Modelling Accrual of Arguments in *ASPIC*$^+$

Henry Prakken
Utrecht University, Department of Information and Computing Sciences and University of Groningen, Faculty of Law, The Netherlands
h.prakken@uu.nl

ABSTRACT

In this paper a new formal model of argument accrual is proposed as an adaptation of the *ASPIC*$^+$ framework for structured argumentation. The new model aims to overcome several weaknesses of existing proposals. It is shown to have desirable formal properties that are in line with standard work on formal argumentation, and to be applicable to a range of situations in legal reasoning.

CCS CONCEPTS

• **Computing methodologies** → *Nonmonotonic, default reasoning and belief revision.*

KEYWORDS

legal argument, argument accrual, formal argumentation systems

ACM Reference Format:
Henry Prakken. 2019. Modelling Accrual of Arguments in *ASPIC*$^+$. In *Seventeenth International Conference on Artificial Intelligence and Law (ICAIL '19), June 17–21, 2019, Montreal, QC, Canada.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3322640.3326703

1 INTRODUCTION

One recurring theme in the computational study of argumentation is that of accrual of arguments, or how several arguments for the same conclusion should be combined. This issue has been especially (although not exclusively) been studied in the context of AI & Law. The main issue is whether accrual should be modelled at the knowledge representation level, by combining different reasons for the same conclusion in antecedents of rules (in [15] called the *KR approach*), or whether it should be modelled at the logical level as a logical operation on arguments (in [15] called the *inference approach*). The present paper is about the inference approach to accrual. To my knowledge, Verheij was the first who pursued this approach in his dissertation [19]. In [15] Prakken proposed three principles that any model of argument accrual should satisfy and then proposed an inference-based model that satisfied these three principles in terms of a combination of Dung's theory of abstract of argumentation frameworks [6] with Pollock's theory of defeasible reasons [14]. The model was implemented by Wietske Visser[1].

[1] Available at http://www.wietskevisser.nl/research/epr/.

ICAIL '19, June 17–21, 2019, Montreal, QC, Canada

ACM ISBN 978-1-4503-6754-7/19/06…$15.00
https://doi.org/10.1145/3322640.3326703

Among the subsequent work is Gordon's recent proposal in [8, 10] in terms of a new version of the Carneades system.

So why write yet another paper on argument accrual? There are several reasons for doing so. First, Prakken's [15] model of accrual, although satisfying the three principles of accrual, has several drawbacks. A computational drawback (which it shares with any KR approach) is that every subset of a set of accruing arguments has to be considered, which leads to an exponential increase in the set of arguments that needs to be considered. Another problem of the model is that it requires the labelling of conclusions of arguments with the premises from which they are derived, which is inelegant and arguably unnatural for knowledge engineers. Third, Prakken's model had a specific problem with (non-)accrual of arguments with strict top rules (discussed but not optimally solved in [15]). Finally, the formal background of Prakken's 2005 approach was only semi-formally sketched, while a full formalisation is desirable for various reasons. These problems are avoided in Gordon's recent proposal. However, there are also reasons to improve on his approach, above all that there are counterexamples to Gordon's claim that his operator on statement labellings is monotonic. This in turn makes that his set of skeptical conclusions cannot be uniquely characterised and that his various semantics do not stand in the same relation as in the standard theory of abstract argumentation frameworks.

Accordingly, the aim of the present paper is to present an inference-based model of argument accrual in the context of the *ASPIC*$^+$ framework that improves on the proposal of [15] by adapting ideas of [8, 10], but in a way that avoids the technical problems with Gordon's proposal. Then the relevance of the new model for AI & Law will be demonstrated with several applications to legal reasoning.

This paper is organised as follows. First in Section 2 the theory of abstract argumentation frameworks and the *ASPIC*$^+$ framework will be introduced. Then in Section 3 the previous approaches of [15] and [8, 10] will be critically examined, after which the new formal model will be presented in Section 4 and formally investigated in Section 5. Then the model will be applied to legal examples in Section 6, after which conclusions will be drawn in Section 7.

2 FORMAL PRELIMINARIES

In this section I summarise the formal frameworks used below.

2.1 Semantics of abstract argumentation frameworks

An *abstract argumentation framework* (*AF*) is a pair $\langle \mathcal{A}, \mathcal{D} \rangle$, where $\mathcal{A}$ is a set of arguments and $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ is a relation of defeat. The theory of *AFs* [6] identifies sets of arguments (called *extensions*) which are internally coherent and defend themselves against defeaters. An argument $A \in \mathcal{A}$ is *defended* by a set by $S \subseteq \mathcal{A}$ if for all $B \in \mathcal{A}$: if B defeats A, then some $C \in S$ defeats B. Then relative to a given AF, $E \subseteq \mathcal{A}$ is *admissible* if E is conflict-free and defends all its members; E is a *complete extension* if E is admissible and $A \in E$ iff A is defended by E; E is a *preferred extension* if E is a $\subseteq$-maximal admissible set; E is a *stable extension* if E is admissible and attacks all arguments outside it; and $E \subseteq \mathcal{A}$ is the *grounded extension* if E is the least fixpoint of operator F, where $F(S)$ returns all arguments defended by S. It holds that any preferred, stable or grounded extension is a complete extension. Finally, for $T \in$ {complete, preferred, grounded, stable}, X is *sceptically* or *credulously* justified under the T semantics if X belongs to all, respectively at least one, T extension.

An alternative way to characterise the various semantics is with *labellings*.

Definition 2.1. A *labelling* of a set $\mathcal{A}$ of an argument an abstract argumentation framework $(\mathcal{A}, \mathcal{D})$ is any pair of non-overlapping subsets (*In*,*Out*) of $\mathcal{A}$. A labelling of $\mathcal{A}$ is a *d-labelling* of $(\mathcal{A}, \mathcal{D})$ iff it satisfies the following constraints:

(1) an argument is *in* iff all arguments defeating it are *out*.
(2) an argument is *out* iff it is defeated by an argument that is *in*.

Then *stable semantics* d-labels all arguments, while *grounded semantics* minimises and *preferred semantics* maximises the set of arguments that are d-labelled *in* and *complete semantics* allows any d-labelling. It has been shown that the T-extensions according to the extension-based definitions coincide with the *in* sets of the corresponding T-d-labellings [4].

2.2 Structured argumentation frameworks in *ASPIC*$^+$

The *ASPIC*$^+$ framework concretises abstract argumentation frameworks by defining the notions of argument and defeat. In its usual formulation it directly generates abstract argumentation frameworks, after which the theory of such frameworks can be used to evaluate the arguments. However, for present purposes it is convenient to work with a formulation in terms of recursive labellings, inspired by similar constructs in [14]. This formulation was in [16] shown to be equivalent to the usual formulation of *ASPIC*$^+$. Using the recursive-labelling form will allow to adapt ideas from [8, 10].

ASPIC$^+$ defines abstract argumentation systems as structures consisting of a logical language $\mathcal{L}$ with negation and two sets $\mathcal{R}_s$ and $\mathcal{R}_d$ of strict and defeasible inference rules defined over $\mathcal{L}$. In the present paper the specification of $\mathcal{L}$ will be left implicit in the examples. Arguments are constructed from a knowledge base (a subset of $\mathcal{L}$) by chaining inferences over $\mathcal{L}$ into acyclic graphs (which are trees if no premise is used more than once). Formally (see for detailed illustrations of the definitions the previous publications on *ASPIC*$^+$, such as [11–13]):

Definition 2.2. [Argumentation System] An *argumentation system* (AS) is a pair $AS = (\mathcal{L}, \mathcal{R})$ where:

- $\mathcal{L}$ is a logical language consisting of propositional or ground predicate-logic literals
- $\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_d$ is a set of strict ($\mathcal{R}_s$) and defeasible ($\mathcal{R}_d$) inference rules of the form $\{\varphi_1, \ldots, \varphi_n\} \rightarrow \varphi$ and $\{\varphi_1, \ldots, \varphi_n\} \Rightarrow \varphi$ respectively (where φ_i, φ are meta-variables ranging over wff in $\mathcal{L}$), such that $\mathcal{R}_s \cap \mathcal{R}_d = \emptyset$. $\varphi_1, \ldots, \varphi_n$ are called the *antecedents* and φ the *consequent* of the rule.[2]

Informally, $n(r)$ is a well-formed formula (wff) in $\mathcal{L}$ which says that the defeasible rule $r \in \mathcal{R}$ is applicable, so that an argument claiming $\neg n(r)$ attacks an inference step in the argument using r. We write $\psi = -\varphi$ just in case $\psi = \neg\varphi$ or $\varphi = \neg\psi$. Also, for any rule r the *antecedents* and *consequent* are denoted, respectively, with $\mathtt{ant}(r)$ and $\mathtt{cons}(r)$.

Definition 2.3. [**Consistency**] For any $S \subseteq \mathcal{L}$, let the *closure of S under strict rules*, denoted $Cl_{R_s}(S)$, be the smallest set containing S and the consequent of any strict rule in $\mathcal{R}_s$ whose antecedents are in $Cl_{R_s}(S)$. Then a set $S \subseteq \mathcal{L}$ is *directly consistent* iff $\nexists\, \psi, \varphi \in S$ such that $\psi = -\varphi$, and *indirectly consistent* iff $Cl_{R_s}(S)$ is directly consistent.

Definition 2.4. [**Knowledge bases**] A *knowledge base* in an $AS = (\mathcal{L}, \mathcal{R}, n)$ is an indirectly consistent set $\mathcal{K} \subseteq \mathcal{L}$.

In this paper $\mathcal{K}$ corresponds to the 'necessary premises' in other *ASPIC*$^+$ publications, which are intuitively certain and therefore not attackable and, moreover, jointly indirectly consistent. We will represent what intuitively are uncertain premises φ as defeasible rules $\Rightarrow \varphi$. This allows us to ignore undermining attack, i.e., attacks on premises, for simplicity. Generalisation of the new approach to the full case with undermining attack is entirely straightforward.

An argument is now formally defined as follows.

Definition 2.5. [**Arguments**] A *argument* A on the basis of a knowledge base $\mathcal{K}$ over an argumentation system AS is a structure obtainable by applying one or more of the following steps finitely many times:

(1) φ if $\varphi \in \mathcal{K}$ with: $\mathtt{Prem}(A) = \{\varphi\}$; $\mathtt{Conc}(A) = \varphi$; $\mathtt{Sub}(A) = \{\varphi\}$; $\mathtt{ImmSub}(A) = \emptyset$; $\mathtt{DefRules}(A) = \emptyset$; $\mathtt{Rules}(A) = \emptyset$; $\mathtt{DefRules}(A) = \emptyset$; $\mathtt{TopRule}(A)$ = undefined.
(2) $A_1, \ldots, A_n \rightarrow \psi$ if $A_1, \ldots, A_n$ are arguments such that $\psi \notin \mathtt{Conc}(\{A_1, \ldots, A_n\})$ and $\mathtt{Conc}(A_1), \ldots, \mathtt{Conc}(A_n) \rightarrow \psi \in \mathcal{R}_s$ with:
$\mathtt{Prem}(A) = \mathtt{Prem}(A_1) \cup \ldots \cup \mathtt{Prem}(A_n)$;
$\mathtt{Conc}(A) = \psi$;
$\mathtt{Sub}(A) = \mathtt{Sub}(A_1) \cup \ldots \cup \mathtt{Sub}(A_n) \cup \{A\}$;
$\mathtt{ImmSub}(A) = \{A_1, \ldots, A_n\}$,
$\mathtt{Rules}(A) = \mathtt{Rules}(A_1) \cup \ldots \cup \mathtt{Rules}(A_n) \cup \{\mathtt{Conc}(A_1), \ldots, \mathtt{Conc}(A_n) \rightarrow \psi\}$;
$\mathtt{DefRules}(A) = \mathtt{Rules}(A) \cap \mathcal{R}_d$;
$\mathtt{TopRule}(A) = \mathtt{Conc}(A_1), \ldots, \mathtt{Conc}(A_n) \Rightarrow \psi$.
(3) $A_1, \ldots, A_n \Rightarrow \psi$ if $A_1, \ldots, A_n$ are arguments such that $\psi \notin \mathtt{Conc}(\{A_1, \ldots, A_n\})$ and $\mathtt{Conc}(A_1), \ldots, \mathtt{Conc}(A_n) \Rightarrow \psi \in \mathcal{R}_d$, with the other notions defined as in (2).

Argument A is *strict* if $\mathtt{DefRules}(A) = \emptyset$ and *defeasible* otherwise.

Each of the functions $\mathtt{Func}$ in this definition is also defined on sets of arguments $S = \{A_1, \ldots, A_n\}$ as follows: $\mathtt{Func}(S) = \mathtt{Func}(A_1) \cup \ldots \cup \mathtt{Func}(A_n)$. Note that the $\Rightarrow$ symbol is overloaded to denote an argument while it also denotes defeasible inference rules.

Arguments can be attacked in two ways: on a defeasibly derived conclusion (rebutting attack) and on a defeasible inference step

[2] Below the brackets around the antecedents will be omitted if there is no danger for confusion.

(undercutting attack). While up to now the definitions were the same as in standard *ASPIC*$^+$, a difference is that in the recursive-labelling version of *ASPIC*$^+$ arguments can only be attacked on their final conclusion or inference, while the recursion will take care of attack on a subargument. This change also induces some changes in other definitions of standard *ASPIC*$^+$. In the following definitions several notions are preceded by a p (for 'Pollock') to distinguish them from the original notions of *ASPIC*$^+$ in e.g. [13].

Definition 2.6. **[p-Attack]** *A p-attacks B* iff *A p-undercuts*, *p-rebuts* or *p-undermines B*, where:

• *A p-undercuts* argument *B* iff $\mathsf{Conc}(A) = -n(r)$ and *B* has a defeasible top rule *r*.

• *A p-rebuts* argument *B* iff $\mathsf{Conc}(A) = -\mathsf{Conc}(B)$ and *B* has a defeasible top rule.

The notion of *p-defeat* can then be defined as follows. First, p-undercutting attacks succeed as *p-defeats* independently of preferences over arguments, since they express exceptions to defeasible inference rules. Next, p-rebutting attacks succeed only if the attacked argument is not stronger than the attacking argument (as usual, $A \prec B$ is defined as $A \preceq B$ and $B \npreceq A$ while $A \approx B$ is defined as $A \preceq B$ and $B \preceq A$).

Definition 2.7. **[p-Defeat]** *A p-defeats B* iff: *A* p-undercuts *B*, or; *A* p-rebuts *B* and $A \nprec B$.

Argumentation systems plus knowledge bases form argumentation theories, which induce p-structured argumentation frameworks.

Definition 2.8. A *p-structured argumentation framework (pSAF)* defined by an argumentation theory $AT = (AS, KB)$ is a triple $\langle \mathcal{A}, C^p, \preceq \rangle$ where $\mathcal{A}$ is the set of all arguments constructed from *KB* in *AS*, $\preceq$ is an ordering on $\mathcal{A}$, and $(X, Y) \in C^p$ iff *X* p-attacks *Y*.

Abstract argumentation frameworks are then generated from *pSAFs* as follows:

*Definition 2.9 (**p-Argumentation frameworks**).* A *p-abstract argumentation framework (pAF) corresponding to a pSAF* = $\langle \mathcal{A}, C^p, \preceq \rangle$ is a pair $(\mathcal{A}, D)$ such that *D* is the p-defeat relation on $\mathcal{A}$ determined by *pSAF*.

Argument evaluation is then recursively defined as follows:

Definition 2.10. [p-labellings.]

(1) A pair of subsets (*In*,*Out*) of $\mathcal{A}$ is a *p-labelling* of a *pAF* iff $In \cap Out = \emptyset$ and for all $A \in \mathcal{A}_{pAF}$ it holds that:
  (a) *A* is labelled *In* iff:
    (i) All arguments that p-defeat *A* are labelled *Out*; and
    (ii) All members of $\mathsf{ImmSub}(A)$ are labelled *In*; and
  (b) *A* is labelled *Out* iff:
    (i) *A* is p-defeated by an argument that is labelled *In*; or
    (ii) A member of $\mathsf{ImmSub}(A)$ is labelled *Out*.

Then *stable semantics* p-labels all arguments, while *grounded semantics* minimises and *preferred semantics* maximises the set of arguments that are p-labelled *in* and *complete semantics* allows any p-labelling.

A link with standard *ASPIC*$^+$ can be made by observing that *A* defeats *B* in standard *ASPIC*$^+$ just in case *A* p-defeats *B* or a proper subargument of *B*. Given this link, it has been shown that the *T*-d-labellings of the *AF* corresponding to a *p-SAF* given the defeat relation coincide with the *T*-p-labellings of the *pSAF* given the p-defeat relation[16].

3 THE ACCRUAL MODELS OF PRAKKEN (2005) AND GORDON (2016,2018)

In this section I critically examine the previous inference-based accounts of accrual in [15] and [8, 10]. I first recall the three principles of accrual proposed in [15].

(1) An accrual is sometimes weaker than its accruing elements, due to the possibility that accruing reasons are not independent. For example, if multiple witnesses who all say the same are from a group of people who are more likely to confirm each other's statements when these statements are false than when they are true, then the accrual of the witness testimony arguments will be weaker than the individual accruing arguments [14].

(2) An accrual makes its elements inapplicable. This is for the same reason: if John and Mary both testify that the same statement is true but the accrual is weaker than their individual testimonies, then no inferences should be drawn from their individual testimonies.

(3) Flawed arguments should not accrue. If an argument is based on grounds that are refuted, it should not enter into the accrual. For example, if John can be shown to make his testimony while hallucinating, his testimony should not enter the accrual with Mary's testimony but should be fully ignored.

3.1 Prakken (2005)

The account of [15] can be directly applied within *ASPIC*$^+$. The main idea is threefold. First, every conclusion drawn by applying a defeasible rule is labelled with the set of conclusions of all arguments to which the rule was applied. Second, a defeasible accrual inference rule is added to $\mathcal{R}_d$, of the following form (in fact, the rule is a scheme for any natural number *i* such that $1 \leq i \leq n$):

$\varphi^{l_1}, \ldots, \varphi^{l_n} \Rightarrow \varphi$ (Accrual)

Finally, the following undercutter scheme is formulated for any *i* such that $1 \leq i \leq n$.

$\varphi^{l_1}, \ldots, \varphi^{l_n} \Rightarrow \neg n(\varphi^{l_1}, \ldots, \varphi^{l_{n-i}} \Rightarrow \varphi)$ (Accrual undercutter)

This undercutter says that when a set of reasons accrues, no proper subset accrues.

This proposal satisfies the three principles of accrual. First, argument orderings are possible according to which an argument applying the Accrual rule to a set *S* is weaker than an argument applying the Accrual rule to a set $S' \subset S$. Second, any argument defeating a subargument of an argument that has the Accrual rule as its top rule also defeats the latter argument. And third, the accrual undercutter makes sure that lesser accruals do not apply if a larger accrual applies. See for more details [15].

On the other hand, this approach has some problems. First, labelling conclusions with sets of premises seems unnatural. More importantly, for any set *S* of reasons for the same conclusion φ, the number of arguments is in the worst case exponential to the cardinality of *S*, since applications of the accrual undercutter can be defeated either directly or on a proper subargument. Finally, on this account arguments with strict top rules cannot enter an accrual, which seems wrong if the argument is defeasible because of defeasible rules applied by subarguments (see [15] for an example). In Section 6 below I will argue that the analysis in [15] is flawed and that the present proposal improves on it.

3.2 Gordon (2018)

In [8, 10] Gordon avoids the first and second weakness of [15] in a new version of the Carneades system [9]. In Carneades an argument has a set of premises, a set of exceptions (in [8, 10] called undercutters) and a conclusion, which is either pro or con a statement. Unlike in *ASPIC*$^+$, all arguments are elementary, that is, they contain a single inference step; they are combined in a recursive definition of and *argument graph*. In the new version all statements in an argument graph are labelled as *in*, *out* or *undecided*. This is achieved by rather involved recursive dependencies between various definitions of concepts like *issues*, *argument weight*, *proof standards*, and *supported* and *applicable* arguments.

The essence of Gordon's new approach to accrual is that an argument can be applicable even if not all its premises are *In*, as long as no premise is undecided. Thus a single argument can be constructed such that it contains all potential reasons for its conclusion, where in general only some of these reasons have to apply for an argument to be *In*. This is how [8] avoids an exponential blow up in the number of arguments to be constructed.

To deal with cycles, Gordon defines in [10] an operator on statement labellings, which in [8] he claims to be monotonic. If this is true, then it has a least fixed point. Accordingly, Gordon defines the set of sceptically acceptable conclusions as the statements that are labelled *In* in the least fixed point of the operator (which he calls the grounded labelling). Moreover, he defines notions of stable and preferred labellings analogously to [6]'s semantics and the p-labellings of *ASPIC*$^+$ defined in the previous section. Being able to do so is valuable, since it uniquely characterises the set of sceptically acceptable conclusions (which agrees with the idea of skeptical acceptance) and it preserves similar relations between the various semantics as in [6], which is desirable for technical reasons that cannot be explained here for reasons of space.

However, there is a problem, since there are counterexamples to the claim in [8] that the operator on statement labellings is monotonic, due to the fat that any proof standard is possible in general. Since listing the definitions of [8] would require several pages, I just list the counterexample here and leave it to the reader to verify it. Consider two arguments a_1 for p and a_2 for $\neg p$, both with no premises or undercutters and with weight > 0.0 where a_1 has higher weight than a_2, and $\{p, \neg p\}$ is an issue and the proof standard for the issue says that a position s satisfies its proof standard iff (1) it has an applicable argument in the current labelling and (2) no other position from the same issue is *In* in the current labelling and (3) some argument for s has higher weight than all arguments for the other positions in the issue that satisfy the previous conditions. Then $F(\emptyset, \emptyset) = (\{p\}, \{\neg p\})$ while $F(\{\neg p\}, \emptyset) = (\{\neg p\}, \{p\})$.

In the next section I will try to retain as much as possible of Gordon's new approach in the context of recursive *ASPIC*$^+$, while avoiding its technical problems.

4 THE NEW PROPOSAL

In this section I modify the recursive version of *ASPIC*$^+$ in order to capture accrual of arguments as a logical operation that overcomes the weaknesses of [15] and [8, 10]. Henceforth I write 'p-defeat' as 'defeat' for simplicity.

The new model of accrual is based on several key ideas. First, arguments for the same conclusion are collected in a so-called accrual set. A key idea here is that accrual sets are defined relative to a labelling of the set $\mathcal{A}$ of arguments, to be able to express that only arguments of which no immediate subargument is *Out* and no undercutter is *In* enter the accrual set. This will help satisfying the third principle of accrual. Note that the labelling referred to in the next definition does not have to be a d-labelling but any any partition of the set of arguments into subsets of arguments labelled *In*, or *Out*, or not labelled.

Definition 4.1. [**Accrual sets**] Given a labelling $l = (In, Out)$ of $\mathcal{A}$, a nonempty set $S \subseteq \mathcal{A}$ is *accrual set* of $\varphi \in \mathcal{L}$, denoted as $s_l(\varphi)$ iff it satisfies both of the following conditions:

1. if $A \in s_l(\varphi)$, then no undercutter of A is *In* and no argument in $\texttt{ImmSub}(A)$ is *Out*;
2. if $\texttt{Conc}(A) = \varphi$ and all undercutters of A are *Out* and all arguments in $\texttt{ImmSub}(A)$ are *In*, then $A \in s_l(\varphi)$.

Let us, given a labelling $l = (In, Out)$ of $\mathcal{A}$, call any argument A such that none of its undercutters are *In* and no argument in $\texttt{ImmSub}(A)$ is *Out weakly applicable in l* and call any argument A such that all its undercutters are *Out* and all arguments in $\texttt{ImmSub}(A)$ are *In* i *strongly applicable in l*. We then see that any accrual set of φ contains all strongly applicable arguments for φ (by condition 2) plus possibly some weakly applicable arguments for φ (by condition 1). The reason for this complication will become clear later on.

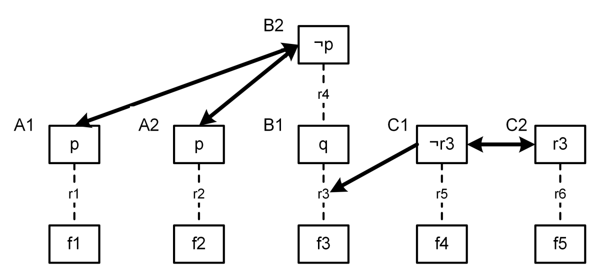

Figure 1: Running example

Example 4.2. [**running example**] Let us illustrate the definitions in this section with the following formal example. Let $\mathcal{K} = \{f_1, f_2, f_3, f_4, f_5\}$, $\mathcal{R}_s = \emptyset$ and $\mathcal{R}_d = \{f_1 \Rightarrow_{r1} p; f_2 \Rightarrow_{r2} p; f_3 \Rightarrow_{r3} q; q \Rightarrow_{r4} \neg p; f_4 \Rightarrow_{r5} \neg r_3; f_5 \Rightarrow_{r6} r_3\}$ (with the rule names according to n attached to the arrows). The following arguments can be constructed (the premise arguments will below be left implicit).

A_1:	$f_1 \Rightarrow p$	A_2:	$f_2 \Rightarrow p$
B_1:	$f_3 \Rightarrow q$	B_2:	$B_1 \Rightarrow \neg p$
C_1:	$f_4 \Rightarrow \neg r_3$	C_2:	$f_5 \Rightarrow r_3$

These arguments are displayed in Figure 1, together with their attack relations: C_1 undercuts B_1, while all symmetric defeat relations are rebuttals. Let l_1 be the empty labelling. Then p has three accrual sets, namely, $\{A_1\}$, $\{A_2\}$ and $\{A_1, A_2\}$, while all other conclusions have singleton sets as accrual sets. Next let l_2 be such that C_1 is *In* while C_2, B_1 and B_2 are *Out* and f_1 and f_2 are undecided. Then $\{A_1, A_2\}$ is the only accrual set for p while $\neg p$ does not have an accrual set.

A second key idea is that the argument preference relation $\leq$ is now redefined as a relation on the powerset of $\mathcal{A}$. It is assumed to satisfy some properties in order to handle cases with strict-and-firm arguments in an intuitive way. This induces the notion of an accrual argumentation framework as follows.

*Definition 4.3 (**l-Argumentation frameworks**).* An *accrual argumentation framework* (*aSAF*) defined by an $AT = (AS, KB)$ is a tuple $\langle \mathcal{A}, C^p, \leq \rangle_{AT}$ where

(1) $\mathcal{A}$ is the set of all arguments on the basis of KB in AS;
(2) C^p is the p-attack relation on $\mathcal{A}$;
(3) $\leq$ is a reflexive binary relation on $2^{\mathcal{A}}$ satisfying:
  (a) $S \leq S'$ for any S' that contains a strict- and firm argument.
  (b) $S \not\leq S'$ for any S that contains a strict- and firm argument.

A third key idea of the new approach is to make the notion of defeat relative to a labelling, so that in case of rebutting arguments only those arguments are compared that have no undercutters that are *In* and immediate subarguments that are *Out*. This will also help satisfying the third principle of accrual.

*Definition 4.4 (**l-defeat**).* Given a labelling $l = (In, Out)$ of the set $\mathcal{A}$ of an *aSAF*, argument A *l-defeats* argument B iff:

(1) A undercuts B; or
(2) A rebuts B and for some $s_l(\mathsf{Conc}(A))$ such that $s_l(\mathsf{Conc}(A))$ and some $s_l(\mathsf{Conc}(B))$ it holds that $s_l(\mathsf{Conc}(A)) \not\prec s_l(\mathsf{Conc}(B))$.

Note that if there is nonempty accrual set for a conclusion of a rebutted argument, then the argument will be *Out* anyway, so its defeat relations are irrelevant.

Example 4.5. [**running example ct'd**] Suppose first that $\{A_1\} \prec \{B_2\}$, $\{A_2\} \prec \{B_2\}$ and $\{B_2\} \prec \{A_1, A_2\}$. Assume first a labelling l in which all arguments are undecided. Then p and $\neg p$ have the same accrual sets as for l_1 in Example 4.2. Then both A_1 and A_2 l-defeat B_2 since $\{B_2\} \prec \{A_1, A_2\}$ but B_2 also l-defeats A_1 since $\{A_1\} \prec \{B_2\}$ and B_2 l-defeats A_2 since $\{A_2\} \prec \{B_2\}$. Assume next a labelling l' in which all premise arguments are *In* while the other arguments are all undecided. Then $\{A_1, A_2\}$ is the only accrual set for p, so now both A_1 and A_2 strictly l'-defeat B_2 since $\{B_2\} \prec \{A_1, A_2\}$. Assume next that the preference relations between all above argument sets are reversed and consider again labelling l'. Then B_2 strictly defeats both A_1 and A_2 because of $\{A_1, A_2\} \prec \{B_2\}$, even though both $\{B_2\} \prec \{A_1\}$ and $\{B_2\} \prec \{A_2\}$. The reason is that $\{A_1, A_2\}$ is still the only accrual set for p.

Finally, inspired by [8], I define a function on labellings and define the new semantics in terms of the fixpoints of this function.

Definition 4.6. [l-labellings.] The *characteristic function F* of an *aSAF* is a total function on the set of all labellings $l(In, Out)$ of $\mathcal{A}$ that returns a labelling $l' = (In', Out')$ of $\mathcal{A}$ according to the following conditions: for all $A \in \mathcal{A}$ it holds that:

(1) $A \in In'$ iff:
  (a) all arguments that l-defeat A are in *Out*; and
  (b) all members of $\mathtt{ImmSub}(A)$ are in *In*; and
(2) $A \in Out'$ iff:
  (a) A is l-defeated by an argument in *In* or
  (b) a member of $\mathtt{ImmSub}(A)$ is in *Out*.

An *l-labelling* of an *aSAF* is any fixpoint of F.

Example 4.7. [**running example ct'd**] Consider first the initial argument ordering of Example 4.7 plus $\{C_1\} \approx \{C_2\}$ and consider the empty labelling l. Then $F(l) = l'$ makes all premise arguments *In* and leaves the other arguments undecided. Then $l'' = F(l')$ also makes A_1 and A_2 in since they have no l' defeaters since they both strictly l'-defeat B_2. So B_2 is *Out* in l'' while the other arguments that were undecided in l' are still undecided in l'. Then $F(l'') = l''$ so l'' is the grounded labelling. There are two preferred labellings: the one extends l'' by making C_2 and B_1 *In* and C_1 *Out* while the other extends l'' by making C_1 *In* and C_2 and B_1 *Out*.

Consider next the second argument ordering of Example 4.7 plus $\{C_1\} \approx \{C_2\}$. Then the grounded labelling equals l' since, although B_2 strictly l'-defeats both A_1 and A_2, its subargument B_2 is undecided. There are two preferred labellings. The first extends l' by making C_2 and B_1 *In* and C_1 *Out* and therefore makes B_2 *In* and both A_1 and A_2 *Out*. The second makes C_1 *In* and C_2 and B_1 and B_2 *Out* so it also makes A_1 and A_2 *In*.

It is easy to see that the F operator always returns a d-labelling in the sense of Definition 2.1. We would now like to define the notions of complete, stable, preferred and grounded l-labellings as the corresponding p-labellings were defined in Section 2.1. However, before this can be done in a way that preserves the relations between these semantics as in Section 2.1, some technical work has to be done, which will be the topic of the next subsection.

Example 4.8. Consider:

A:	$\Rightarrow p$	B:	$\Rightarrow \neg p$
C:	$\Rightarrow r$	D:	$C \Rightarrow \neg p$
E:	$\Rightarrow \neg r$		

Suppose $\{A\} \prec \{B\}$ while $\{B, D\} \prec \{A\}$ and $\{C\} \approx \{E\}$ (this is an example of an accrual that is weaker than some of its elements).

I first show why why weakly applicable arguments should be allowed to enter accrual sets. Suppose otherwise; then the only accrual set for $\neg p$ is $\{B\}$ and since $\{A\} \prec \{B\}$, we have that $F(\emptyset, \emptyset) = (\{B\}, \{A\})$ and $F(\{B\}, \{A\}) = (\{B\}, \{A\})$. But we also have that $F(\{A, C\}, \{B, D, E\}) = (\{A, C\}, \{B, D, E\})$, so F is not monotonic and we can have multiple grounded extensions.

I next explain why weakly applicable arguments do not *have* to be in an accrual set and why there can be multiple accrual sets. In other words, in defining defeat we should, for each of the two conclusions involved in the rebuttal, alternatively consider all possible extensions of the set of strongly applicable arguments with weakly applicable arguments. The reason is that if we only compare the two sets with all such arguments, then in the current example, if we start with the empty labelling to obtain the grounded extension, we have to compare $\{A\}$ with $\{B, D\}$ only, so the outcome is that A is in the grounded l-labelling while B and D are not. This is undesirable since the fact that C is only weakly applicable indicates that a reasonable evaluator could decide to adopt E and thereby reject D, in which case the comparison would only be between A and B. With the current definitions the outcome is that the grounded l-labelling is empty. Note also that otherwise (if the only accrual set for B is $s_l(B) = \{B, D\}$) the characteristic function F is again not monotonic, since then $F(\emptyset, \emptyset) = (\{A\}, \{B, D\})$ while $F(\{E\}, \{C\}) = (\{B, E\}, \{A, C, D\})$.

I finally show how the new proposal deals with strict arguments, i.e., arguments that cannot be attacked. First, since $\mathcal{K}$ is assumed to be indirectly consistent, it cannot be that two accrual sets for arguments with jointly contradictory conclusions both contain a strict argument, since otherwise the empty set is indirectly inconsistent, so $\mathcal{K}$ is indirectly inconsistent. Next, if $s_l(A)$ contains a strict argument A' and B's conclusion contradicts A's conclusion, then $s_l(B) \prec s_l(A)$ by condition (3) of Definition 4.1. So all arguments in $s_l(B)$, which are all defeasible by assumption that $\mathcal{K}$ is indirectly consistent, are defeated by A' and are *Out* in $F(l)$ since A' has no defeaters. Moreover, any defeasible member of $s_l(A)$ is *In* in $F(l)$ unless it has an undercutter that is not *Out* in $F(l)$. In sum, the new proposal arguably treats strict arguments in an intuitively acceptable way.

5 FORMAL INVESTIGATION OF THE NEW PROPOSAL

In this section I investigate the formal properties of the new proposal. First, we would like that the new semantics preserves the Dungean relations between the semantics and, second, we would like that if each statement has at most one argument, then the new definitions reduce to the old ones. I now address both issues in turn.

5.1 Verifying the Dungean relations between the semantics

What is to be proven is that the grounded l-labelling is unique and included in any other l-labelling, that every stable l-labelling is preferred and that every grounded, stable or preferred l-labelling is complete. I first prove that the characteristic function of an *aSAF* is monotonic. To this end I first define an ordering on pairs of subsets of $\mathcal{A}$. Overloading the symbol $\subseteq$ I say that $(S_1, S_2) \subseteq (T_1, T_2)$ iff $S_1 \subseteq T_1$ and $T_1 \subseteq T_2$. I state without proof that this ordering on pairs of sets is a complete partial order, as required for applying fixpoint theory to conclude that F is monotonic. Then the following lemma can be proven.

Lemma 5.1. *For any pair l and l' of labellings such that $l \subseteq l'$ it holds that if A l'-defeats B then A l-defeats B.*

Proof. Suppose A l'-defeats B. If A undercuts B then A also l-defeats B since undercutting defeat does not depend on l' or l. If A rebuts B then there exist pairs $s_{l'}(\mathsf{Conc}(A))$ and $s_{l'}(\mathsf{Conc}(B))$ such that $s_{l'}(\mathsf{Conc}(A)) \nprec s_{l'}(\mathsf{Conc}(B))$. We prove that these sets are also accrual sets for A and B given l. For any argument C in any of these two sets: if $C \notin In' \cup Out'$ then $C \notin In \cup Out$, if $C \in In'$ then $C \notin Out$ and if $C \in Out'$ then $C \notin In$. So if C is strongly applicable in l' then C is weakly or strongly applicable in l, and if C is weakly applicable in l' then C is also weakly applicable in l. Moreover, since $l \subseteq l'$, all arguments for $\mathsf{Conc}(A)$, respectively, $\mathsf{Conc}(B)$ that are strongly applicable in l are also strongly applicable in l', so they are in $s_{l'}(\mathsf{Conc}(A))$, respectively, $s_{l'}(\mathsf{Conc}(B))$, so they are in $s_l(\mathsf{Conc}(A))$, respectively, $s_l(\mathsf{Conc}(B))$. Hence $s_{l'}(\mathsf{Conc}(A))$ and $s_{l'}(\mathsf{Conc}(B))$ are also accrual sets according to l. And since the preference ordering on $2^{\mathcal{A}}$ does not depend on labellings, we have that A also l-defeats B. □

It can now be shown that F is monotonic.

Proposition 5.2. *For any pair of labellings l_1 and l_2 such that $l_1 \subseteq l_2$ it holds that $F(l_1) \subseteq F(l_2)$.*

Proof. Let $l_1 = (In_1, Out_1)$ and $l_2 = (In_2, Out_2)$ and let $F(l_1) = (In'_1, Out'_1)$ and $F(l_2) = (In'_2, Out'_2)$. Suppose $A \in In'_1$. Then all l_1-defeaters of A are in Out_1, so they are also in Out_2. Moreover, by Lemma 5.1 there are no l_2-defeaters of A that are not also l_1-defeaters of A. So all l_2-defeaters of A are in Out_2, so $A \in In'_2$.

Suppose next $A \in Out'_1$. Then some l_1-defeater of A is in In_1, so it is also in In_2. But then $A \in Out'_2$. □

Because of this result the *grounded l-labelling* can be defined as the smallest l-labelling, a *complete l-labelling* as any l-labelling, a *stable l-labelling* as a total l-labelling and a *preferred l-labelling* as any maximal l-labelling. The original relations between these semantics are then preserved, as well as the guarantee of existence of complete and preferred l-labellings (recall that $\subseteq$ on $2^{\mathcal{A}}$ is a complete partial order).

5.2 Reduction to *ASPIC*$^+$ if no proper accruals

The next thing to prove is that if there are no proper accruals, then everything reduces to the original *ASPIC*$^+$ semantics. For reasons of space I have to confine myself to stating the two lemmas on which the proof depends. First I make the obvious assumption that $\{A\} \preceq \{B\}$ just in case $A \preceq B$. This assumption immediately yields the following lemma.

Lemma 5.3. *If $s_l(\mathsf{Conc}(A)) = \{A\}$ and $s_l(\mathsf{Conc}(B)) = \{B\}$ then if A l-defeats B then A defeats B.*

Another useful Lemma follows from the proof in [16] of the equivalence of d-labellings and p-labellings for *ASPIC*$^+$.

Lemma 5.4. *For any p-labelling (l-labelling): $A \in Out$ iff some subargument of A is defeated (l-defeated) by an argument that is In.*

With these lemmas the proof of the following proposition is straightforward and has a similar structure as the proof of Proposition 5.2.

Proposition 5.5. *Let aSAF be such that each element from $\mathcal{L}$ is the conclusion of at most one argument and let pSAF be obtained from aSAF by letting $A \preceq B$ just in case $\{A\} \preceq \{B\}$. Then:*

(1) *every l-labelling of aSAF is a p-labelling of pSAF;*
(2) *every p-labelling of pSAF is an l-labelling of aSAF*

5.3 Verifying the principles of accrual

I next verify that the new proposal satisfies [15]'s three principles of accrual. First, accruals can be weaker than their elements, since no assumptions are made about the preference relation on $2^{\mathcal{A}}$. See Example 4.8 for an illustration.

Next, that an accrual makes its elements inapplicable can most clearly be seen if all arguments in the relevant accrual sets are strongly applicable. Then for each of the two conclusions involved in a rebuttal there are unique accrual sets. Thus it is ensured that if a 'larger' accrual set is too weak, then the conclusion cannot be saved by a stronger 'smaller' accrual set. see again Example 4.8 for an illustration. More generally, the second principle is respected by requiring that all strongly applicable arguments for a conclusion are in the accrual set for that conclusion. At first sight it would seem that the possibility to include weakly applicable arguments in an accrual set violates this principle. However, above we have seen why this is still a good idea. Moreover, in preferred and stable l-labellings the set of undecided arguments is minimal.

Finally, the principle that flawed arguments may not accrue is respected by leaving arguments of which an undercutter is *In* or an immediate subargument is *Out* outside any accrual set.

6 APPLICATIONS

6.1 Accrual of strict reasons

First I revisit an example from [15] to illustrate how the present proposal overcomes a weakness of the proposal in [15].

Example 6.1. Suppose in a criminal case the issue is whether the suspect was in Amsterdam on the morning of October 22nd, 2004 and consider a witness Albert who testifies he saw the suspect in the Vondelpark at 9.00am that day and a witness Bill who testifies that he saw the suspect at Central Station at 10.30am that day. When combined with some obvious background knowledge, both statements on their own deductively imply that the subject was in Amsterdam that morning. Therefore, the two witness testimonies cannot be combined in an accrual for the latter proposition, since strict inferences do not accrue. Schematically:

A: *testifies*$(a, V(s, 9.00))$, so presumably $V(s, 9.00)$, so deductively $A(s, \textit{morning})$

B: *testifies*$(b, S(s, 10.30))$, so presumably $S(s, 10.30)$, so deductively $A(s, \textit{morning})$

In [15] Prakken wrote:

> At first sight, the result that the witness testimonies do not accrue with respect to the proposition *morning* would seem counterintuitive. Yet the principle that accrual of strict arguments does not make sense is beyond doubt so that the intuition must be at fault.

However, the principle that accrual of strict arguments does not make sense is only beyond doubt for strict-and-firm arguments, so for unattackable arguments. Indeed, on the present account, the two witness testimonies accrue for the conclusion that the suspect was in Amsterdam. Suppose a third witness Cathy says that the suspect was not in Amsterdam that morning:

C: *testifies*$(c, \neg A(s, \textit{morning}))$, so presumably $\neg A(s, \textit{morning})$.

There is one accrual set for A, which is also the unique accrual set for B, namely, $\{A, B\}$. Note that both A and B rebut C but not conversely. To see whether these rebutting attacks succeed as defeats, $\{A, B\}$ has to be compared with $\{C\}$.

6.2 Arguing about legislative proposals

In [17] a debate about a legislative proposal was modelled in *ASPIC*$^+$ with two schemes of defeasible inference rules (which themselves are also schemes). For good consequences a set of defeasible rules of the following form was assumed for each $n \geq 1$:

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

Therefore (presumably), action A is good.

A similar scheme was assumed for bad consequences. Thus in fact the KR approach to accrual was applied, which as noted above suffers from the same exponential blow-up as [15]'s inference approach. I now sketch how the same example can be modelled in the present approach.

The debate is about a proposal of a previous Dutch government to impose mandatory minimum sentences for recidivism in case of serious crimes. Three good consequences were opposed to two bad consequences of the proposed bill. The good consequences were GC1: the bill will reduce crime; GC2: the bill is meeting the call from the public for more severe sentences; GC3: the bill has the symbolic effect of underlining norms. The bad consequences were BC1: the bill will harden the relationship between the judiciary and the legislator; BC2: the bill will force judges to impose severe sentences for not so serious cases. In addition, the argument for GC1 was attacked on its causal premise that the bill will reduce crime and the argument for GC2 was attacked on its evaluative premise that meeting the call from the public for more severe sentences is good. Schematically:

GC_1: *LessCrime*, *LessCrimeGood* $\Rightarrow$ *BillGood*
GC_2: *MeetsCall*, *MeetsCallGood* $\Rightarrow$ *BillGood*
GC_3: *UnderlinesNorms*, *UnderlinesNormsGood* $\Rightarrow$ *BillGood*

BC_1: *HardensRelation*, *HardensRelationBad* $\Rightarrow \neg$*BillGood*
BC_2: *ForcesJudges*, *ForcesJudgesBad* $\Rightarrow \neg$*BillGood*

C $\ldots \Rightarrow \neg$*LessCrime*
E $\ldots \Rightarrow \neg$*MeetsCallGood*

Let us denote the premise subarguments of GC_1 with GC_{1a} and GC_{1b} likewise with the other arguments. While in the approach of [17] 6 arguments for *BillGood* and 3 arguments for $\neg$*BillGood* have to be considered, now only 3, respectively 2 arguments have to be considered. In addition, there is now, unlike in [17], no need for undercutters of lesser accruals. The contents of the accrual sets for *BillGood* and $\neg$*BillGood* depend on how the two premise attacks are resolved. Let us for ease of illustration order the sets of arguments for these conclusions in terms of their number of element.

To construct the grounded labelling, let l be the empty labelling. Then all nonempty subsets of $\{GC_1, GC_2, GC_3\}$ are an accrual set for `BillGood` and all nonempty subsets of $\{BC_1, BC_2\}$ are an accrual set for $\neg$`BillGood`. So any argument for *BillGood* l-defeats any argument for $\neg$*BillGood* and vice versa. If both C and E are strictly preferred over their target, then C strictly l-defeats GC_1 by p-rebutting its first subargument, while E strictly l-defeats GC_2 by p-rebutting

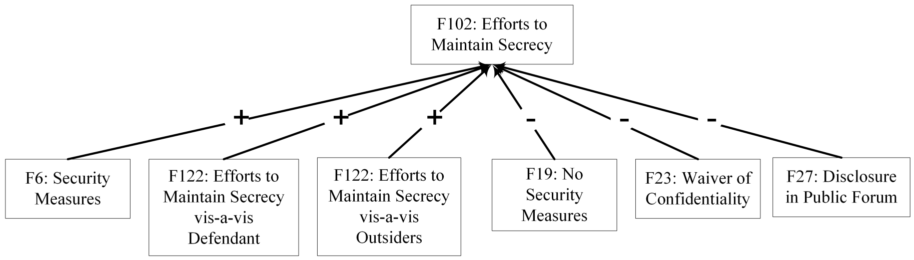

Figure 2: Partial CATO factor hierarchy

its second subargument. So $l' = F(l)$ makes $C, E, G, GC_{1b}, GC_{2a}, GC_{3a}, GC_{3b}, BC_{1a}, BC_{1b}, BC_{2a}, BC_{2b}$ *In*, it makes $GC_{1a}, GC_1, GC_{3b}, GC_3$ *Out* and it leaves $\{GC_3, BC_1, BC_2\}$ undecided. Then both accrual sets are uniquely determined relative to l: we have $s'_l(\texttt{BillGood}) = \{GC_3\}$ and $s'_l(\neg\texttt{BillGood}) = \{BC_1, BC_2\}$. So both BC_1 and BC_2 strictly l-defeat GC_3. So $l'' = F(l')$ equals l' except that B_1 and B_2 moves from undecided to *In* and GC_3 moves from undecided to *Out*. Then $F(l'') = l''$ so l'' is the grounded labelling. So $\neg\texttt{BillGood}$ is justified in grounded semantics.

To illustrate credulous reasoning in grounded semantics, assume now that $\{C\}$ is still strictly preferred over $\{GC_{1a}\}$ but that $\{E\} \approx \{GC_{2b}\}$. Then what changes is that E and GC_{2b} defeat each other, so $l' = F(l)$ now also leaves these two arguments undecided. This induces an additional $s'_l(\texttt{BillGood})$ besides $\{GC_3\}$, namely, $\{GC_2, GC_3\}$. Then both of GC_2, GC_3 defeat both of BC_1, BC_2 and vice versa, so $F(l') = l'$ and both BillGood and its negation are defensible in grounded semantics. Moreover, we have that the following two labellings are maximal fixpoints of the operator on labellings: the first one is l extended with GC_2, GC_3 in *In* and BC_1, BC_2 in *Out* while the second is with GC_2, GC_3 in *Out* and BC_1, BC_2 in *In*. The same grounded and preferred labellings are obtained if GC_{2a} strictly l- and $F(l)$-defeats E, although then there is a unique accrual set for BillGood, namely, $\{GC_2, GC_3\}$.

6.3 Factor-based reasoning

Factor-based reasoning as modelled in, for instance, CATO [2], is a natural application domain for models of accrual. In factor-based domains there are no clear rules but only factors pro and con a conclusion, which can in different cases occur in different constellations. Then sets of factors pro and con must be weighed, which can be seen as argument accrual. Below I sketch how the present approach can be used to model this weighing process.

In [1] Al-Abdulkarim et al. model (among other things) reasoning with CATO-style factor hierarchies in abstract dialectical frameworks (ADFs, see [5]). Briefly, ADFs are directed graphs where the nodes represent statements and the links between nodes can be typed. Al-Abdulkarim et al. consider the case where links either *support* or *attack* a statement. Each statement in an ADF has an acceptance condition defined in terms of its parents. In general, ADFs allows cycles, which requires fixpoint semantics to deal with them. For example, if there can be con links from n to n' and vice versa while the acceptance condition of n is that n' is rejected while the acceptance condition of n' is that n is accepted.

Al-Abdulkarim et al. represent factor hierarchies in ADFs in the obvious way where each (base or intermediate) factor and the top-level statement is an ADF-statement and each pro or con relation between factors becomes a support or attack link in the ADF. Since factor hierarchies are acyclic, Al-Abdulkarim et al. do not have to use the (rather involved) semantics of ADFs, since for acyclic ADFs all semantics coincide and the status of each node can be computed with a simple bottom-up procedure, starting with the leaves of a factor hierarchy. This allows them to encode acceptance conditions as rules in a Prolog program.

For example, consider a snapshot of a factor hierarchy displayed in Figure 2. On the basis of their opinion on the relative priority between factor sets [1] propose the following acceptance conditions:

Reject F102 if F19
Reject F102 if F23
Reject F102 if F27 and NOT F123 (NOT is negation as failure)
Accept F102 if F6
Accept F102 if F122
Accept F102 if F123
Reject F102

The last clause is the default value of F102 if no other rule applies. The rules are prioritised by the order in which they are listed, with the rule with the highest priority listed first.

Let us now see how this can be represented in the current version of *ASPIC*$^+$. First, each pro, respectively, con link in a factor hierarchy between factors f and f' is represented as a defeasible rule $f \Rightarrow f'$, respectively, $f \Rightarrow \neg f'$. If desired, then a positive or negative default value for a factor f can be specified by including a rule $\Rightarrow f$ or $\Rightarrow \neg f$. Next, the acceptance conditions translate to preferences between accrual sets of arguments with particular top rules. Let us in our example denote any argument for $F102$, respectively, $\neg F102$ as A_x, respectively, B_x, where x is the number of the factor which is the antecedent of the argument's top rule. Let us also for any set S of arguments pro, respectively, con $F102$ denote S^- as any nonempty subset of S and S^+ as any superset with arguments for $F102$, respectively, $\neg F102$ (the latter including the default value for $F102$ if specified). Then the above Prolog program corresponds to the following set of preferences:

$\{A_6, A_{122}, A_{123}\}^- \prec \{B_{19}\}^+$
$\{A_6, A_{122}, A_{123}\}^- \prec \{B_{23}\}^+$
$\{A_6, A_{122}\}^- \prec \{B_{27}\}^+$

$\{B_{27}\} \prec \{A_{123}\}^+$

The default rule $\Rightarrow \neg F102$ can be included in these preferences in such a way that it is relevant only if no other argument for $F102$ or $\neg F102$ can be constructed and that otherwise its inclusion does not change any of the above preference relations.

Al-Abdulkarim et al. claim that their Prolog program with the acceptance conditions is like a rule base in *ASPIC*$^+$. However, we now see that it is instead like a specification of a preference relation on sets of arguments.

Embedding Al-Abdulkarim et al.'s method in the above way in the present version of *ASPIC*$^+$ may have a number of benefits. For example, the method thus also applies to cases with defeat cycles. Also, the preference relation on $\mathcal{A}$ can depend on more than just the arguments' top rules. For instance, in our example preferences may also depend on how strongly a factor pro or con $F102$ is derived (this is impossible in ADFs, in which a statement's acceptability status is fully stated in terms of the status its parents). Also, the method overcomes a limitation of Al-Abdulkarim et al.'s method in that since the Prolog program implicitly expresses a preference between its clauses, in many cases only the winning arguments are constructed. By contrast, in *ASPIC*$^+$ the losing arguments are also constructed, which may, for instance, be beneficial for purposes of explanation. Finally, by embedding the method in *ASPIC*$^+$ any way of reasoning about determining the base factors in a factor hierarchy can be embedded.

6.4 Rules, Principles and Exclusionary Reasons

Not all reasons accrue. Raz [18] distinguished ordinary from exclusionary reasons. While ordinary reasons for or against a conclusion have to be weighed, exclusionary reasons decide the issue on their own and exclude (in *ASPIC*$^+$ terminology undercut) other reasons on the same issue. This distinction is related to Dworkin's [7] distinction between principles and rules. Verheij et al. [20] convincingly argue this this is only a matter of degree. Rules only exclude those reasons that were taken into account when adopting the rule; if new reasons come up, then they must still be weighed together with the reasons provided by the rule.

Let us see whether Verheij et al.'s account can be modelled in the present approach. Assume the issue is whether some legal effect e occurs, that there is one reason p_1 pro e and one reason c_1 con e, and that p_1 is ruled to be an exclusionary reason pro e. This yields the following defeasible rules, the first two of which express reasons, the third is the rule that expresses that p_1 is an exclusionary reason for e and the fourth and fifth are undercutters ensuring that other reasons concerning e besides the rule are not weighed.

r_1: $p_1 \Rightarrow e$
r_2: $c_1 \Rightarrow \neg e$
r_3: $p_1 \Rightarrow \neg r_2$

Let us construct the grounded labelling in case $\mathcal{K} = \{p_1, c_1\}$. With the empty labelling, there are accrual sets for both e and $\neg e$, namely, $\{A_1\}$ where A_1 applies r_1, and $\{A_2\}$ where A_2 applies r_2. Regardless of their relative preference, in $l' = F(l)$ there only is an accrual set for e, namely $\{A_1\}$, since argument A_3 applying r_3 is *In* in l', so A_2 cannot be in an accrual set for e. Then although A_1 and A_2 do not defeat each other, this does not matter since A_2 is *Out* in $l'' = F(l')$ since it is undercut by A_3, so A_1 is *In* in l''.

Assume next that new reasons p_2 pro e and c_2 con e come up. This yields the following additional rules:

r_4: $p_2 \Rightarrow e$
r_5: $c_2 \Rightarrow \neg e$

With the same knowledge base $\mathcal{K} = \{p_1, c_1\}$ the empty labelling yields three accrual sets for e, namely, $\{A_1\}$, $\{A_4\}$ (where A_4 applies r_4) and $\{A_1, A_2\}$, and also three accrual sets for $\neg e$, namely, $\{A_3\}$, $\{A_5\}$ (where A_5 applies r_5) and $\{A_3, A_5\}$. Regardless of their relative preference, in l' the only accrual sets are $\{A_1, A_4\}$ for e (since their immediate subarguments are *In* in $l' = F(l)$) and $\{A_5\}$ for $\neg e$. Then the outcome is decided by the preference relation between these sets. If the original con factor c_1 should also be allowed to enter the new weighing process, then this can be achieved by adding r_6: $c_2 \Rightarrow \neg r_3$, which undercuts the undercutter r_3 in case c_2 holds.

6.5 Clarifying some terminology from argumentation theory

In argumentation theory often a distinction is made between various ways in which multiple 'reasons' can support a conclusion. This is also relevant for legal reasoning; see, for instance, the discussion in Section 3.A.3 of [3] on the various types of support in evidential legal arguments. The terminology varies; here I adopt the terms recently used by [21]. Let us consider the special case of two potential reasons p and q for the same conclusion r. Informally, if both p and q are needed to build a convincing argument for r, then the resulting argument for r is called *linked*, if each reason on its own suffices to build a convincing argument for r, then the argument is called *convergent* while if each of p and q can on its own be used to build an argument but the argument that uses both is more convincing, the argument is called *cumulative*. Cumulative arguments are thus a special case of what in the present paper are called accruals; more precisely, they are the special case in which an accrual is always stronger than its elements. These definitions are only a first attempt and much discussion in the literature is on making them more precise. In [22] formal definitions of these and related concepts are given in the context of *ASPIC*$^+$. However, accruals are in their classification, following [15] regarded as special kinds of linked arguments, namely, as arguments with a top rule corresponding to [15]'s accrual inference scheme (see Section 3.1 above). Since the present paper has replaced [15]'s approach with a new approach, the question arises how accruals, and cumulative arguments in particular, can be formally characterised given the present approach to argument accrual.

First I briefly summarise the essence of [22]'s definitions of linked and convergent arguments. An argument for conclusion r is, relative to two propositions p and q, *linked* if it has a top rule $p, q \rightsquigarrow r$ and *convergent* if it either has a top rule $p \rightsquigarrow r$ or a top rule $q \rightsquigarrow r$ (the arrow notation $\rightsquigarrow$ here reflects that in all these cases the rule can be either defeasible or strict). On this account, a convergent 'argument' with two reasons p and q is in fact a set of two separate arguments for the same conclusion q. After having given these definitions, [22] continue by combining them with the notion of 'serial' arguments, which are arguments that chain multiple inference steps. However, this complication is for present purposes irrelevant, so I will just focus on the case where there are several potential reasons for the same conclusion, where each of these reasons is either given as a fact or as a conclusion of an argument. How can cumulative arguments then be characterised?

Since the accrual sets of Definition 4.1 contain arguments for the same conclusion, we would at first sight seem to have lost the ability to distinguish between convergent and accruals/cumulative arguments. However, the key is the argument ordering. In the literature on argumentation theory the problem is often characterised in terms of whether an argument can convince an audience (see e.g. the initial discussions in [21]). In the present setting, this is the question whether an argument survives the competition with its counterarguments given all relevant attack and defeat relations between arguments. And the defeat relation depends on the argument ordering. We can say that two reasons p and q are *linked* as regards claim r if both p and q are needed to build an argument for r, so if there is a rule $p, q \rightsquigarrow r$ but no rules $p \rightsquigarrow r$ and $q \rightsquigarrow r$. By contrast, for both convergent and cumulative reasons there are rules $p \rightsquigarrow r$ and $q \rightsquigarrow r$ but no rule $p, q \rightsquigarrow r$. Then let A_1 be an argument for r with top rule $p \rightsquigarrow r$ and A_2 an argument with top rule $q \rightsquigarrow r$. We then say that A_1 and A_2 are *convergent* whenever $\{A_1\} \approx \{A_1, A_2\}$ and $\{A_2\} \approx \{A_1, A_2\}$, they are *cumulative accruals* (or *cumulative* for short) whenever $\{A_1\} \prec \{A_1, A_2\}$ and $\{A_2\} \prec \{A_1, A_2\}$, and they are *non-cumulative accruals otherwise*. In other words, A_1 and A_2 are convergent iff together the arguments are just as 'strong' in the argument ordering as either A_1 or A_2 alone, they are cumulative if together they are stronger than either A_1 or A_2 alone, and they are non-cumulative accruals if together they are stronger or weaker than either A_1 or A_2 alone. This account in terms of relative strength is arguably a more natural way to characterise the difference between convergent and accrual arguments than an account that reduces accruals to a special kind of linked arguments.

Gordon [8] defines the notions of linked, convergent and cumulative arguments as properties of individual arguments, and he does so in terms of different weight functions on arguments. An argument is *linked* iff its conclusion has maximum weight just in case all its premises are in and minimum weight otherwise, it is *convergent* iff its conclusion has maximum weight just in case at least one of its premises is in and minimum weight otherwise, and it is *cumulative* iff its weight increases the more of its premises are *In*. Thus Gordon defines, just as the present proposal, the difference between convergent and cumulative arguments in terms of the strength of an argument, although in his case this strength is absolute instead of relative. Unlike in the present proposal, Gordon does not address the accrual of multiple arguments for the same conclusion. An advantage of accruing multiple arguments is that thus arguments based on different inference rules can be naturally accrued. For example, we may have two arguments for the same conclusion based on the argument schemes from expert opinion and from statistical generalisation. It is not obvious how such arguments can be naturally combined in a single Carneades-style argument without turning the inference rules into premises.

7 CONCLUSION

In this paper a new formal model of argument accrual has been proposed as an adaptation of the *ASPIC*$^+$ framework for structured argumentation. The model advances the state-of-the art in that it overcomes several weaknesses of previously proposed models of argument accrual. It was shown to satisfy [15]'s three principles of accrual, to reduce to 'standard' *ASPIC*$^+$ if each statement has at most one argument, and to have desirable formal properties that are in line with standard work on formal argumentation. Its usefulness was illustrated by applications to a range of situations in legal reasoning. Among other things, this has clarified the relation with recent work on factor-based legal reasoning using abstract dialectical frameworks. Future work should turn the current semantic specification into computational procedures and should further investigate its applicability to realistic argumentation scenarios in legal and other kinds of reasoning.

REFERENCES

[1] L. Al-Abdulkarim, K.D. Atkinson, and T.J.M. Bench-Capon. A methodology for desinging to reason with legal cases using abstract dialectical frameworks. *Artificial Intelligence and Law*, 24:1–50, 2016.
[2] V. Aleven. Using background knowledge in case-based legal reasoning: a computational model and an intelligent learning environment. *Artificial Intelligence*, 150:183–237, 2003.
[3] T.J. Anderson and W. Twining. *Analysis of Evidence.* Cambridge University Press, Cambridge, second edition, 2005.
[4] P. Baroni, M. Caminada, and M. Giacomin. An introduction to argumentation semantics. *The Knowledge Engineering Review*, 26:365–410, 2011.
[5] G. Brewka and S. Woltran. Abstract dialectical frameworks. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference*, pages 102–111. AAAI Press, 2010.
[6] P.M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n–person games. *Artificial Intelligence*, 77:321–357, 1995.
[7] R.M. Dworkin. Is law a system of rules? In R.M. Dworkin, editor, *The Philosophy of Law*, pages 38–65. Oxford University Press, Oxford, 1977.
[8] T.F. Gordon. Defining argument weighing functions. *Journal of Applied Logics – IfCoLog Journal of Logics and their Application*, 5:747–773, 2018.
[9] T.F. Gordon, H. Prakken, and D.N. Walton. The Carneades model of argument and burden of proof. *Artificial Intelligence*, 171:875–896, 2007.
[10] T.F. Gordon and D.N. Walton. Formalizing balancing arguments. In P. Baroni, T.F. Gordon, T. Scheffler, and M. Stede, editors, *Computational Models of Argument. Proceedings of COMMA 2016*, pages 327–338. IOS Press, Amsterdam etc, 2016.
[11] S. Modgil and H. Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397, 2013.
[12] S. Modgil and H. Prakken. The ASPIC+ framework for structured argumentation: a tutorial. *Argument and Computation*, 5:31–62, 2014.
[13] S. Modgil and H. Prakken. Abstract rule-based argumentation. In P. Baroni, D. Gabbay, M. Giacomin, and L. van der Torre, editors, *Handbook of Formal Argumentation*, volume 1, pages 73–141. College Publications, London, 2018.
[14] J.L. Pollock. *Cognitive Carpentry. A Blueprint for How to Build a Person.* MIT Press, Cambridge, MA, 1995.
[15] H. Prakken. A study of accrual of arguments, with applications to evidential reasoning. In *Proceedings of the Tenth International Conference on Artificial Intelligence and Law*, pages 85–94, New York, 2005. ACM Press.
[16] H. Prakken. Relating ways to instantiate abstract argumentation frameworks. In K.D. Atkinson, H. Prakken, and A.Z. Wyner, editors, *From Knowledge Representation to Argumentation in AI, Law and Policy Making. A Festschrift in Honour of Trevor Bench-Capon on the Occasion of his 60th Birthday*, pages 167–189. College Publications, London, 2013.
[17] H. Prakken. Formalising debates about law-making proposals as practical reasoning. In M. Araszkiewicz and K. Płeszka, editors, *Logic in the Theory and Practice of Lawmaking*, pages 301–321. Springer, Berlin, 2015.
[18] J. Raz. *Practical Reason and Norms.* Princeton University Press, Princeton, 1975.
[19] B. Verheij. *Rules, reasons, arguments: formal studies of argumentation and defeat.* Doctoral dissertation University of Maastricht, 1996.
[20] B. Verheij, J.C. Hage, and H.J. van der Herik. An integrated view on rules and principles. *Artificial Intelligence and Law*, 6:3–26, 1998.
[21] D.N. Walton and T.F. Gordon. Cumulative arguments in artificial intelligence and informal logic. *Revista Iberoamericana de Argumentacion*, 14:1–28, 2017.
[22] Bin Wei and H. Prakken. Defining the structure of arguments with AI models of argumentation. In F. Bex, F. Grasso, N. Green, F. Paglieri, and C. Reed, editors, *Argument Technologies: Theory, Analysis, and Applications*, pages 1–22. College Publications, London, 2017.