mathcal{A}_{10}, \mathcal{A}_3)$ and $(\mathcal{A}_{10}, \mathcal{A}_{11})$ would succeed. As a consequence, $\mathcal{A}_3$ and $\mathcal{A}_{11}$ become unacceptable and $\mathcal{A}_4$ becomes acceptable. This is shown in the right-hand side of Figure 4.

This provides us with four measurable effects of a revision process on an Argumentation Theory:

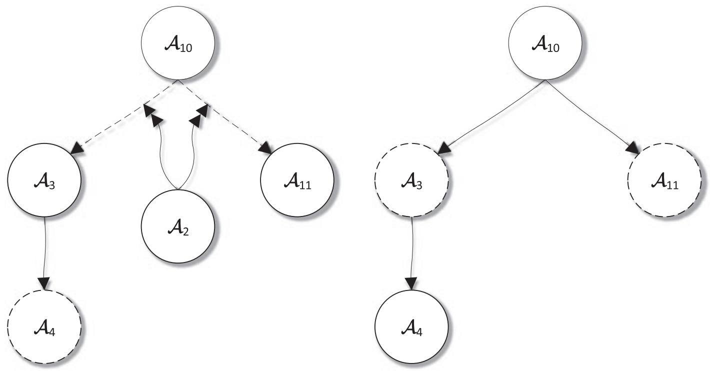

FIGURE 4. EAFs comparing the statuses of $\mathcal{A}_3$, $\mathcal{A}_4$ and $\mathcal{A}_{11}$.

- Argument loss—during a simple contraction, the arguments subject to the contraction will be lost from the theory. If those arguments are then sub-arguments in other arguments, those other arguments will also be lost.
- Acceptability loss—during a contraction or revision, previously acceptable arguments may become unacceptable. If those arguments defend (i.e. attack attackers of) other arguments, those other arguments may also become unacceptable and so on.
- Argument gain—during a revision or expansion, new arguments are added to the theory. These arguments may also be sub-arguments of other arguments that are also added to the theory.
- Acceptability gain—during a contraction or revision, previously unacceptable arguments may become acceptable. If those arguments defend other arguments, those arguments may also become acceptable and so on.

To capture these four effects, each is defined in terms of a function that examines the difference between two argumentation theories. The definition of each function takes two argumentation theories and outputs a set of arguments that satisfy the specified property (i.e. argument loss or gain, or acceptability loss or gain). Recall that $\Pi_{\mathcal{L}}$ represents the set of all argumentation theories based on the language $\mathcal{L}$, and $\mathcal{A}rgs(\Pi_{\mathcal{L}})$ is the set of all arguments in $\Pi_{\mathcal{L}}$. We use the superscript $\pm$ to denote an argumentation theory that has been changed in some way, but does not necessarily satisfy the properties of a specific contraction or revision.

To capture these four effects, each is defined in terms of a function related to expansions and simple contractions. The definition of each function involves an argumentation system $\mathcal{AS}$ in an Argumentation Theory $\mathcal{AT}$, and a formula ϕ in the knowledge base in $\mathcal{AS}$. Υ_G is used to represent the set of all argumentation theories in a change graph. Where the type of change is required to be explicit, we append to $\mathcal{AT}$ a superscript $+$ (for expansion) or $-$ (for simple contraction). Where it is necessary to represent a change that can be of either type, we append $\pm$.

In Section 5, we already defined two functions that encapsulate argument drop and argument gain. We will revisit these definitions here as part of the determination of minimal change.

DEFINITION 7.1
The *argument drop function* Δ_A:
$\Delta_A\colon \Pi_{\mathcal{L}} \times \Pi_{\mathcal{L}} \rightarrow 2^{\mathcal{A}rgs(\mathcal{AT})}$,
$\Delta_A(\mathcal{AT}, \mathcal{AT}^{\pm}) = \{\mathcal{A} \mid \mathcal{A} \in \mathcal{A}rgs(\mathcal{AT}), \mathcal{A} \notin \mathcal{A}rgs(\mathcal{AT}^{\pm})\}$

When expanding, arguments are simply added to the theory without consideration of acceptability. Therefore no arguments are dropped by an expansion.

Proposition 7.2
For any $\{\phi_1,\dots,\phi_n\} \subseteq \mathcal{L}$ with $\mathcal{AT}^{\pm} = \mathcal{AT} + \{\phi_1\} + \dots + \{\phi_n\}$ then $\Delta_A(\mathcal{AT}, \mathcal{AT}^{\pm}) = \emptyset$

The second function, the *acceptability drop function*, identifies those arguments that lose acceptability when removing or adding a formula.

Definition 7.3
The *acceptability drop function* Δ_S for an extension E in an Argumentation Theory $\mathcal{AT}$:
Δ_S: $\Pi_{\mathcal{L}} \times \Pi_{\mathcal{L}} \to 2^{\mathcal{A}rgs(\mathcal{AT})}$,
$\Delta_S(\mathcal{AT}, \mathcal{AT}^{\pm}) = \{\mathcal{A} \mid \mathcal{A} \in E(\mathcal{AT}), \mathcal{A} \in \mathcal{A}rgs(\mathcal{AT}^{\pm}), \mathcal{A} \notin E(\mathcal{AT}^{\pm})\}$

An expansion or simple-contraction with respect to an atomic argument always drops a conflict-free set.

Proposition 7.4
For any $\{\phi_1,\dots,\phi_n\} \subseteq \mathcal{L}_n$, $\Delta_S(\mathcal{AT}, \mathcal{AT} \pm \{\phi\})$ is conflict-free in $\mathcal{AF}_{\mathcal{AT}}$

The third function, the *argument gain function*, identifies those arguments that the system gains when removing or adding a formula.

Definition 7.5
The *argument gain function* Γ_A:
Γ_A: $\Pi_{\mathcal{L}} \times \Pi_{\mathcal{L}} \to 2^{\mathcal{A}rgs(\mathcal{AT}^{\pm})}$,
$\Gamma_A(\mathcal{AT}, \mathcal{AT}^{\pm}) = \{\mathcal{A} \mid \mathcal{A} \notin \mathcal{A}rgs(\mathcal{AT}),\ \mathcal{A} \in \mathcal{A}rgs(\mathcal{AT}^{\pm})\}$

An important property of simple contraction is that no arguments are gained following the operation.

Proposition 7.6
For any $\{\phi_1,\dots,\phi_n\} \subseteq \mathcal{L}_n$ with $\mathcal{AT}^{\pm} = \mathcal{AT} - \{\phi_1\} - \dots - \{\phi_n\}$ then $\Gamma_A(\mathcal{AT}, \mathcal{AT}^{\pm}) = \emptyset$

The final function, the *acceptability gain function*, identifies those arguments that were acceptable in the input argumentation system and have remained in the output system, but have lost acceptability.

Definition 7.7
The *acceptability gain function* Γ_S for an extension E in an Argumentation Theory $\mathcal{AT}$:
Γ_S: $\Pi_{\mathcal{L}} \times \Pi_{\mathcal{L}} \to 2^{\mathcal{A}rgs(\mathcal{AT})}$,
$\Gamma_S(\mathcal{AT}, \mathcal{AT}^{\pm}) = \{\mathcal{A} \mid \mathcal{A} \notin E(\mathcal{AT}),\ \mathcal{A} \in E(\mathcal{AT}^{\pm})\}$

The drop and gain functions for each type of change (i.e. argument and acceptability) are linked to each other, in that, for a given change, no arguments that are dropped are also gained, and vice versa, and no arguments that gain acceptability also lose it, and vice versa.

Proposition 7.8
For $X \in \{A, S\}$, $\Delta_X(\mathcal{AT}, \mathcal{AT}^{\pm}) \cap \Gamma_X(\mathcal{AT}, \mathcal{AT}^{\pm}) = \emptyset$

Given these properties, it can be proven that any Argumentation Theory can always be contracted by any subset of the arguments in the theory.

Proposition 7.9
For any $\mathcal{S} \subseteq \mathcal{A}rgs(\mathcal{AT})$, $\mathcal{AT} \dot{-} \mathcal{S}$ is always defined.

The same does not hold unconditionally for revision, however. Recall that a revision must (i) have all input arguments acceptable in the derived framework and (ii) lead to a well-formed Argumentation Theory. A simple example, however, can show how these principles are violated.

Consider arguments $\mathcal{A}_1$: p and $\mathcal{A}_2$: $\neg p$. It is not possible for both $\mathcal{A}_1$ and $\mathcal{A}_2$ to be acceptable under a unique extension semantics in an Argumentation Theory containing the arguments, and thus $\mathcal{AT}\dot{+}\{\mathcal{A}_1,\mathcal{A}_2\}$ is undefined. For $\mathcal{AT}\dot{+}\mathcal{S}$ to be defined, therefore, $\mathcal{S}$ must be conflict-free.

Recall that when an Argumentation Theory is revised, existing arguments may need to be contracted to ensure the input arguments are all acceptable. This allows us to take advantage of proposition 7.9 in proving that, provided $\mathcal{S}$ is conflict-free, $\mathcal{AT}\dot{+}\mathcal{S}$ is always defined.

PROPOSITION 7.10
For any conflict-free set of arguments $\mathcal{S}$ (based on $\mathcal{L}$), $\mathcal{AT}\dot{+}\mathcal{S}$ is always defined.

In practice, we use these four measures towards the determination of *path costs* in a change graph for each path from $\mathcal{AT}$ to every Argumentation Theory in Θ. In the present work, we assume that all four functions for measuring minimal change have equal weighting.

DEFINITION 7.11
Given a change graph $CG(\mathcal{AT},\Theta)=\langle\Upsilon,\Omega\rangle$, the *path cost* from $\mathcal{AT}$ to $\mathcal{AT}'\in\Theta$ is
$V(\mathcal{AT},\mathcal{AT}')=|\,\Delta_A(\mathcal{AT},\mathcal{AT}')\cup\Delta_S(\mathcal{AT},\mathcal{AT}')\cup\Gamma_A(\mathcal{AT},\mathcal{AT}')\cup\Gamma_S(\mathcal{AT},\mathcal{AT}')\,|$

Comparing the output of V for the paths from $\mathcal{AT}$ to each Argumentation Theory in Θ allows the minimal change(s) to be identified. The next section puts this approach to minimal change in Argument Revision to work in two scenarios.

8 Argument Revision in dialogue

In this section, we show how Argument Revision can assist a dialogue participant in selecting their next move. We will not use a specific dialogue framework to illustrate, but instead assume it is a game in the style of McBurney & Parsons [32] with the following properties:

(1) The dialogue has a set of participants P, where $|P|\geq 2$.
(2) Each participant $p\in P$ has a commitment store C_p^t that contains all publicly-shared beliefs after the move at turn ($t\in\mathbb{N}$) in the dialogue.
(3) C_p^t is closed under strict and defeasible rule application, with the closure operator Cl defined thus:

- $C_p^t\subseteq Cl(C_p^t)$
- if $\{\varphi_1,\ldots,\varphi_n,\ulcorner\varphi_1,\ldots,\varphi_n\rightsquigarrow\varphi\urcorner\}\subseteq Cl(C_p^t)$, then $\varphi\in Cl(C_p^t)$

where $\rightsquigarrow$ represents a consequence operator of unspecified type (strict or defeasible).
(4) $Cl(C_p^t)$ must be consistent.
(5) The communication language of the framework contains a locution *claim*$(\Phi\subseteq\mathcal{L})$ (or similar) such that when performed as move t $Cl(C_p^t)=Cl(C_p^{t-1}\cup\Phi)$.
(6) The communication language of the framework contains a locution *retract*$(\Phi\subseteq\mathcal{L})$ (or similar) such that when performed as move t $\forall\phi\in\Phi$, $\phi\notin C_p^t$. Note that $\phi\notin C_p^t$ does not guarantee $\phi\notin Cl(C_p^t)$ (see Section 8.1 below).

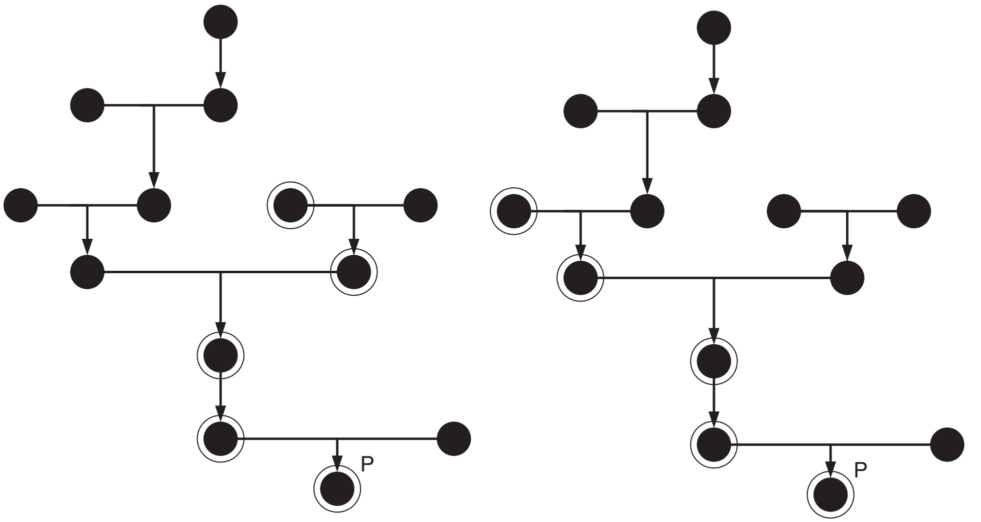

FIGURE 5. Stability adjustment from [45] (L) and alternative adjustment (R).

8.1 Commitment retraction

In certain dialogue frameworks, when a participant cannot defend a statement, *S*, against a counter-argument from their opponent, they ordinarily should retract *S*. However, retracting *S* alone is sometimes insufficient, because they may still hold commitment to other statements which *S* is a consequence of and thus their commitment store is inconsistent under closure. Consider, for instance, the following commitment store of player α (after some move *t*) from which ψ is to be retracted:

$$C^t_\alpha = \{\phi, \psi, \lceil \phi \Rightarrow \psi \rceil\}$$

Retracting ψ leaves $C^t_\alpha = \{\phi, \lceil \phi \Rightarrow \psi \rceil\}$; however $Cl(C^t_\alpha) = \{\phi, \psi, \lceil \phi \Rightarrow \psi \rceil\}$. It is therefore necessary to make further retractions to ensure ψ can no longer be inferred.

In dialogue protocols such as *PPD*, the goal of ensuring that a retracted statement can no longer be inferred in a commitment store is called *stability*, with the process used to reach it termed a *stability adjustment* [45]. Stability is achieved through retracting other statements so as to eliminate the possibility of inferring the originally retracted statement, with [45] using the diagram to the left in Figure 5 to illustrate such an adjustment following the retraction of a statement *P* (where filled circles are statements, circled statements are those to be retracted and the arrows indicate inference).

A limitation of [45]'s specification of stability adjustments is that they do not explain exactly which statements should be selected. Consider the structure on the right in Figure 5. This shows the same inference tree for *P*, but with a different stability adjustment with respect to *P*. The same number of retractions has taken place, which raises the question of how an agent should choose which adjustment to make.

Argument revision, and its determination of minimal change, offers a solution to this problem. By considering the retraction of commitments as an Argument Revision process, a participant can determine which, if any, methods are minimal, in the senses of minimal change specified in Section 7.

Such a process would not, however, be carried out with respect to only existing commitments. It is also important to consider what the participant might *potentially* become committed to at a future point in the dialogue. Consider again the sample commitment store provided above. It might be that, with respect to commitments, retracting ϕ is minimal (compared to retracting $\ulcorner\phi \Rightarrow \psi\urcorner$); but if ϕ is a premise in several as-yet unstated arguments in the Argumentation Theory from which the participant derives their beliefs, those arguments would be rendered incommunicable in the remainder of the dialogue (unless the course of the dialogue changed such that a future move returned ϕ to the commitment store).

Consider a dialogue participant α whose knowledge is derived from the Argumentation Theory $\mathcal{AT}_\alpha$ with the following knowledge base and rules:

$$\mathcal{K}(\mathcal{AT}_\alpha) = \left\{ \begin{array}{ll} a, & b, \\ d, & e, \\ f, & \ulcorner a \Rightarrow_{r_3} g\urcorner, \\ \ulcorner e \in \bar{f}\urcorner, & \ulcorner d \in \overline{\ulcorner e \in \bar{f}\urcorner}\urcorner, \\ \ulcorner j \in \overline{\ulcorner h \Rightarrow_{r_6} i\urcorner}\urcorner, & \ulcorner d \Rightarrow_{r_4} \ulcorner a \Rightarrow_{r_1} c\urcorner\urcorner, \\ \ulcorner d \Rightarrow_{r_5} \ulcorner b \Rightarrow_{r_2} c\urcorner\urcorner & \end{array} \right\}$$

$$\mathcal{R}(\mathcal{AT}_\alpha) = \left\{ \begin{array}{ll} (a \Rightarrow_{r_1} c), & (b \Rightarrow_{r_2} c), \\ (a \Rightarrow_{r_3} g), & (d \Rightarrow_{r_4} \ulcorner a \Rightarrow_{r_1} c\urcorner), \\ (d \Rightarrow_{r_5} \ulcorner b \Rightarrow_{r_2} c\urcorner) & \end{array} \right\}$$

The arguments in $\mathcal{AT}_\alpha$ are as follows:

$\mathcal{A}_1$: a

$\mathcal{A}_2$: b

$\mathcal{A}_3$: d

$\mathcal{A}_4$: e

$\mathcal{A}_5$: f

$\mathcal{A}_6$: $\ulcorner a \Rightarrow_{r_3} g\urcorner$

$\mathcal{A}_7$: $\ulcorner e \in \bar{f}\urcorner$

$\mathcal{A}_8$: $\ulcorner d \in \overline{\ulcorner e \in \bar{f}\urcorner}\urcorner$

$\mathcal{A}_9$: $\ulcorner j \in \overline{\ulcorner h \Rightarrow_{r_6} i\urcorner}\urcorner$

$\mathcal{A}_{10}$: $\ulcorner d \Rightarrow_{r_4} \ulcorner a \Rightarrow_{r_1} c\urcorner\urcorner$

$\mathcal{A}_{11}$: $\ulcorner d \Rightarrow_{r_5} \ulcorner b \Rightarrow_{r_2} c\urcorner\urcorner$

$\mathcal{A}_{12}$: $\dfrac{\mathcal{A}_3(d)}{\ulcorner a \Rightarrow_{r_1} c\urcorner} r_4$

$\mathcal{A}_{13}$: $\dfrac{\mathcal{A}_3(d)}{\ulcorner b \Rightarrow_{r_2} c\urcorner} r_5$

$\mathcal{A}_{14}$: $\dfrac{\mathcal{A}_1(a)}{c} r_1$

$\mathcal{A}_{15}$: $\dfrac{\mathcal{A}_2(b)}{c} r_2$

$\mathcal{A}_{16}$: $\dfrac{\mathcal{A}_1(a)}{g} r_3$

Assume that at a given time t, $C_\alpha^t = \{a, b, c, d, \ulcorner a \Rightarrow_{r_1} c\urcorner, \ulcorner b \Rightarrow_{r_2} c\urcorner, \ulcorner d \Rightarrow_{r_4} \ulcorner a \Rightarrow_{r_1} c\urcorner\urcorner, \ulcorner d \Rightarrow_{r_5} \ulcorner b \Rightarrow_{r_2} c\urcorner\urcorner\}$ $(Cl(C_\alpha^t) = C_\alpha)$ but α is now required to retract c. Retracting c on its own is insufficient to achieve stability, because:

$$Cl(C_\alpha^t \backslash \{c\}) = \{a, b, c, d, \ulcorner a \Rightarrow_{r_1} c\urcorner, \ulcorner b \Rightarrow_{r_2} c\urcorner, \ulcorner d \Rightarrow_{r_4} \ulcorner a \Rightarrow_{r_1} c\urcorner\urcorner, \ulcorner d \Rightarrow_{r_5} \ulcorner b \Rightarrow_{r_2} c\urcorner\urcorner\}$$

TABLE 5. Outputs of the functions for measuring minimal change

$\pm\{\phi\}$	Δ_A	Δ_S	Γ_A	Γ_S	Σ
$-\{d\}$	$\{\mathcal{A}_3, \mathcal{A}_{12}, \mathcal{A}_{13}, \mathcal{A}_{14}, \mathcal{A}_{15}\}$	$\{\mathcal{A}_5\}$	$\emptyset$	$\emptyset$	6
$-\{a\}$	$\{\mathcal{A}_1, \mathcal{A}_{14}, \mathcal{A}_{16}\}$	$\emptyset$	$\emptyset$	$\emptyset$	3
$-\{b\}$	$\{\mathcal{A}_2, \mathcal{A}_{15}\}$	$\emptyset$	$\emptyset$	$\emptyset$	2
$-\{\lceil d \Rightarrow_{r_4} \lceil a \Rightarrow_{r_1} c\rceil\rceil\}$	$\{\mathcal{A}_{10}, \mathcal{A}_{12}, \mathcal{A}_{14}\}$	$\emptyset$	$\emptyset$	$\emptyset$	3
$-\{\lceil d \Rightarrow_{r_5} \lceil a \Rightarrow_{r_2} c\rceil\rceil\}$	$\{\mathcal{A}_{11}, \mathcal{A}_{13}, \mathcal{A}_{15}\}$	$\emptyset$	$\emptyset$	$\emptyset$	3

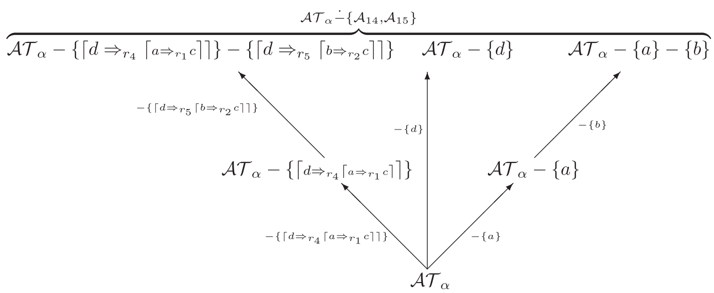

FIGURE 6. Change graph for $\mathcal{AT}_\alpha \dot{-} \{\mathcal{A}_{14}, \mathcal{A}_{15}\}$.

It is therefore necessary to make further retractions. c is the conclusion of two arguments, $\mathcal{A}_{14}$ and $\mathcal{A}_{15}$, thus with respect to α's beliefs this retraction is characterizable as the simple contraction $\mathcal{AT}_\alpha - \{\mathcal{A}_{14}, \mathcal{A}_{15}\}$. Table 8.1 shows the outputs of the acceptability drop and gain functions, and the total cost ($\Sigma = | \Delta_A \cup \Delta_S \cup \Gamma_A \cup \Gamma_S |$) of each simple contraction performed in achieving $\mathcal{AT}_\alpha - \{\mathcal{A}_{14}, \mathcal{A}_{15}\}$. Figure 6 shows the change graph for the contraction.

Given the outputs of the four functions at each edge on the change graph, α computes the overall costs of each path from $\mathcal{AT}_\alpha$ to the set of argumentation theories $\mathcal{AT}_\alpha \dot{-} \{\mathcal{A}_{14}, \mathcal{A}_{15}\}$:

$$V(\mathcal{AT}_\alpha, \mathcal{AT}_\alpha - \{d\}) = 6 \qquad V(\mathcal{AT}_\alpha, \mathcal{AT}_\alpha - \{a\} - \{b\}) = 3 + 2 = 5$$

$$V(\mathcal{AT}_\alpha, \mathcal{AT}_\alpha - \{\lceil d \Rightarrow_{r_4} \lceil a \Rightarrow_{r_1} c\rceil\rceil\} - \{\lceil d \Rightarrow_{r_5} \lceil a \Rightarrow_{r_2} c\rceil\rceil\}) = 2 + 2 = 4$$

Thus it can be seen that when α wishes to retract c, the minimal way of doing so is to retract the premises a and b.

This example has illustrated two important principles. The first is that considering argument acceptability is important when determining minimal change in a system of argumentation based on [22]'s abstract theory. Had acceptability not been considered, retracting d would have had a cost of 5, the same as retracting a and b. Second, the example also shows that a revision with the fewest steps is not always minimal. The identified minimal revision involves two steps—the removal of a, then the removal of b, reaching the goal with 5 changes overall. Removing only d would also have reached the goal of the contraction, but would have done so with a total of 6 changes.

8.2 *Dishonesty*

In this section, we build on our previous work in [43] in which we characterize dishonesty in dialogue as a form of internal argument revision. Dishonesty is a common human behaviour, with types of dishonesty ranging from innocent and trivial 'white lies', to those with more serious consequences. In either case, it is necessary for the speaker of the lie to maintain it, or face negative consequences (an upset child, or charges of corruption).

8.2.1 Why be dishonest?

In [43] we offer two possible reasons as to why a dialogue participant would be dishonest—1) where a participant wishes to keep certain beliefs and arguments private;[4] and 2) where it is less costly (w.r.t. minimal change) to be dishonest than retract an existing commitment (and associated arguments).

To briefly illustrate a situation where a dialogue participant may want to keep certain arguments private, we provide the following example from [43]. Consider the following knowledge base and rules in an argumentation system, where a government minister is attempting to convince a political opponent that they should not spend money on welfare. The real reason is that the government thinks a war is about to happen that will need to be paid for, but they do not want to share this. Instead, the position to not spend money on welfare is justified by saying money needs spent on education, despite the education department spending being under-budget. For clarity, we leave implicit the arguments for rules, preferences and contraries.

- $\mathcal{K}_p = \left\{ have\_money, education\_under\_budget, war \right\}$
- $\mathcal{R}_d = \left\{ \begin{array}{l} war, have\_money \Rightarrow_{r1} pay\_for\_war \\ have\_money \Rightarrow_{r2} pay\_for\_education \\ have\_money \Rightarrow_{r3} pay\_for\_welfare \\ education\_under\_budget \Rightarrow_{r4} \neg pay\_for\_education \end{array} \right\}$
- $P_{\mathcal{L}} = \{war\}$ (private formulae)

And with:

- $r2 < r1$, $r3 < r1$, $r2 < r4$, $r3 < r2$
- $\overline{pay\_for\_war} = \{pay\_for\_education, pay\_for\_welfare\}$
- $\overline{pay\_for\_education} = \{pay\_for\_war, pay\_for\_welfare\}$
- $\overline{pay\_for\_welfare} = \{pay\_for\_war, pay\_for\_education\}$

The arguments in the theory are as follows:

$\mathcal{A}_1$: $have\_money$

$\mathcal{A}_2$: $education\_under\_budget$

$\mathcal{A}_3$: war

$\mathcal{A}_4$: $\dfrac{\mathcal{A}_3(war), \mathcal{A}_1(have\_money)}{pay\_for\_war} r_1$

$\mathcal{A}_5$: $\dfrac{\mathcal{A}_1(have\_money)}{pay\_for\_education} r_2$

$\mathcal{A}_6$: $\dfrac{\mathcal{A}_1(have\_money)}{pay\_for\_welfare} r_3$

$\mathcal{A}_7$: $\dfrac{\mathcal{A}_2(education\_under\_budget)}{\neg pay\_for\_education} r_4$

The set of private arguments (based on the set of private formulae) is $P_{Args} = \{\mathcal{A}_3, \mathcal{A}_4\}$.

[4]Other scenarios exist where a participant may wish to keep beliefs private (e.g. in a medical context) but our intention here is to show that dishonesty is characterizable as a form of belief and, by extension, Argument Revision.

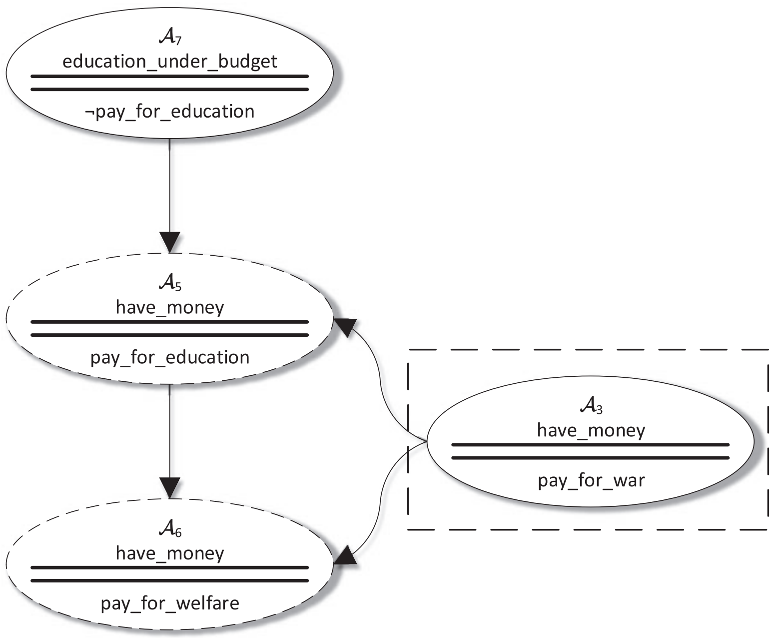

FIGURE 7. Argumentation Framework for the politician example.

The resultant framework, evaluated under grounded semantics, is shown in Figure 7, with islands (those arguments with no attack interactions) being omitted for clarity. Acceptable arguments have a solid border, unacceptable arguments have a dashed border and private arguments are surrounded by a dashed box.

The politician needs to defeat the argument $\mathcal{A}_6$. Since $\mathcal{A}_3$ is a private argument, the only remaining defeater of $\mathcal{A}_6$ is $\mathcal{A}_5$; however, $\mathcal{A}_5$ is not acceptable because it is defeated by $\mathcal{A}_7$. Thus, no honest defeater of $\mathcal{A}_6$ can be offered, without sharing a private argument. In terms of argument revision, to be dishonest with respect to $\mathcal{A}_5$ is characterised as the revision $\mathcal{AT}\dot{+}\{\mathcal{A}_5\}$. In this case, that revision would be carried out by contracting $\mathcal{A}_3$ $\mathcal{A}_7$.

8.2.2 Dishonesty in dialogue

To illustrate why a participant would be dishonest instead of retracting during a dialogue, we return to the example presented in Section 8.1. Let us again assume that α is in a position where, ordinarily, they must retract c from the commitment store:

$$C^t_\alpha = \{a, b, c, d, \lceil a \Rightarrow_{r_1} c\rceil, \lceil b \Rightarrow_{r_2} c\rceil, \lceil d \Rightarrow_{r_4} \lceil a \Rightarrow_{r_1} c\rceil\rceil, \lceil d \Rightarrow_{r_5} \lceil b \Rightarrow_{r_2} c\rceil\rceil\}$$

Let us further assume that the reason for the retraction is that α's opponent has advanced the following arguments:

$$\mathcal{B}_1\colon h, \quad \mathcal{B}_2\colon \lceil h \Rightarrow_{r_6} i\rceil, \quad \mathcal{B}_3\colon \lceil i \in \overline{c}\rceil, \quad \mathcal{B}_4\colon \frac{\mathcal{B}_1(h)}{i} r_6$$

TABLE 6. Outputs of the functions for measuring minimal change

$\pm\{\phi\}$	Δ_A	Δ_S	Γ_A	Γ_S	Σ
$+\{j\}$	$\emptyset$	$\emptyset$	$\{\mathcal{A}_{17}\}$	$\emptyset$	1

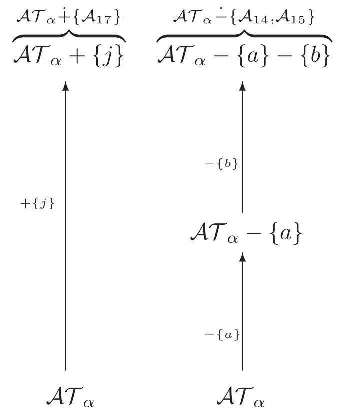

FIGURE 8. Change graph for $\mathcal{AT}_\alpha \dot{+} \{\mathcal{A}_{17}\}$ vs. (the minimal path for) $\mathcal{AT}_\alpha \dot{-} \{\mathcal{A}_{14}, \mathcal{A}_{15}\}$.

That is, an argument for i and another for i being an asymmetric contrary of (and thus defeats arguments for) c. α has no arguments that defeat i which leads to the retraction of c. Recall, however, that $\lceil j \in \overline{\lceil h \Rightarrow_{r_6} i \rceil} \rceil \in \mathcal{K}(\mathcal{AT}_\alpha)$. If α had an argument for j, they could undercut the rule $h \Rightarrow i$ and restore the acceptability of the arguments for c. α can thus be dishonest by claiming an argument for j in the dialogue. In order to do this, α must assess the impact of j on their existing Argumentation Theory and, by extension, beliefs.

The easiest method of introducing an argument for j is to revise $\mathcal{AT}_\alpha$ with respect to the atomic argument $\mathcal{A}_{17}$: j, i.e. $\mathcal{AT} \dot{+} (\mathcal{A}_{17}\}$. The change graph for $\mathcal{AT} \dot{+} (\mathcal{A}_{17}\}$ is shown in Figure 8 alongside the subgraph for $\mathcal{AT}_\alpha \dot{-} \{\mathcal{A}_{14}, \mathcal{A}_{15}$ that represents the minimal path. The the outputs of the functions for measuring minimal change are shown in Table 6.

The path cost from $\mathcal{AT}$ to the single theory in $\mathcal{AT} \dot{+} \{\mathcal{A}_{17}\}$ is:

$$V(\mathcal{AT}_\alpha, \mathcal{AT}_\alpha + \{j\}) = 1$$

Recall that the lowest cost retractions (removal of a and b) resulted in 5 changes to $\mathcal{AT}_\alpha$. We thus observe that in this situation, it is better (with respect to minimal change) for α to be dishonest by claiming an argument for j to defend c, instead of retracting their arguments for c.

This example has demonstrated two uses of Argument Revision in a dialogical context. An honest agent can use Argument Revision to determine the best way of retracting a statement; a dishonest agent can extend this by determining whether or not being dishonest is better than any possible retractions.

9 Related work

Adapting AGM-style belief revision operators and postulates to account for the dynamics of argumentation is, in general, not new, however these are usually restricted to Dung-style abstract argumentation frameworks. For instance, several works have incorporated dynamics into abstract argumentation by considering the outcome of an abstract argumentation framework [10, 17, 21]. Briefly, the outcome of a framework is a description of which arguments should be accepted and those which should now (under some semantics). The models proposed in [10, 21] permit only expansions of (or additions to) a framework and do not include contractions. Postulates for the revision of an argumentation system are presented in in [17], but these conflate revision and contraction; an outcome can describe arguments to be both accepted and not accepted. In the present work, separation between revision and contraction is maintained, as in the original AGM theory. Nevertheless, revision and contraction in our model can, at the abstract level, be seen as special cases of [17]'s model: in a revision, all arguments in the required outcome are accepted; conversely, in a contraction all arguments in the required outcome are not accepted.

Several models of dynamics in abstract argumentation incorporate postulates based on those proposed by Katsuno and Mendelzon [30]. These include the enforcement-based approaches of [9, 18], the translation-based approach of [16] and the extension-based approach of [20]. Katsuno and Mendelzon's postulates are based on the AGM postulates, but instead focus on belief *update*—where revision deals with changes to knowledge of the world, update deals with changes to the world itself. Our present work has focused specifically on adapting the AGM postulates directly to a system of argumentation, however examining the applicability of Katsuno and Mendelzon's postulates to our proposed model is a direction for future work.

AGM-style postulates for the revision of abstract argumentation frameworks are proposed by Baumann and Brewka [6]. These postulates are similar to those we propose in section 6 insofar as they are faithful to the original AGM postulates. There are, however, key differences. First, Baumann and Brewka's are specified in terms of a new Dung-logic, which used *k-models* instead of extensions for determining revisions. This new logic also incorporates a definition of strong equivalence, permitting an analogue of the sixth AGM revision postulate. Our model of Argument Revision is based upon an the ASPIC$^+$ system of argumentation in which the notion of strong equivalence is at present unclear. Nevertheless, should such a definition be forthcoming in the future it would be a relatively trivial step to define an analogous sixth postulate. Baumann and Brewka also provide analogues of the seventh and eighth AGM postulates (composite revisions); our aim in this article was to focus on the basic AGM postulates, and leave for future work the examination of composite revisions in our model.

Another significant difference between our postulates and those proposed by Baumann and Brewka is that they have maintained separation between postulates 3 and 4, whereas we conflated them into a single postulate. This arises from our differing approaches; by defining a new Dung logic that encodes the information contained within an argumentation framework, Baumann and Brewka are able to maintain the subset relation between revision and expansion. We on the other hand have chosen to specify postulates in terms of the argumentation frameworks themselves (via ASPIC$^+$ structure), which breaks this relation, because the goal of a revision may be achieved through adding new arguments.

Finally, we have also provided postulates for contraction, which in turn permitted analogues for the Levi and Harper identities; Baumann and Brewka focus only on expansion and revision.

Postulates for dynamics in structured argumentation are proposed in [36], which is based on Argument Theory Change (ATC) [35, 39]. As we noted in Section 2, ATC defines a new model (based on Dung's abstract frameworks) that incorporates a model of dynamics; in the present work, we have applied a model of dynamics to an existing model of argumentation. The postulates defined in [36] adapt the argumentation postulates of *success*, *consistency*, *inclusion*, *vacuicty*, *core-retainment* and *uniformity*. These have certain similarities and differences to our proposed postulates for Argument Revision: *success* is analogous to ($\mathcal{AT}\dot{+}2$) (all input arguments are acceptable); *consistency* is analogous to ($\mathcal{AT}\dot{+}5$) (all revised theories are well-formed, which ensures consistency in extensions); *vacuicty* is analogous to ($\mathcal{AT}\dot{+}3\backslash 4$) (if the union of the extension with all sub-arguments of the input arguments is conflict-free, then simply add the input arguments). *Inclusion*, *core-retainment* and *uniformity*, on the other hand, do not have analogous postulates in our model; for inclusion, this is because we assume that an argumentation theory may be revised by means other than adding only the required new information to the knowledge base (e.g. adding some other information may bring about a change of acceptability). For core-retainment, although there is no analogous postulate, its spirit is captured via change graphs, which impose a minimality constraint on changes to an argumentation theory. Finally, although the AGM model includes a postulate that is analogous to uniformity we chose not to include such a postulate because, as noted in section 6, defining strict equivalence between arguments in ASPIC$^+$ requires a level of analysis beyond the scope of this article.

10 Conclusions

In this article, we have presented Argument Revision, a model of dynamics for a system of structured argumentation (ASPIC$^+$). The model is influenced by the AGM theory of belief revision but with certain differences to account for the richer structure provided by ASPIC$^+$. Two types of argument change were specified—argument revision, where the goal is to ensure certain arguments are acceptable in the system, and argument contraction, where the goal is to ensure certain arguments are not acceptable in the system. These argument change processes are similar to their belief revision counterparts but with two key differences. First, the Argument Revision operators each return a set of potential theories that satisfy the goal (revision or contraction). Second, because consistency is handled by argument evaluation, we do not impose this on the resultant theories; instead, we require that they be well-formed and modular.

Postulates for argument revision and contraction were specified, which reflect the spirit of the first five basic AGM postulates for each process. We then demonstrated that argument revision can be expressed in terms of argument contraction and vice versa, identities that are similar the Levi and Harper identities (respectively) in belief revision.

We demonstrated that an Argumentation Theory can be revised and contracted by modifying the knowledge base in its argumentation system to add or remove premises. This is achieved through the simplest application of the specified revision and contraction operators, revising and contracting with respect to the atomic arguments that the premises create. A structure called a change graph was defined to model the series of steps required to secure a specific revision or contraction. A given series of steps can be assessed in terms of four measurable effects on the arguments in an Argumentation Theory: argument drop, argument gain, acceptability drop and acceptability gain. Using these measures, a path cost function identifies the minimal change(s) in a change graph.

To illustrate the use of Argument Revision in dialogue, two examples were presented to demonstrate that a participant in a dialogue can apply Argument Revision to determine whether or not they should concede an opinion to an opponent (and if so, exactly what form that concession should take), or instead be dishonest; a decision made on the basis of minimal change with respect to beliefs.

Understanding the dynamics of argumentation systems is a crucial component in the development of computational models of argument, with applications in areas as diverse as inter-agent communication and decision support. In Argument Revision, we have presented here a mechanism for managing change in argumentation systems which draws upon the AGM approach but which offers significantly more flexibility for open, real-world systems and lays a foundation for exploring a wide range of techniques for argument-based negotiation, persuasion and dialogue in general.

Acknowledgements

The authors are indebted to Guillermo Simari, Keith Edwards and the anonymous reviewers for their comments and feedback on earlier versions of this work.

Funding

This work was supported by The Engineering and Physical Sciences Research Council (EPSRC) of the UK government [EP/G060347/1]; and The Leverhulme Trust [RPG-2013-076].

References

[1] C. E. Alchourrón, P. Gärdenfors and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, **50**, 510–530, 1985.

[2] C. E. Alchourrón and D. Makinson. On the logic of theory change: contraction functions and their associated revision functions. *Theoria*, **48**, 14–37, 1982.

[3] L. Amgoud, L. Bodenstaff, M. Caminada, P. McBurney, S. Parsons, H. Prakken, J. van Veenen and G. A. W. Vreeswijk. Final review and report on formal argumentation system. Deliverable D2.6, ASPIC IST-FP6-002307, 2006.

[4] L. Amgoud and S. Vesic. Revising option status in argument-based decision systems. *Logic and Computation*, **22**, 1019–1058, 2012.

[5] H. Arló-Costa and I. Levi. Contraction: On the decision-theoretical origins of minimal change and entrenchment. *Synthese*, **152**, 129–154, 2006.

[6] R. Baumann and G. Brewka. AGM meets abstract argumentation: expansion and revision for Dung frameworks. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)*, pp. 2734–2740. Buenos Aires, Argentina, 2015.

[7] F. Bex, J. Lawrence, M. Snaith and C. Reed. Implementing the argument web. *Communications of the ACM*, **56**, 66–73, 2013.

[8] F. Bex, S. Modgil, H. Prakken and C. Reed. On logical reifications of the argument interchange format. *Journal of Logic and Computation*, **23**, 951–989, 2013.

[9] P. Bisquert, C. Cayrol, F. Dupin de Saint-Cyr and M. Lagasquie-Schiex. Enforcement in argumentation is a kind of update. In *Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM'2013)*, pp. 30–43, 2013.

[10] R. Booth, S. Kaci, T. Rienstra and L. van der Torre. A logical theory about dynamics in abstract argumentation. In *Proceedings of the 7th International Conference on Scalable Uncertainty Management (SUM 2013)*, pp. 148–161. Springer, Washington DC, USA, 2013.

[11] M. Caminada. An algorithm for computing semi-stable semantics. In *Proceedings of ECSQARU 2007*, pp. 222–234, 2007.

[12] M. Caminada. Comparing two unique extension semantics for formal argumentation: Ideal and eager. In *Proceedings of BNAIC 2007*, pp. 81–87, 2007.

[13] M. Caminada and L. Amgoud. On the evalutation of argumentation formalisms. *Artificial Intelligence*, **171**, 286–310, 2007.

[14] C. Cayrol, F. D. de Saint-Cyr and M. Lagasquie-Schiex. Change in abstract argumentation frameworks: Adding an argument. *Artificial Intelligence Research*, **38**, 49–84, 2010.

[15] C. Chesñevar, J. McGinnis, S. Modgil, I. Rahwan, C. Reed, G. Simari, M. South, G. Vreeswijk and S. Willmott. Towards an argument interchange format. *The Knowledge Engineering Review*, **21**, 293–316, 2006.

[16] S. Coste-Marquis, S. Konieczny, J. Mailly and P. Marquis. A translation-based approach for revision of argumentation frameworks. In *Proceedings of the 14th European Conference on Logics in Artificial Intelligence*, pp. 397–411. Springer, Madeira, Portugal, 2014.

[17] S. Coste-Marquis, S. Konieczny, J.G. Mailly, and P. Marquis. On the revision of argumentation systems: Minimal change of arguments statuses. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR 14)*, pp. 52–61, 2014.

[18] S. Coste-Marquis, S. Konieczny, J. G. Mailly and P. Marquis. Extension enforcement in abstract argumentation as an optimization problem. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)*, pp. 2876–2882, 2015.

[19] J. Devereux and C. Reed. Strategic argumentation in rigorous persuasion dialogue. In *Proceedings of the Sixth International Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2009)*, Lecture Notes in Computer Science, pp. 94–113, 2009.

[20] M. Diller, A. Haret, T. Linsbichler, S. Rümmele and S. Woltran. An extension-based approach to belief revision in abstract argumentation. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)*, pp. 2926–2932. Buenos Aires, Argentina, 2015.

[21] S. Doutre, A. Herzig and L. Perrussel. A dynamic logic framework for abstract argumentation. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR 14)*, pp. 62–71, 2014.

[22] P. M. Dung. On the acceptability of arguments and its fundemental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, **77**, 321–357, 1995.

[23] M. A. Falappa, G. Kern-Isberner, M. D. L. Reis and G. R. Simari. Prioritized and non-prioritized multiple change on belief bases. *Journal of Philosophical Logic*, **41**, 77–113, 2012.

[24] M. A. Falappa, G. Kern-Isberner and G. R. Simari. Belief revision and argumentation theory. In *Argumentation in Artificial Intelligence*, I. Rahwan and G. R. Simari, eds, pp. 314–360. Springer, 2009.

[25] M. A. Falappa, G. Kern-Isberner and G. R. Simari. Explanations, belief revision and defeasible reasoning. *Artificial Intelligence*, **141**, 1–28, 2002.

[26] E. L. Fermé, D. M. Gabbay and G. R. Simari, eds. *Trends in Belief Revision and Argumentation Dynamics*. College Publications, 2013.

[27] P. Gärdenfors. *Knowledge in Flux*. MIT Press, 1988.

[28] P. Gärdenfors. Belief revision: an introduction. In *Belief Revision*, P. Gärdenfors, ed., pp. 1–27. Cambridge University Press, 1992.

[29] G. Harman. *Change in View: Principles of Reasoning*. The MIT Press, 1986.

[30] H. Katsuno and A. O. Mendelzon. On the difference between updating a knowledge base and revising it. In *Proceedings of the 2nd International Conference on Principles of Knowledge*

Representation and Reasoning (KR'91), pp. 387–394. San Francisco, CA, USA, Morgan Kaufmann, 1991.

[31] P. Krümpelmann, M. Thimm, M. A. Falappa, A. Garca, G. Kern-Isberner and G. R. Simari. Selective revision by deductive argumentation. In *Theory and Applications of Formal Argumentation*, S. Modgil, N. Oren and F. Toni, eds, pp. 147–162. Barcelona, Spain, Springer, 2012.

[32] P. McBurney and S. Parsons. Games that agents play: A formal framework for dialogues between autonomous agents. *Journal of Logic, Language and Information*, **11**, 315–334, 2002.

[33] S. Modgil. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, **173**, 901–1040, 2009.

[34] S. Modgil and H. Prakken. Reasoning about preferences in structured extended argumentation frameworks. In P. Baroni, F. Cerutti and G. R. Simari, editors, *Proceedings of the 3rd International Conference on Computational Models of Argument (COMMA 2010)*, pp. 347–358, Desenzano del Garda, Italy, IOS Press, 2010.

[35] M. O. Moguillansky, N. D. Rotstein, M. A. Falappa, A. J. García and G. R. Simari. Argument theory change through defeater activation. In *Proceedings of the 3rd International Conference on Computational Models of Argument (COMMA 2010)*, P. Baroni, F. Cerutti and G. R. Simari, eds, pp. 359–366. Desenzano del Garda, Italy, IOS Press, 2010.

[36] M. O. Moguillansky, R. Wassermann and M. A. Falappa. Inconsistent-tolerant base revision through argument theory change. *Logic Journal of IGPL*, **20**, 2012.

[37] J. L. Pollock. Defeasible reasoning. *Cognitive Science*, **11**, 481–518, 1987.

[38] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, **1**, 93–124, 2010.

[39] N. D. Rotstein, M. O. Moguillansky, M. A. Falappa, A. J. García and G. R. Simari. Argument theory change: Revision upon warrant. In *Proceedings of the 2nd International Conference on Computational Models of Argument (COMMA 2008)*, Ph. Besnard, S. Doutre and A. Hunter, eds, pp. 336–347. Toulouse, France, IOS Press, 2008.

[40] H. Rott and S. O. Hansson. Safe contraction revisited. In *David Makinson on Classical Methods for Non-Classical Problems*, pp. 35–70. Springer, 2014.

[41] H. Rott and M. Pagnucco. Severe withdrawal (and recovery). *Journal of Philosophical Logic*, **28**, 501–547, 1999.

[42] M. Snaith and C. Reed. Justified argument revision in agent dialogue. In *Proceedings of the Ninth International Workshop on Argumentation in Mult-Agent Systems (ArgMAS 2012)*, 2012.

[43] M. Snaith and C. Reed. Argument revision as a means of supporting dishonesty. In *Trends in Belief Revision and Argument Dynamics*, E. L. Fermé, D. M. Gabbay and G. R. Simari, eds, pp. 229–246. College Publications, 2013.

[44] G. A. Vreeswijk. Abstract argumentation systems. *Artificial Intelligence*, **90**, 225–279, 1997.

[45] D. N. Walton and E. C. W. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press, New York, 1995.

[46] T. L. van der Weide and F. Dignum. Reasoning about and discussing preferences between arguments. In *Proceedings of the Eighth International workshop on Argumentation in Multi-Agent Systems (ArgMAS 2011)*, P. McBurney, S. Parsons and I. Rahwan, eds, pp. 117–135, Taipei, Taiwan, Springer, 2011.

Received 27 November 2014

A A Proofs

THEOREM A.1 (*Def* +)

$$\mathcal{AT}\dot{+}\mathcal{S} = \bigcup_{\mathcal{S}'\in DS^{\dot{+}}(\mathcal{S})} \bigcup_{\mathcal{AT}^-\in(\mathcal{AT}\dot{-}\mathcal{S}')} (\mathcal{AT}^- + \mathcal{S})$$

PROOF. From Definition 5.3, every revision disruption set for $\mathcal{S}$ contains at least one element from every disruption set for $\mathcal{S}$, thus:

$$\forall \mathcal{S}_1 \in DS^{\dot{-}}(\mathcal{S}), \forall \mathcal{S}' \in DS^{\dot{+}}(\mathcal{S}), |\mathcal{S}_1 \cap \mathcal{S}'| \geq 1.$$

Therefore, contracting $\mathcal{AT}$ with respect to a revision disruption set for $\mathcal{S}$ means there are no disruption sets for $\mathcal{S}$ acceptable in the contracted theory, thus:

$$\forall \mathcal{S}' \in DS^{\dot{+}}(\mathcal{S}), \forall \mathcal{AT}^- \in (\mathcal{AT}\dot{-}\mathcal{S}'), \nexists \mathcal{S}_1 \in DS^{\dot{-}}(\mathcal{S}) \text{ s.t. } \mathcal{S}_1 \subseteq E(\mathcal{AT}^-).$$

It follows, from Definition 5.2, that $E(\mathcal{AT}^-) \cup Sub(\mathcal{S})$ is conflict-free and thus from postulate $\mathcal{AT}\dot{+}3\backslash 4$:

$$\forall \mathcal{AT}^+ \in (\mathcal{AT}^- + \mathcal{S}), \mathcal{S} \subseteq E(\mathcal{AT}^+).$$

From Definition 4.5, if all arguments in $\mathcal{S}$ are acceptable in a theory, that theory is in $\mathcal{AT}\dot{+}\mathcal{S}$:

$$\nexists \mathcal{AT}' \in \Pi_{\mathcal{L}} \text{ s.t. } S \subseteq E(\mathcal{AT}') \text{ and } \mathcal{AT}' \notin (\mathcal{AT}\dot{+}\mathcal{S}).$$

Therefore, every theory that is contracted by a revision disruption set for $\mathcal{S}$ and then subsequently expanded with respect to $\mathcal{S}$ is a revision of $\mathcal{AT}$ by $\mathcal{S}$:

$$\bigcup_{\mathcal{S}'\in DS^{\dot{+}}(\mathcal{S})} \bigcup_{\mathcal{AT}^-\in(\mathcal{AT}\dot{-}\mathcal{S}')} (\mathcal{AT}^- + \mathcal{S}) \subseteq (\mathcal{AT}\dot{+}\mathcal{S})$$

From Definition 4.5, all arguments in $\mathcal{S}$ are acceptable in all revisions of some theory $\mathcal{AT}$ by $\mathcal{S}$. Therefore, from Definition 5.3, for every revision disruption set $\mathcal{S}'$ for $\mathcal{S}$ there exists a theory in $\mathcal{AT}\dot{+}\mathcal{S}$ where no arguments in $\mathcal{S}'$ are acceptable:

$$\forall \mathcal{S}' \in DS^{\dot{+}}(\mathcal{S}), \exists \mathcal{AT}^+ \in (\mathcal{AT}\dot{+}\mathcal{S}) \text{ s.t. } \mathcal{S}' \cap E(\mathcal{AT}^+) = \emptyset$$

From Definition 4.3, if no arguments in $\mathcal{S}'$ are acceptable in a theory, that theory is in $\mathcal{AT}\dot{-}\mathcal{S}'$:

$$\nexists \mathcal{AT}' \in \Pi_{\mathcal{L}} \text{ s.t. } S' \cap E(\mathcal{AT}') = \emptyset \text{ and } \mathcal{AT}' \notin (\mathcal{AT}\dot{-}\mathcal{S}')$$

Therefore, $\mathcal{AT}\dot{-}\mathcal{S}'$ contains all theories in which no arguments in $\mathcal{S}'$ are acceptable. When $\mathcal{S}' \in DS^{\dot{+}}(\mathcal{S})$ then $\forall \mathcal{AT}^- \in \mathcal{AT}\dot{-}\mathcal{S}'$, $E(\mathcal{AT}^-) \cup Sub(\mathcal{S})$ is conflict-free; therefore, $\mathcal{AT}^-$ is a revision of $\mathcal{AT}$ by $\mathcal{S}$.

Since $\mathcal{AT} \dot{-} \mathcal{S}'$ contains all theories in which no arguments in $\mathcal{S}'$ are acceptable, and $DS^{\dot{+}}(\mathcal{S})$ contains all sets of arguments that make $\mathcal{S}$ unacceptable, it follows that:

$$\forall \mathcal{AT} \in \Pi_{\mathcal{L}} \text{ s.t. } \mathcal{S} \subseteq E(\mathcal{AT}),\ \mathcal{AT} \in \bigcup_{\mathcal{S}' \in DS^{\dot{+}}(\mathcal{S})} \bigcup_{\mathcal{AT}^- \in (\mathcal{AT} \dot{-} \mathcal{S}')} (\mathcal{AT}^- + \mathcal{S})$$

It therefore follows that:

$$(\mathcal{AT} \dot{+} \mathcal{S}) \subseteq \bigcup_{\mathcal{S}' \in DS^{\dot{+}}(\mathcal{S})} \bigcup_{\mathcal{AT}^- \in (\mathcal{AT} \dot{-} \mathcal{S}')} (\mathcal{AT}^- + \mathcal{S})$$

And therefore:

$$\mathcal{AT} \dot{+} \mathcal{S} = \bigcup_{\mathcal{S}' \in DS^{\dot{+}}(\mathcal{S})} \bigcup_{\mathcal{AT}^- \in (\mathcal{AT} \dot{-} \mathcal{S}')} (\mathcal{AT}^- + \mathcal{S})$$ ■

THEOREM A.2
(*Def* $\dot{-}$)

$$\mathcal{AT} \dot{-} \mathcal{S} = \bigcup_{\mathcal{S}' \in 2^{\mathcal{S}}} \bigcup_{\mathcal{S}_D \in DS^{\dot{-}}(\mathcal{S}')} \bigcup_{\mathcal{AT}^+ \in (\mathcal{AT} \dot{+} \mathcal{S}_D)} (\mathcal{AT}^+ - \mathcal{S})$$

PROOF. Throughout, let $\mathcal{S}' \in 2^{\mathcal{S}}$.

From Definition 4.3, $\forall \mathcal{AT}^- \in (\mathcal{AT} \dot{-} \mathcal{S}')$, either $\mathcal{S}' \cap \mathcal{A}rgs(\mathcal{AT}^-) = \emptyset$ or $\exists \mathcal{S}'' \in DS^{\dot{-}}(\mathcal{S}')$ s.t. $\mathcal{S}'' \subseteq E(\mathcal{AT}^-)$ (all subsets of $\mathcal{S}$ are either defeated or nonexistent in a theory contracted with respect to $\mathcal{S}$).

From Definition 5.2, no arguments in $\mathcal{S}'$ are acceptable in a theory that is revised with respect to any contraction disruption set for $\mathcal{S}'$, thus:

$$\forall \mathcal{S}_D \in DS^{\dot{-}}(\mathcal{S}'), \forall \mathcal{AT}^+ \in (\mathcal{AT} \dot{+} \mathcal{S}_D), \mathcal{S}' \cap E(\mathcal{AT}^+) = \emptyset.$$

From Definition 4.5, $(\mathcal{AT} \dot{+} \mathcal{S}'')$ contains all theories in which all arguments in $\mathcal{S}''$ are acceptable:

$$\nexists \mathcal{AT}' \in \Pi_{\mathcal{L}} \text{ s.t. } S'' \subseteq E(\mathcal{AT}') \text{ and } \mathcal{AT}' \notin (\mathcal{AT} \dot{+} \mathcal{S}'').$$

It therefore follows that the set

$$\bigcup_{\mathcal{S}'' \in DS^{\dot{-}}(\mathcal{S}')} (\mathcal{AT} \dot{+} \mathcal{S}'')$$

contains all theories in which all arguments in $\mathcal{S}'$ are defeated. Thus, the set

$$\bigcup_{\mathcal{S}_D \in DS^{\dot{-}}(\mathcal{S}')} \bigcup_{\mathcal{AT}^+ \in (\mathcal{AT} \dot{+} \mathcal{S}_D)} (\mathcal{AT}^+ - \mathcal{S})$$

contains all theories in which all arguments in $\mathcal{S}'$ are defeated and all arguments in $\mathcal{S} \setminus \mathcal{S}'$ are completely removed (simple-contracted).

Therefore, the set

$$\bigcup_{\mathcal{S}'\in 2^{\mathcal{S}}} \bigcup_{\mathcal{S}_D\in DS^{\dot{-}}(\mathcal{S}')} \bigcup_{\mathcal{AT}^+\in(\mathcal{AT}\dot{+}\mathcal{S}_D)} (\mathcal{AT}^+ - \mathcal{S})$$

contains all theories in which all arguments in $\mathcal{S}$ are either unacceptable or completely removed and thus:

$$\mathcal{AT}\dot{-}\mathcal{S} = \bigcup_{\mathcal{S}'\in 2^{\mathcal{S}}} \bigcup_{\mathcal{S}_D\in DS^{\dot{-}}(\mathcal{S}')} \bigcup_{\mathcal{AT}^+\in(\mathcal{AT}\dot{+}\mathcal{S}_D)} (\mathcal{AT}^+ - \mathcal{S})$$ ■

PROPOSITION A.3
7.2 For any $\{\phi_1,\ldots,\phi_n\}\subseteq\mathcal{L}$ with $\mathcal{AT}^{\pm}=\mathcal{AT}+\{\phi_1\}+\ldots+\{\phi_n\}$ then $\Delta_A(\mathcal{AT},\mathcal{AT}^{\pm})=\emptyset$

PROOF. Consider the opposite: $\Delta_A(\mathcal{AT},\mathcal{AT}^{\pm})\neq\emptyset$. Thus, $\forall\mathcal{A}\in\Delta_A(\mathcal{AT},\mathcal{AT}^{\pm})$, $\mathcal{A}\in\mathcal{A}rgs(\mathcal{AT})$ and $\mathcal{A}\notin\mathcal{A}rgs(\mathcal{AT}^{\pm})$, and hence $Prem(\mathcal{A})\subseteq\mathcal{K}(\mathcal{AT})$ and $Prem(\mathcal{A})\not\subseteq\mathcal{K}(\mathcal{AT}^{\pm})$. However, $\mathcal{K}(\mathcal{AT})\subseteq\mathcal{K}(\mathcal{AT}+\{\phi\})$ and hence $\mathcal{K}(\mathcal{AT})\subseteq\mathcal{K}(\mathcal{AT}+\{\phi_1\}+\ldots+\{\phi_n\})$. Contradiction! ■

PROPOSITION A.4
7.4 For any $\{\phi_1,\ldots,\phi_n\}\subseteq\mathcal{L}_n$, $\Delta_S(\mathcal{AT},\mathcal{AT}\pm\{\phi\})$ is conflict-free in $\mathcal{AF}_{\mathcal{AT}}$

PROOF. Consider an extension E under some single-extension semantics subsumed by complete semantics. From Definition 7.3, $\mathcal{A}\in\Delta_S(\mathcal{AT},\mathcal{AT}\pm\{\phi\})$ iff $\mathcal{A}\in E(AT)$ and $\mathcal{A}\notin E(\mathcal{AT}\pm\{\phi\})$. Thus $\Delta_S(\mathcal{AT},\mathcal{AT}\pm\{\phi\})\subseteq E(\mathcal{AT})$. Since $E(\mathcal{AT})$ is conflict-free [22], all subsets of $E(\mathcal{AT})$ are also conflict-free. ■

PROPOSITION A.5
7.6 For any $\{\phi_1,\ldots,\phi_n\}\subseteq\mathcal{L}_n$ with $\mathcal{AT}^{\pm}=\mathcal{AT}-\{\phi_1\}-\ldots-\{\phi_n\}$ then $\Gamma_A(\mathcal{AT},\mathcal{AT}^{\pm})=\emptyset$

PROOF. Consider the opposite: $\mathcal{A}\notin\mathcal{A}rgs(\mathcal{AT})$, $\mathcal{A}\in\mathcal{A}rgs(\mathcal{AT}^{\pm})$. Thus, $Prem(\mathcal{A})\not\subseteq\mathcal{K}(\mathcal{AT})$ and $Prem(\mathcal{A})\subseteq\mathcal{K}(\mathcal{AT}^{\pm})$. However, $\mathcal{K}(\mathcal{AT}^{\pm})=\mathcal{K}(\mathcal{AT})\backslash\{\phi_1,\ldots,\phi_n\}$, hence $Prem(\mathcal{A})\subseteq\mathcal{K}(\mathcal{AT})$ and $\mathcal{A}\in\mathcal{A}rgs(\mathcal{AT})$. Contradiction. ■

PROPOSITION A.6
7.8 For $X\in\{A,S\}$, $\Delta_X(\mathcal{AT},\mathcal{AT}^{\pm})\cap\Gamma_X(\mathcal{AT},\mathcal{AT}^{\pm})=\emptyset$

PROOF. Consider the opposite in terms of argument drop and gain: $\mathcal{A}\in\Delta_A(\mathcal{AT},\mathcal{AT}^{\pm})$, $\mathcal{A}\in\Gamma_A(\mathcal{AT},\mathcal{AT}^{\pm})$. From definitions 7.1 and 7.5, $\mathcal{A}\notin\mathcal{A}rgs(\mathcal{AT}^{\pm})$ and $\mathcal{A}\in\mathcal{A}rgs(\mathcal{AT}^{\pm})$ respectively. Contradiction!
Consider the opposite in terms of acceptability drop and gain: $\mathcal{A}\in\Delta_S(\mathcal{AT},\mathcal{AT}^{\pm})$, $\mathcal{A}\in\Gamma_S(\mathcal{AT},\mathcal{AT}^{\pm})$. From definitions 7.3 and 7.7, $\mathcal{A}\notin E(\mathcal{AT}^{\pm})$, $\mathcal{A}\in E(\mathcal{AT}^{\pm})$. Contradiction. ■

PROPOSITION A.7
7.9 For any $\mathcal{S}\subseteq\mathcal{A}rgs(\mathcal{AT})$, $\mathcal{AT}\dot{-}\mathcal{S}$ is always defined.

PROOF. $\forall\mathcal{A}\in\mathcal{A}rgs(\mathcal{AT})$, for any $\phi\in Prem(\mathcal{A})$, $\mathcal{A}\in\Delta_A(\mathcal{AT},\mathcal{AT}-\{\phi\})$. Thus, in the extreme case, $\bigcup_{\mathcal{A}\in\mathcal{S}} Prem(\mathcal{A})=\{\phi_1,\ldots,\phi_n\}$, $S\subseteq\Delta_A(\mathcal{AT},\mathcal{AT}-\{\phi_1,\ldots,\phi_n\})$. ■

PROPOSITION A.8
7.10 For any conflict-free set of arguments $\mathcal{S}$, $\mathcal{AT}\dot{+}\mathcal{S}$ is always defined.

PROOF. Since for any $\mathcal{S}' \subseteq \mathcal{A}rgs(\mathcal{AT})$ $\mathcal{AT} \dot{-} \mathcal{S}'$ is always defined (Proposition 7.9), it follows that $\forall \mathcal{AT}^- \in (\mathcal{AT} - \mathcal{A}rgs(\mathcal{AT}))$, $\mathcal{A}rgs(\mathcal{AT}^-) = \emptyset$ and by extension $E(\mathcal{AT}^-) = \emptyset$. Thus $E(\mathcal{AT}^-) \cup Sub(\mathcal{S})$ is conflict-free; therefore $\mathcal{AT}^- \dot{+} \mathcal{S}$ is defined and by extension $\mathcal{AT} \dot{+} \mathcal{S}$ is defined. ■