Some reflections on two current trends in formal argumentation

Henry Prakken

[1] Department of Information and Computing Sciences, Utrecht University
[2] Faculty of Law, University of Groningen

Abstract. This paper discusses two recent developments in the formal study of argumentation-based inference: work on preference-based abstract argumentation and on classical (deductive) argumentation. It is first argued that general models of the use of preferences in argumentation cannot leave the structure of arguments and the nature of attack and defeat unspecified. Then it is claimed that classical argumentation cannot model some common forms of defeasible reasoning in a natural way. In both cases it will be argued that the recently proposed *ASPIC*$^+$ framework for structured argumentation does not suffer from these limitations. In the final part of the paper the work of Marek Sergot on argumentation-based inference will be discussed in light of the preceding discussion.

1 Introduction

Argumentation is a form of reasoning that makes explicit the reasons for the conclusions that are drawn and how conflicts between reasons are resolved. The study of argumentation has an inferential side, focused on how conclusions can be drawn from a body of uncertain, incomplete and/or inconsistent information, and a dialogical side, focused on how intelligent agents can resolve their conflicts of opinion by engaging in dialogue. This paper is about the inferential side of argumentation.

The study of argumentation in AI is nowadays very popular. This is good, since the idea of argumentation has much intellectual and application potential. Arguments are a natural concept in science, in politics, in business, in professions like law and medicine and in everyday conversation, so important applications are easy to imagine. However, to realise this potential, insight is needed in the strengths and limitations of argumentation formalisms. The aim of this paper is to critically examine two current research strands, namely, research on preference-based abstract argumentation and on classical (deductive) argumentation. (With 'deductive argumentation' I mean any form of argumentation in which arguments can only be attacked on their premises, that is, in which the rules for constructing arguments are assumed to be certain. With 'classical argumentation' I mean the special case of deductive argumentation in which arguments are valid standard propositional or first-order inferences.)

I shall first sketch a brief history of the study of argumentation in AI in light of the topic of this paper, only surveying work that is still relevant or influential today. I shall then briefly outline my latest work on argumentation-based inference, my version of the *ASPIC*$^+$ framework, published in [58]. In the main part of the paper I shall use this framework to critically examine the above-mentioned research strands, arguing that both are inherently limited and that *ASPIC*$^+$ does not share these limitations. As for preference-based abstract argumentation I shall argue that general models of the use of preferences in argumentation cannot leave the structure of arguments and the nature of attack and defeat unspecified. As for classical argumentation, I shall argue that it cannot model some common forms of defeasible reasoning in a natural way. In the final part of the paper I shall discuss Marek Sergot's work on argumentation-based inference in light of the preceding discussion.

2 A brief history of formal and computational research on argumentation

Historically, the formal study of argumentation-based inference mainly originated from research on nonmonotonic logic and logic programming, while it was also influenced by research in AI & Law. From the second half of the 1980s, argumentation was proposed as a new way to model nonmonotonic inference [44, 50, 43, 70, 76], culminating in [25]. around the same time, several AI & Law researchers proposed formal models of legal argument, making use of and extending general work on nonmonotonic logic; e.g. [31, 56, 33, 73, 62]. All early AI work on argumentation specified the structure of arguments. A key element in much of this work was a distinction between strict and defeasible inference rules, going back to a similar distinction made in e.g. default logic [66] and work on inheritance networks [35]. Even Dung in his landmark 1995 paper stood in this tradition. Dung did two things: he developed the new idea of abstract argumentation frameworks, and he used this idea to reconstruct and compare a number of then mainstream nonmonotonic logics and logic-programming formalisms, namely, default logic, the first version of Pollock's system for defeasible reasoning [50] and several logic-programming semantics. However, these days the second part of his paper is largely forgotten[3] and his paper is almost exclusively cited for its first part, on AFs.

Let me say more about the work that followed [25], starting with a very brief review of the main notions. An *abstract argumentation framework* (AF) is a pair $\langle AR, attacks \rangle$, where AR is a set arguments and $attacks \subseteq AR \times AR$ is a binary relation. The theory of AFs then addresses how sets of arguments (called *extensions*) can be identified which are internally coherent and defend themselves against attack. A key notion here is that of an argument being *acceptable with respect to*, or *defended by* a set of arguments: $A \in AR$ is defended by $S \subseteq AR$ if for all $A \in S$: if $B \in AR$ attacks A, then some $C \in S$ attacks B. Then relative to a given AF various types of extensions can be defined as follows (here E is *conflict-free* if no argument in E attacks an argument in E):

- E 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 maximal (with respect to set inclusion) admissible set;
- E is a *stable extension* if E is admissible and attacks all arguments outside it;

[3] The same holds for a third part on relations with cooperative game theory.

– E is a *grounded extension* if E is the least fixpoint of operator F, where $F(S)$ returns all arguments defended by S.

In the first years after publication of this landmark paper it gave rise to three kinds of follow-up work. Some continued to use AFs as Dung did in his paper, namely, to reconstruct and compare existing systems as instances of AFs. For example, Hadassah Jakobovits [40, 39] showed that a later version of Pollock's system for defeasible reasoning [51, 52] has preferred semantics and Claudette Cayrol [22] related various forms of classical argumentation to Dung's stable semantics and (with Leila Amgoud in [4]) to Dung's grounded semantics for AFs. Others further developed the theory of AFs. For example, an alternative, labelling-based version of semantics for AFs was developed [74, 39, 18] (Jakobovits also generalised this semantics).

In my own work I used AFs in yet another way. Immediately when I read Dung's 1995 paper I realised that it was a breakthrough, and I decided to continue my evolving work with Giovanni Sartor in the context of Dung's idea of abstract argumentation frameworks. We had been working on defining a new argumentation system with rule priorities in the language of extended logic programming. We now decided to design it in such a way that it results in a Dung AF, so that its semantics is given by the theory of AFs[4]. While defining the structure of arguments was relatively easy, the hard part was designing the *attacks* relation. In fact, in an attempt to be more in line with natural language, we renamed *attack* to *defeat* and we reserved the term *attack* for more basic, purely syntactical forms of conflicts between arguments, such as having contradictory conclusions. Then we combined our *attack* relation with rule preferences that resolved some attacks, resulting in a *defeat* relation on AR that is a subset of our *attack* relation. We then applied all of Dung's definitions to our pair $\langle AR, defeat\rangle$[5]. (Arguably the work on assumption-based argumentation, starting with [14] and going back to [42], is another example of designing new systems within Dung's abstract approach, although it was not until [27] that this was formally proven.)

All this work makes, in my opinion, proper use of Dung's ideas on abstract argumentation and shows that when used properly, Dung's idea of abstract argumentation frameworks is very valuable. However, in my opinion this is less obviously the case for another, more recent way to work with AFs, namely, extending them with new elements without specifying the structure of arguments. In my opinion, this approach has some inherent limitations. The approach was first applied for preferences, by e.g. [3] (later work has added, for example, values [8] and constraints [23] to AFs). [3] added to AFs a a preference relation on AR, resulting in *preference-based argumentation frameworks* ($PAFs$), which are a triple $\langle AR, attacks, \preceq\rangle$. An argument A then *defeats* an argument B if A attacks B and $A \not\prec B$. Thus each PAF generates an AF of the form $\langle AR, defeats\rangle$, to which Dung's theory of AFs can be applied. In a way, this idea is an abstraction of my work with Giovanni Sartor in [62] but there is a crucial

[4] We thus followed Dung's advice when I first met him, which was that more work needed to be done on the structure of the arguments. (All I could say on that occasion was that I extremely admired his paper.) Dung himself gave fuller treatment of argument structure in later work with Bob Kowalski, Francesca Toni and others, on assumption-based argumentation.

[5] Strictly speaking the formalism was not based on [25] but on [24]. However, a reconstruction in terms of [25] is straightforward.

difference, since in [62] (and also in *ASPIC*$^+$) the structure of arguments is crucial in determining how preferences must be applied to attacks. Since $PAFs$ do not specify the structure of arguments, they cannot model various subtle differences at this point, as I shall explain in detail below in Section 4. More generally, I shall argue that the approach to extend AFs at the abstract level with preferences, without specifying the structure of arguments or the nature of attack and defeat, is prone to run into problems.

A recent development that does specify the structure of arguments and the nature of attack and defeat is work on classical argumentation, e.g. [4, 10, 11, 32]. The idea here is that arguments are classical propositional or first-order proofs from inconsistent sets of premises and that arguments can only be attacked on their premises. More recently, this work was by [1, 2] generalised to any underlying Tarskian deductive logic. This research strand is definitely interesting but in light of the above-mentioned earlier work on argumentation with defeasible rules, it raises the question to what extent forms of argumentation can be reduced to inconsistency handling in deductive logic. This question is not new: it also arose in the field of nonmonotonic logic in the 1980s and 1990s; see e.g. [17, 30, 52]. While these discussions were not fully conclusive, there was at least an awareness among nonmonotonic logicians of those days that classical-logic approaches have some potential limitations, which need to be taken care of. In Section 5 I shall argue that this awareness is less apparent in current research on classical and deductive argumentation. Moreover, I shall take a firm position in this debate, arguing that classical and other deductive approaches cannot model some common forms of defeasible reasoning in a natural way, and that therefore any model of argumentation that claims to be general must leave room for defeasible inference rules.

3 The *ASPIC*$^+$ Framework

While I am still proud of my work with Giovanni Sartor in [62], over the years I came to see its limitations. For example, it has a simple logic-programming like language, it does not model premise attack, it models a specific use of preferences, and it has a specific semantics. Most importantly, its rules can only be used to express domain-specific knowledge, such as 'birds fly' or legal rules. By contrast, the work of especially John Pollock [50, 51, 52] and Gerard Vreeswijk [76, 77] provides general accounts of structured argumentation, since their strict and defeasible inference rules (Pollock called them 'conclusive' and 'prima facie reasons') are meant to capture general patterns of inference: their strict rules can, for instance, express the laws of propositional and first-order logic, and their defeasible rules can capture general epistemological principles (Pollock) or argumentation schemes [75, 13, 59]. For this reason I and others in the European ASPIC project integrated and further developed the work of Pollock, Vreeswijk and Prakken & Sartor. The first published version of the *ASPIC* system [20] still has a simple notion of a rule, only suitable for expressing domain-specific knowledge, and has no knowledge base or preferences. However, the version I presented in [58], now called the *ASPIC*$^+$ framework, claims to be a general framework for structured argumentation. It abstracts as much as possible from the nature of the logical language and the inference rules and from the ways in which preferences can be used to distinguish between attack and defeat, it generalises classical negation to an arbitrary contrariness relation and it adds premise attack to Pollock's notions of rebutting and undercutting attack. In my work on *ASPIC*$^+$ the framework is not used as a computational formalism but as a theoretical framework for expressing, analysing and relating specific systems. I am especially interested in identifying conditions under which instantiations of *ASPIC*$^+$ satisfy [20]'s rationality postulates for argumentation-based inference.

The *ASPIC*$^+$ framework assumes an unspecified logical language $\mathcal{L}$ with a binary contrariness relation and defines arguments as inference trees formed by applying strict or defeasible inference rules of the form $\varphi_1, \ldots, \varphi_n \rightarrow \varphi$ and $\varphi_1, \ldots, \varphi_n \Rightarrow \varphi$, where $\varphi_1, \ldots, \varphi_n$ are the *antecedents* and φ the *consequent* φ of the rule. The framework applies to any set of strict and defeasible inference rules formulated over $\mathcal{L}$. As said above, they can be used to express domain-specific knowledge but also to capture general patterns of reasoning.

Informally, that an inference rule is strict means that if its antecedents are accepted, then its consequent must be accepted *no matter what*, while that an inference rule is defeasible means that if its antecedents are accepted, then its consequent must be accepted *if there are no good reasons not to accept it*. In other words, if an inference rule is strict, then it is rationally impossible to accept its antecedents while refusing to accept its consequent, while if an inference rule is defeasible, it is rationally possible to accept its antecedents but not its consequent. From the distinction between strict and defeasible rules at least two design decisions and one rationality postulate follow: (1) arguments cannot be attacked on applications of their strict rules, (2) it does not make sense to make strict rules subject to a priority mechanism (since they must always be applied), and (3) extensions must be closed under application of strict rules but not under application of defeasible rules. In principled instantiations of *ASPIC*$^+$ the set of strict rules will be determined by the choice of the logical language $\mathcal{L}$: its formal semantics will tell which inference rules over $\mathcal{L}$ are valid and can therefore be added to $\mathcal{R}_s$. If the strict rules are thus determined by the semantics of $\mathcal{L}$, then they are normally all domain-independent; domain-specific inference rules are only needed if $\mathcal{L}$ does not have the means to express conditionals, as is the case in, for example, logic-programming languages, in which only literals can be expressed.

The basic notion of *ASPIC*$^+$ is that of an argumentation system.

Definition 1. *[Argumentation system] An* argumentation system *is a tuple* $AS = (\mathcal{L}, ^-, \mathcal{R}, \leq)$ *where*

- $\mathcal{L}$ *is a logical language.*
- $^-$ *is a contrariness function from* $\mathcal{L}$ *to* $2^{\mathcal{L}}$*, such that if* $\varphi \in \overline{\psi}$ *then if* $\psi \notin \overline{\varphi}$ *then* φ *is called a* contrary *of* ψ*, otherwise* φ *and* ψ *are called* contradictory*. The latter case is denoted by* $\varphi = -\psi$ *(i.e.,* $\varphi \in \overline{\psi}$ *and* $\psi \in \overline{\varphi}$*).*
- $\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_d$ *is a set of strict (*$\mathcal{R}_s$*) and defeasible (*$\mathcal{R}_d$*) inference rules such that* $\mathcal{R}_s \cap \mathcal{R}_d = \emptyset$.
- $\leq$ *is a partial preorder on* $\mathcal{R}_d$.

Arguments are constructed from a knowledge base, which is assumed to contain three kinds of formulas.

Definition 2. *[Knowledge bases] A* knowledge base *in an argumentation system* $(\mathcal{L}, ^-, \mathcal{R}, \leq)$ *is a pair* $(\mathcal{K}, \leq')$ *where* $\mathcal{K} \subseteq \mathcal{L}$ *and* $\leq'$ *is a partial preorder on* $\mathcal{K}_p$.

Here, $\mathcal{K} = \mathcal{K}_n \cup \mathcal{K}_p \cup \mathcal{K}_a$*, the* necessary, ordinary *and* assumption *premises, where these subsets of* $\mathcal{K}$ *are disjoint .*

Intuitively, arguments can only be attacked on their ordinary and assumption premises. Attacks on assumption premises always result in defeat while attacks on ordinary premises are resolved with preferences. Hence no preferences can be defined on $\mathcal{K}_n$ or $\mathcal{K}_a$.

Arguments can be constructed step-by-step by chaining inference rules into trees. Arguments thus contain subarguments, which are the structures that support intermediate conclusions (plus the argument itself and its premises as limiting cases). In what follows, for a given argument the function `Prem` returns all its premises, `Conc` returns its conclusion and `Sub` returns all its sub-arguments.

Definition 3. *[Argument] An* argument A *on the basis of a knowledge base* $(\mathcal{K}, \leq')$ *in an argumentation system* $(\mathcal{L}, ^-, \mathcal{R}, \leq)$ *is:*

1. φ *if* $\varphi \in \mathcal{K}$ *with:* $\texttt{Prem}(A) = \{\varphi\}$*;* $\texttt{Conc}(A) = \varphi$*;* $\texttt{Sub}(A) = \{\varphi\}$*;*
2. $A_1, \ldots A_n \rightarrow/\Rightarrow \psi$ *if* $A_1, \ldots, A_n$ *are arguments such that there exists a strict/defeasible rule* $\texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \rightarrow/\Rightarrow \psi$ *in* $\mathcal{R}_s/\mathcal{R}_d$*.*
   $\texttt{Prem}(A) = \texttt{Prem}(A_1) \cup \ldots \cup \texttt{Prem}(A_n)$, $\texttt{Conc}(A) = \psi$, $\texttt{Sub}(A) = \texttt{Sub}(A_1) \cup \ldots \cup \texttt{Sub}(A_n) \cup \{A\}$.

An argument is *strict* if all its inference rules are strict and *defeasible* otherwise, and it is *firm* if all its premises are in $\mathcal{K}_n$ and *plausible* otherwise.

Example 1. Consider a knowledge base in an argumentation system with

$$\mathcal{R}_s = \{p, q \rightarrow s;\ u, v \rightarrow w\}; \mathcal{R}_d = \{p \Rightarrow t;\ s, r, t \Rightarrow v\}$$
$$\mathcal{K}_n = \{q\}; \mathcal{K}_p = \{p, u\}; \mathcal{K}_a = \{r\}$$

An argument for w is displayed in Figure 1. The type of a premise is indicated with a superscript and defeasible inferences are displayed with dotted lines. Formally the

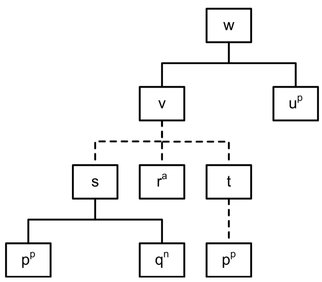

Fig. 1. An argument

argument and its subarguments are written as follows:

A_1: p A_5: $A_1 \Rightarrow t$
A_2: q A_6: $A_1, A_2 \rightarrow s$
A_3: r A_7: $A_5, A_3, A_6 \Rightarrow v$
A_4: u A_8: $A_7, A_4 \rightarrow w$

We have that

$\texttt{Prem}(A_8) = \{p, q, r, u\}$
$\texttt{Conc}(A_8) = w$
$\texttt{Sub}(A_8) = \{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8\}$

Combining an argumentation system and a knowledge base with an *argument ordering* results in an *argumentation theory*. The argument ordering is a partial preorder $\preceq$ on arguments (with its strict counterpart $\prec$ defined in the usual way). It could be defined in any way, for example, in terms of the orderings $\leq$ on $\mathcal{R}_d$ and $\leq'$ on $\mathcal{K}_p$. See Section 6 of [58] for two ways of doing so, expressing a weakest- and last-link principle.

Definition 4. *[Argumentation theories] An* argumentation theory *is a triple* $AT = (AS, KB, \preceq)$ *where* AS *is an argumentation system,* KB *is a knowledge base in* AS *and* $\preceq$ *is a partial preorder on the set of all arguments on the basis of* KB *in* AS *(below denoted by* $\mathcal{A}_{AT}$*).*

Arguments can be attacked in three ways: attacking a conclusion of a defeasible inference, attacking the defeasible inference itself, or attacking a premise.

Definition 5. *[***ASPIC**$^+$ ***attacks****]* A attacks B *iff* A undercuts, rebuts *or* undermines B, *where:*

• A undercuts *argument* B *(on* B'*) iff* $\texttt{Conc}(A) \in \overline{r}$ *for some* $B' \in \texttt{Sub}(B)$ *such that* B'*'s top rule is defeasible and is named by* r *in* $\mathcal{L}$.[6]

• A rebuts *argument* B *(on* B'*) iff* $\texttt{Conc}(A) \in \overline{\varphi}$ *for some* $B' \in \texttt{Sub}(B)$ *of the form* $B''_1, \ldots, B''_n \Rightarrow \varphi$*. In such a case* A contrary-rebuts B *iff* $\texttt{Conc}(A)$ *is a contrary of* φ*.*

• *Argument* A undermines B *(on* B'*) iff* $\texttt{Conc}(A) \in \overline{\varphi}$ *for some* $B' = \varphi$*,* $\varphi \in \texttt{Prem}_{a/p}(B)$. *In such a case* A contrary-undermines B *iff* $\texttt{Conc}(A)$ *is a contrary of* φ *or if* $\varphi \in \mathcal{K}_a$*.*

In Example 1 argument A_8 can be undercut in two ways: by an argument with conclusion φ such that $\varphi \in \overline{r_5}$ (where r_5 names rule $p \Rightarrow t$), which undercuts A_8 on A_5, and by an argument with conclusion φ such that $\varphi \in \overline{r_7}$ (where r_7 names rule $s, r, t \Rightarrow v$), which undercuts A_8 on A_7. Moreover, argument A_8 can be rebutted on A_5 with any argument for a conclusion φ such that $\varphi \in \overline{t}$ and on A_7 with any argument for a conclusion φ such that $\varphi \in \overline{v}$. Moreover, if $\overline{t} = -t$ and the rebuttal has a defeasible top rule, then A_5 in turn rebuts the argument for $\overline{t}$. However, A_8 itself does not rebut that argument, except in the special case where $w \in \overline{\overline{t}}$. Finally, argument A_8 can be undermined with an argument that has conclusion φ such that $\varphi \in \overline{p}$, $\overline{r}$ or $\overline{u}$.

Attacks combined with the preferences defined by an argument ordering yield three kinds of defeat. For undercutting attack no preferences are needed to make it succeed,

[6] This definition assumes that defeasible inference rules are named in $\mathcal{L}$; the precise nature of this naming convention will be left implicit.

since undercutters state exceptions to the rule they attack, and for contrary-rebutting and -undermining no preferences are needed since such contrary attacks already embody some kind of preference (cf. e.g. attacks on negation-as-failure assumptions in logic programming).

Definition 6. *[Successful rebuttal, undermining and defeat]*

– A successfully rebuts B *if* A *rebuts* B *on* B' *and either* A *contrary-rebuts* B' *or* $A \nprec B'$.
– A successfully undermines B *if* A *undermines* B *on* φ *and either* A *contrary-undermines* B *or* $A \nprec \varphi$.
– A defeats B *iff* A *undercuts or successfully rebuts or successfully undermines* B.

The success of rebutting and undermining attacks thus involves comparing the conflicting arguments at the points where they conflict. The definition of successful undermining exploits the fact that an argument premise is also a subargument.

The semantics of *ASPIC*$^+$'s argumentation theories is then given by linking them to Dung's abstract argumentation frameworks as follows:

Definition 7 (Argumentation framework). *An* abstract argumentation framework (AF) corresponding to an argumentation theory $\langle AS, KB, \preceq\rangle$ *is a pair* $\langle AR, attacks\rangle$ *such that:*

– AR *is the set* $\mathcal{A}_{AT}$ *as defined by Definition 3,*
– $attacks$ *is the defeat relation on* AR *given by Definition 6.*

A variant of this definition is where AR only contains the consistent arguments from $\mathcal{A}_{AT}$. Now one way to define a consequence notion for statements is to say that a statement is *justified* if it is the conclusion of a justified argument. An alternative is to say that it is justified if each extension contains an argument with the statement as conclusion (but the argument does not have to be the same in all extensions).

In [58] the extensions induced by Definition 7 are all shown to satisfy [20]'s rationality postulates of consistency, closure under strict inference and closure under subarguments, under complete, stable, preferred and grounded semantics. In [47] these results are also proven for the case where all arguments have consistent premises, so that *ASPIC*$^+$ can additionally capture classical-logic approaches to argumentation.

Several results testify to the generality of the *ASPIC*$^+$ framework. In [47] forms of classical argumentation were shown to be a special case of *ASPIC*$^+$ with the language of propositional logic, with only ordinary premises, with as strict rules all propositionally valid inferences and with no defeasible rules. Then the weakest-link preference mechanism of [58] was used to yield a preference-based version of classical argumentation that satisfies [20]'s rationality postulates. Furthermore, in [58] assumption-based argumentation was shown to be a special case of *ASPIC*$^+$ with only strict inference rules, only assumption-type premises and no preferences (the proof exploited the link between assumption-based argumentation and Dung's abstract frameworks as proven in [27]). Because of this result, the sufficient conditions identified in [20] and [58] for satisfying [20]'s rationality postulate of consistency also apply to assumption-based argumentation, which in general does not satisfy this postulate.

4 A critique of abstract preference-based argumentation frameworks

In 1998, Amgoud & Cayrol [3] introduced the notion of preference-based abstract argumentation frameworks ($PAFs$). Recall that these add to Dung's abstract frameworks a binary preference relation on arguments, which is a preorder $\preceq$. As mentioned above, they in fact follow the same approach at the abstract level as [62] at a more concrete level: they decompose Dung's [25] *attacks* relation into a more basic relation capturing purely syntactic forms of conflict, which they call "defeat". They then say that A "attacks" B if A "defeats" B and $A \not\prec B$. Note that [3] unlike [62] and the *ASPIC*$^+$ framework do not rename 'attack' to 'defeat'. To enable the comparison with *ASPIC*$^+$, I will reverse their uses of 'attack' and 'defeat'.

For reasons of clarity, I now reformulate Dung's definition of acceptability in terms of the *attacks* and $\preceq$ relation.

Definition 8. *Given a $PAF = \langle \mathcal{A}, attacks, \preceq \rangle$ an argument A is* acceptable with respect to *a set of arguments S if for all B attacking A such that $B \not\prec A$, there exists a C in S such that C attacks B and $C \not\prec B$.*

It is easy to see that if A *defeats* B is defined as A attacks B and $A \not\prec B$, then this is equivalent to

> An argument A is acceptable with respect to a set of arguments S if all B defeating A are defeated by a $C \in S$.

This is the formulation I will use below. It clearly reveals that Dung's semantics directly apply to $PAFs$ if Dung's *attacks* relation is replaced by the just-defined defeat relation. It also reveals that the *attacks* relation of [3], although it has the same name as Dung's *attacks* relation, is in fact a different relation: the role of Dung's *attacks* relation is now played by the *defeats* relation induced by a PAF.

I shall now show that in general $PAFs$ (and similar abstract frameworks like [8]'s value-based argumentation frameworks, or $VAFs$) model preference-based argumentation at a too high level of abstraction: I shall argue that in general a proper modelling of preferences in argumentation requires that the structure of arguments and the nature of attack are made explicit. To start with, there are reasonable notions of attack that result in defeat irrespective of preferences, such as *ASPIC*$^+$'s undercutting attack. A framework that does not make the structure of arguments explicit cannot distinguish between preference-dependent and preference-independent attacks. At first sight it might seem that this problem can be solved by allowing two abstract kinds of attack, called preference-dependent and preference-independent attack, and to apply the argument ordering only to the first type of attack. However, this solution still faces problems, since it cannot recognise that in general the question which preference must be used to resolve an attack depends on the structure of arguments.

Consider the following example in *ASPIC*$^+$, with $\mathcal{K}_n = \mathcal{K}_a = \emptyset; \mathcal{K}_p = \{p, q\}$, $\mathcal{R}_s = \emptyset$, $\mathcal{R}_d = \{p \Rightarrow r; q \Rightarrow \neg r; \neg r \Rightarrow s\}$, where the contrariness relation over $\mathcal{L}$ corresponds to classical negation in the obvious way. We then have the following arguments:

$A_1 = p \qquad\qquad B_1 = q$
$A_2 = A_1 \Rightarrow r \quad B_2 = B_1 \Rightarrow \neg r$
$\qquad\qquad\qquad\quad B_3 = B_2 \Rightarrow s$

We have that A_2 and B_2 attack each other and A_2 attacks B_3, since it directly rebuts its subargument B_2 (see Figure 2).

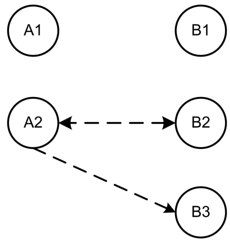

Fig. 2. The attack graph

Assume that the defeasible rules are ordered as follows: $q \Rightarrow \neg r < p \Rightarrow r$, $p \Rightarrow r < \neg r \Rightarrow s$ and let us apply the last-link argument ordering, which orders arguments according to the preferences of their last-applied defeasible rules (this ordering is, for instance, suitable for reasoning with legal rules). Then the following argument ordering is generated: $B_2 \prec A_2$ since $q \Rightarrow \neg r < p \Rightarrow r$, and $A_2 \prec B_3$ since $p \Rightarrow r < \neg r \Rightarrow s$. A PAF modelling then generates the following single defeat relation: A_2 defeats B_2 (see Figure 3). Then we have a single extension (in whatever semantics), namely, $\{A_1, B_1, A_2, B_3\}$. So not only A_2 but also B_3 is justified. However, this violates [20]'s

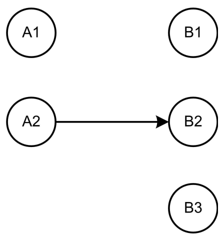

Fig. 3. The PAF defeat graph

rationality postulate of subargument closure of extensions, since B_3 is in the extension while its subargument B_2 is not. The cause of the problem is that the PAF modelling of this example cannot recognise that the reason why A_2 attacks B_3 is that A_2 directly attacks B_2, which is a subargument of B_3. So the PAF modelling fails to capture that in order to check whether A_2's attack on B_3 succeeds, we should compare A_2 not with

B_3 but with B_2, as happens in *ASPIC*$^+$. Now since $B_2 \prec A_2$ we also have that A_2 defeats B_3 (see Figure 4), so in *ASPIC*$^+$ the single extension (in whatever semantics) is $\{A_1, B_1, A_2, B_3\}$ and we have that A_2 is justified and both B_2 and B_3 are overruled, so closure under subarguments is respected. Moreover, recall that in [58] *ASPIC*$^+$ is shown to always satisfy this postulate. These problems are not due to the inclusion of

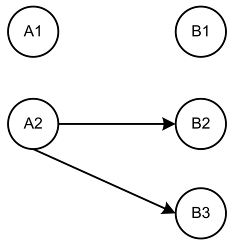

Fig. 4. The *ASPIC*$^+$ defeat graph

defeasible inference rules or the last-link ordering. Consider the following example in classical argumentation (imagine a version of *ASPIC*$^+$ with no defeasible rules, with $\mathcal{L}$ the language of propositional logic, with $\mathcal{R}_s$ consisting of all propositionally valid inferences and with consistent arguments):[7] $\mathcal{K}_n = \mathcal{K}_a = \emptyset; \mathcal{K}_p = \{p, q, \neg p\}$, where $q > \neg p > p$. The following arguments can be constructed:

$A_1 = p$ $\qquad$ $B = \neg p$
$A_2 = q$
$A_3 = A_1, A_2 \rightarrow p \wedge q$

We have that A_1 and B attack each other and B attacks A_3 (on p). Suppose arguments are compared in terms of their premises, and premise sets are compared according to [5]'s *democratic criterion*:

$$S \geq_s S' \text{ iff for all } r' \in S' \setminus S \text{ there exists an } r \in S \setminus S' \text{ such that } r >' r'.$$

Then we have that $A_1 \prec B$ and $B \prec A_3$. The PAF for this example then generates an extension containing A_2, A_3 and B under any semantics, which again violates subargument closure. Moreover, the closure of this extension under strict rules is inconsistent, so this PAF also violates the consistency postulate. In *ASPIC*$^+$we instead obtain that B defeats A_3 on A_1, so the correct outcome is obtained, namely, an extension with only A_2 and B and not including A_3.

The lesson that can be learned from these examples is that in general the choice of preference to resolve an attack depends on the structural nature of the attack, and the problem with $PAFs$ is that they cannot model the structural nature of attacks. Note

[7] As said above, such an instantiation is formally defined in [47] and then shown to satisfy [20]'s rationality postulates.

that the same observations hold for value-based argumentation frameworks and any similar abstract framework for preference-based argumentation. It may be that there are instantiations of $PAFs$ in which these problems do not arise, but even if they exist, we can still conclude that $PAFs$ are abstract but not general (a phrase coined by [21]). Their abstract nature gives the mistaken appearance of generality, while they make design choices that are only correct for certain classes of instantiations. More generally, the lesson to be learned is that for any extension of abstract argumentation that does not make the structure of arguments or the nature of attack explicit, such as extensions with constraints [23] or with weighted attacks [28], a careful analysis is needed whether these phenomena can indeed be modelled at the abstract level. One way to give such an analysis is by combining the extensions with the *ASPIC*$^+$ framework and investigating the resulting properties, as I have just done for $PAFs$.

In fact, there is one abstract framework that escapes the above criticism, namely, [45]'s *extended argumentation frameworks* ($EAFs$). Such frameworks extend Dung's AFs with attacks on attacks. Very briefly, since the theory of $EAFs$ does not put any constraints on attacks on attacks, the first above example can be modelled in $EAFs$ by letting the preference $q \Rightarrow \neg r < p \Rightarrow r$ attack B_2's attack on A_2. Of course, the theory of $EAFs$ does not give any guidance on how to model attacks on attacks, but such guidance can come from combining $EAFs$ with a framework for structured argumentation, just as is done in *ASPIC*$^+$ with Dung's AFs. For an initial proposal in this vein see [46].

5 A critique of work on classical and deductive argumentation

As already said above, much current formal and computational work on argumentation is on abstract argumentation, as introduced by [25]. However, to be useful and realistic, abstract models must be combined with accounts of the structure of arguments and the nature of attack and defeat. While this should be obvious, it is less obvious what such accounts should be. In the present section I shall argue that work on classical and, more generally, deductive argumentation is of limited applicability and that many, if not most forms of argumentation can only be modelled in a natural way by combining strict and defeasible inference rules.

5.1 Defeasible vs. plausible reasoning

Let us ask the question whether classical logic (or some other deductive logic) suffices for defining the inference rues with which arguments can be constructed. John Pollock, one of the fathers of our field, gave a negative answer. According to him any full theory of argumentation should give an account of the interplay between deductive and defeasible reasons:

> It is logically impossible to reason successfully about the world around us using only deductive reasoning. All interesting reasoning outside mathematics involves defeasible steps. [52, p.41]

> ... we cannot get around in the world just reasoning deductively from our prior beliefs together with new perceptual input. This is obvious when we look at the varieties of reasoning we actually employ. We tend to trust perception, assuming that things are the way they appear to us, even though we know that sometimes they are not. And we tend to assume that facts we have learned perceptually will remain true, as least for a while, when we are no longer perceiving them, but of course, they might not. And, importantly, we combine our individual observations inductively to form beliefs about both statistical and exceptionless generalizations. None of this reasoning is deductively valid. [53, p. 173]

In the 1980's and early 1990's Pollock's view was quite in agreement with most research on nonmonotonic logic at that time. Default logic [66], still one of the most influential nonmonotonic logics, added defeasible inference rules to the proof theory of classical logic. Systems for inheritance with exceptions [36] combined strict and defeasible inheritance rules. In 1992 Guillermo Simari and Ron Loui fully formalised [44]'s initial ideas on argumentation with strict and defeasible inference rules [70]. This work in turn led to the development of Defeasible Logic Programming [29]. [43] proposed the idea of abstract argumentation structures with strict and defeasible rules and showed how a number of existing nonmonotonic logics could be reconstructed as such structures. Gerard Vreeswijk further developed these ideas in his abstract argumentation systems [76, 77]. In 1994 Donald Nute published the first version of Defeasible Logic, which also combines strict and defeasible inference rules [48]. Finally, [62] formalized an argumentation logic with strict and defeasible inference rules and defeasible priorities explicitly as an instance of [25]'s abstract argumentation frameworks.

However, a more recent research strand is to model argumentation as inconsistency handling in classical (or some other deductive) logic [10, 4, 49, 11, 1, 32]. In terms of $ASPIC^+$ this work regards all inference rules as strict. Accordingly, arguments can in these approaches only be attacked on their premises, while in systems with defeasible inference rules, they can also be attacked if all their premises are accepted, since the premises only presumptively support their conclusion. Here the philosophical distinction between *plausible* and *defeasible* reasoning is relevant; cf. [67, 68] and [76, Ch. 8]. Following Rescher, Vreeswijk describes plausible reasoning as sound (i.e., deductive) reasoning on an uncertain basis and defeasible reasoning as unsound (but still rational) reasoning on a solid basis. In these terms, models of deductive argumentation formalise plausible reasoning while $ASPIC^+$ combines plausible and defeasible reasoning.

5.2 Can defeasible reasoning be reduced to plausible reasoning?

The current attempts to model argumentation on the basis of classical/deductive logic have their parallel in the history of nonmonotonic logic, in which there have been several attempts to reduce nonmonotonic reasoning to some kind of inconsistency handling in classical logic, e.g. [38, 54, 16, 7][8]. If such a reduction is possible then there is no

[8] Assumption-based argumentation [14, 26] is similar but more general; on the one hand it only allows for premise attack and thus in fact only allows for strict rules, on the other hand it does not commit to classical logic as the source of its rules.

need for new logics but just for a proper way of modelling inconsistency handling in deductive logic, which, so it is said, is well-understood [11, p. 16].

However, these approaches have been criticised for producing counterintuitive results due to the use of the material implication, which is claimed to be logically too strong for representing defeasible conditionals; cf. e.g. [17, 55, 30, 19]. Let us examine this debate for so-called 'default reasoning', which is the kind of defeasible reasoning where empirical generalisations ('defaults' for short) are applied to particular facts to infer new particular facts. Recall that defeasible reasoning is unsound reasoning from a certain basis. For example, given that quakers are normally pacifists, that republicans are normally not pacifists and that Nixon was both a quaker and a republican, a defeasible reasoner is interested in what can be concluded about whether Nixon was a pacifist. Note that there is nothing inconsistent in these givens. The reason that they are jointly consistent is that 'If Q then normally P'and Q' does not deductively imply P since things could be abnormal: Nixon could be an abnormal quaker or republican. A defeasible reasoner does not want to reject any of the above statements. Instead such a reasoner, given knowledge about how the world normally is, wants to assume whenever possible that things are normal, in order to jump to conclusions about Nixon in the absence of evidence to the contrary.

Now there are two ways to formalise such normality assumptions. The first is to add defeasible inference rules to those of classical logic, which formalize the defeasible jumps to conclusions. Thus the normality assumption is captured by the defeasible nature of the new inference rules. This is what systems with defeasible rules do. The second way is to make the implicit normality assumptions explicit as additional premises. More precisely, they are added to the antecedents of material implications expressing the default, and their assumed truth can be expressed as an additional premise. This is what inconsistency handling approaches in classical logic do. Let us call this the normality assumption approach.

Let us formally illustrate this with another well-known example from the literature on nonmonotonic logic.

(1) Birds normally fly
(2) Penguins normally don't fly
(3) All penguins are birds
(4) Penguins are abnormal birds with respect to flying
(5) Tweety is a penguin

From these natural-language statements any defeasible reasoner will conclude that Tweety can fly.

Let us formalise the normality assumption approach in the classical-logic instantiation of *ASPIC*$^+$ described at the end of the previous section and defined in [47].

(1) $bird \wedge \neg ab_1 \supset canfly$
(2) $penguin \wedge \neg ab_2 \supset \neg canfly$
(3) $penguin \supset bird$
(4) $penguin \supset ab_1$
(5) $penguin$

The idea is that the normality assumptions of a defeasible reasoner are expressed as additional statements $\neg ab_1$ and $\neg ab_2$ in the knowledge base K. Assume first that all statements are in the ordinary premises $\mathcal{K}_p$. This agrees with [4, 11], in which all premises can be attacked. I shall first show that this idea does not work. Recall that a defeasible reasoner regards (1-5) as given and is interested in what follows from it about Tweety's flying abilities. A defeasible reasoner does not want to give up any of (1-5). However, note that $\{1, 2, 3, 4, 5\} \cup \{\neg ab_1, \neg ab_2\}$ is minimally inconsistent (with respect to set inclusion) so if we take any single element out, the rest can be used to build an argument against it. This means that we can formally build arguments against any of (1-5), which a defeasible reasoner is not prepared to do.

Let us therefore add (1-5) to the axioms $\mathcal{K}_n$, so that they cannot be attacked. In principle this could be an acceptable way of reducing defeasible to plausible reasoning, since it precludes the construction of arguments against what is explicitly given.[9] Then we have the argument $\{1, 2, 3, 4, 5\} \cup \{\neg ab_2\} \rightarrow \neg \mathit{canfly}$, which has no counterargument. Note in particular that $\{4, 5\} \vdash ab_1$, so any argument with premise $\neg ab_1$ will be strictly defeated by a strict-and-firm underminer (in *ASPIC*$^+$ strict-and-firm arguments are strictly preferred to all other arguments). So at first sight it would seem that this refined approach adequately models default reasoning.

However, this approach still has problems, as can be illustrated by changing our example a little: above it was given as a matter of fact that Tweety is a penguin but in reality the particular 'facts' of a problem are not simply given but derived from information sources (sensors, testimonies, databases, the internet, and so on). Now, as described by Pollock in the above quotations, in reality none of these sources is fully reliable so inferring facts from them can only be done under the assumption that things are normal. So let us change the example by saying that Tweety was observed to be a penguin and that animals that are observed to be penguins *normally* are penguins. We change 5 to $5'$ and we add 6 to $\mathcal{K}_n$:

(5') $observed\_as\_penguin$
(6) $observed\_as\_penguin \wedge \neg ab_3 \supset penguin$

Moreover, we add $\neg ab_3$ to $\mathcal{K}_p$. We can still build an argument for the conclusion that Tweety cannot fly, namely, $\{1, 2, 3, 4, 5'\} \cup \{\neg ab_2, \neg ab_3\} \rightarrow \neg \mathit{canfly}$. However, now we can build an attacker of this argument, namely $\{1, 2, 3, 4, 5', 6\} \cup \{\neg ab_1, \neg ab_2\} \rightarrow ab_3$. At first sight, it would seem that we can still obtain the intuitive outcome by introducing a priority mechanism and saying that the first argument is preferred over the second since the assumption $\neg ab_3$ is preferred over the assumption $\neg ab_1$. However, the problem is that this is an ad-hoc solution; there is no general principle on which such a preference can be based. The heart of the problem is the fact that the material implication satisfies contraposition, a property which is too strong for default statements. The problem is quite fundamental since, as stressed by Pollock, ultimately all our knowledge about the world is derived with the help of perceptions; and derivations

[9] Alternatively, a priority mechanism could be used to let the unwanted arguments be defeated, but this does not prevent the problems described next, while a defeasible reasoner does not even want to consider such arguments.

from perceptions are inherently defeasible. We must therefore conclude that any full model of argumentation must address the issue of interleaving reasoning with strict and defeasible inference rules.[10].

In *ASPIC*$^+$ the example can be correctly formalised in several ways. The simplest is to model the above defaults as domain-specific inference rules, by replacing the $\supset$ symbol in (1,2,5') with $\Rightarrow$ and replacing the conjunctions in these statements with commas. Formalisations with general defeasible inference rules are also possible, by introducing a connective for default conditionals in $\mathcal{L}$ and adding modus ponens but not modus tollens for this connective to $\mathcal{R}_d$.

How does assumption-based argumentation deal with this example? It is easy to find modellings that yield the intuitive outcome, for example, by replacing all material implications with inference rules (where (5') becomes an inference rule with empty antecedent). However, as remarked above, assumption-based argumentation does in general not satisfy the consistency postulate and special cases that do satisfy this postulate may still yield the unwanted outcome. For example, if the inference rules are closed under so-called transposition (which in [20] and [58] is shown to suffice for consistency), then in my just-suggested modelling we also have the following transposed versions of the rule versions of (4) and (6):

(4') $\neg ab_1 \rightarrow \neg penguin$
(6') $observed\_as\_penguin \wedge \neg penguin \rightarrow ab_3$

Then an argument for ab_3 can be constructed by applying (4') and then (6) to $\neg ab_1$, which in all semantics prevents $penguin$ from being justified.

My analysis in this section is not meant to be original. In fact, in the literature on nonmonotonic logic several discussions of this kind can be found, e.g. in [17, 30] or my own [55]; see also [19]. Moreover, the reader may have noted the formal similarity of this example with the Yale Shooting scenario as discussed by [34]. My point is rather that insights that were once well known are in danger of being forgotten today. See, for example, [37]'s proposal for formalising 'experts are normally truthful, except when they have vested interests in what they are saying' in classical-logic argumentation:

(1) $expert \supset truthful$
(2) $expert \wedge vested\_interests \supset \neg truthful$

[10] It might be thought that a general principle for preferring $\neg ab_3$ over $\neg ab_1$ is that our perceptions must be given greater priority to counter the fundamentalist sceptic, who maintains that since perceptions are fallible, it is impossible to obtain any knowledge since all knowledge is ultimately based on perceptions. Granted that the idea of defeasible reasoning is a convincing reply to the sceptic, in my opinion it does not imply that inferences from perceptions are always more certain than other forms of defeasible inference. In fact, the strength of perceptive inferences is highly context-dependent, as, for example, much empirical research on eyewitness reliability shows. Moreover, this viewpoint does not explain why the fact that penguins are exceptional birds with respect to flying is a reason to believe that the perception that Tweety is a penguin is flawed. On the contrary, the opposite point of view (that it is not such a reason) strengthens the attack on the sceptic.

Then being an expert implies not having vested interests, and having vested interests implies not being an expert, so it cannot even be consistently stated that somebody is an expert with vested interests. This is clearly undesirable if the problem we are modelling is to verify what can be concluded about the truthfulness of a given expert of whom we know that he has vested interests in what he is saying.

I end this section with a brief discussion of my own research experiences in modelling actual argumentation. I have carried out three substantial case studies [61, 57, 60] and supervised two further case studies [71], [12, Ch. 6], all in the legal domain. In all these case studies we found that the reasoning interleaves deductive and defeasible inference, with particular emphasis on defeasible inference. The facts of a case come from sources of evidence, and as stressed by Pollock (see the above quotations) inferences from such sources are always defeasible. Then classification rules are applied to the facts and such rules are often defeasible (see e.g. [9]'s discussion of open texture in the law). Finally, legal rules are applied to the classified facts and legal rules are also inherently defeasible [33, 69]. Another finding was that in none of the case studies premise attack played a significant role. If a premise was challenged or attacked at all, then almost always support for the premise was given in the form of a defeasible argument, so that the attack subsequently took the form of a rebuttal or undercutter.

6 Marek Sergot's work on argumentation-based inference

Although Marek Sergot's main research interests lie outside argumentation, he still published several papers on argumentation, such as [41], [6] and [72]. In light of the topic of this paper, his [41] is particularly relevant, since it applies argumentation to a genuine problem of bioinformatics and thus provides a clear illustration of the practical benefits of the argumentation paradigm. In fact, as early as in 1998 Sergot had already, with Trevor Bench-Capon in [9], proposed the idea of rule-based argumentation systems. Although their paper was about legal reasoning, the ideas were of much wider relevance. It is fair to say that much of the work in the 1990's on argumentation-based inference was foreseen in this paper.

Nevertheless, in their [9], Sergot & Bench-Capon at first sight seemed to commit to a deductive approach to argumentation, which is the approach that I have critically discussed in this paper. After sketching the general idea of rule-based argumentation, they say the following about how arguments can be challenged.

> When a system of conflicting rules is used to generate contradictory conclusions, then the proofs which are constructed in the process do take on the nature of arguments. An argument, like a proof, starts from some assumptions or premises and moves by rules of inference to a conclusion. In the case of a proof we know that the rules of inference are truth preserving: if we accept the premises it is not open to us to deny the conclusion. But we can properly refuse to accept a proof, by denying the premises on which it is based. It is for this reason that arguments can be identified with proofs; that arguments are persuasive rather than compelling; that arguments may be sound (in that they apply valid rules of inference to the premises they are given) but weak (in

> that the premises may be questionable); and that two equally sound arguments may give contradictory conclusions. It is always open to someone to reject the conclusion of an argument. [9, pp. 19-20]

Clearly this quote describes what I in this paper have called deductive argumentation. However, a few years later Marek Sergot endorsed my PhD thesis, in which I criticised the possibility of reducing defeasible reasoning to inconsistency handling in classical logic, and from what I remember, he fully agreed with me. So we should not read the above quote too strictly, as excluding other forms of attack.

In [41] Sergot and his colleagues apply Dung's abstract frameworks to a problem of bioinformatics, namely, predicting the structure of of protein based on its sequence. They describe how an abstract AF captures the expert knowledge used by a researcher to interpret the output of a biological search engine. The search engine matches an unannotated protein sequence with a database of protein structures. The question to be answered by the researcher is whether the match is positive or negative, that is, whether the match is a good indicator of protein structure or not. Arguments are called claims and connect single features of a sequence to either a 'yes' or a 'no' answer to this question. For example, a long match is a good indicator of protein structure while a short match is not a good indicator of protein structure. Attack relations between arguments are not computed from their logical form but handcoded by the expert. For example, the argument 'The identity between the match sequence and the query sequence is low, so the match is negative' might attack the above argument based on a long match. Attacks are not necessarily symmetric.

In fact, the model thus created by the expert is a Dung AF plus an additional element, namely, for each argument the information whether it supports or opposes the conclusion that the match is positive. This additional element is in fact used in drawing conclusions from the AF: first the union of all preferred extensions is taken and if all arguments in the union have the same conclusion, then that conclusion is drawn, otherwise no conclusion is drawn (Note, by the way, that the method does not prevent that arguments for opposite conclusions are in the same extension.)

While thus there is some structure in the arguments, it is very limited: there is no chaining of inferences and there is no distinction between types of inference rules, while the nature of attacks is left implicit. Moreover, there is no distinction between attack and defeat: presumably the expert implicitly encoded his preferences in his assignments of attacks. Does this mean that this is an application where the internal structure of arguments does not matter? I don't think so. The paper does not describe on which grounds the expert assigned the attack relations but an analysis in terms of *ASPIC*$^+$ may bring some clarity. The paper's examples only contain asymmetric attacks. At first sight this may seem surprising but when looking though an *ASPIC*$^+$ lens an explanation suggests itself: it may be that attackers state exceptions to rules of thumb underlying an argument. Consider again the following arguments:

A: The match is long, so the match is a good indicator of structure.
B: The identity between the match sequence and the query sequence is low, so the match is not a good indicator of structure.

The expert said that B attacks A while A does not attack B. It may be that the expert has thus expressed a rule-exception structure: low identity is an exception to the rule of thumb that long matches are positive.

Apparently the expert did not assign any attack relation on the ground that arguments support contradictory conclusions that a match is positive, respectively, negative. Strictly speaking the framework thus violates [20]'s consistency postulate. However, a reconstruction in *ASPIC*$^+$ is possible that respects this postulate. Let us formalise the above example in *ASPIC*$^+$ in such a way that both the nature of B as an undercutter of a rule of thumb and the contradictoriness of the conclusions of A and B is respected. The easiest way is with domain-specific inference rules:

r_1: The match is long $\Rightarrow$ the match is a good indicator of structure
r_2: The identity between the match sequence and the query sequence is low $\Rightarrow$
the match is not a good indicator of structure
r_3: The identity between the match sequence and the query sequence is low $\Rightarrow \neg r_1$.

Then r_3 can be used to undercut application of r_1 while r_2 can be used to draw the opposite conclusion. Inference can then simply be modelled by checking whether in all extensions there is an argument for the conclusion that the match is positive (respectively, negative). This is the second variant of skeptical inference described above just below Definition 7.

All in all [41] is a fascinating paper, since it applies formal argumentation to a genuine scientific application in a way that appears to be useful (the authors show that adding their AF to the search engine improves its performance). Moreover, although as I just argued the paper does assume some structure of arguments, it also shows that not all applications of argumentation need the full expressiveness of *ASPIC*$^+$ or, say, classical or assumption-based argumentation.

7 Conclusion

As I said in the introduction, the study of argumentation in AI is nowadays very popular, which is good, since our field has a lot of intellectual and application potential: unlike fixpoints and minimal models, arguments are a natural concept in many fields and professions, as well as in everyday conversation. However, I fear that if the characteristics of actual argumentation are ignored and the historic roots of our field are forgotten, this potential may not be realised. I have illustrated the first point in two ways. I first argued that if the use of preferences to resolve attacks is modelled without making the structure of arguments and the nature of attack explicit, then problems arise with respect to the rationality postulates of consistency and subargument closure. I then argued that if the defeasible nature of commonsense inference rules is not formalised as such, some common forms of defeasible reasoning cannot be represented in a natural way. With the latter I also illustrated the second point, by reminding the reader of similar discussions in the history of nonmonotonic logic.

As for deductive models of argumentation my conclusion is that they apply to just a minority of argumentation problems, namely, only those problems that can be modelled as inconsistency handling. Most realistic argumentation problems cannot be modelled as such, since they involve defeasible reasoning steps. This does not mean that research on deductive argumentation should stop; it definitely has its place in the study of argumentation but it must be combined with other aspects into a full account of argumentation-based inference. In all modesty I claim that *ASPIC*$^+$ provides such a full account, but I am sure that many readers will disagree. I look forward to their alternatives, as long as these give defeasibe inference the place it deserves.

8 Afterword: some personal observations on Marek Sergot

In 1993 Marek Sergot was the external examiner of my PhD thesis titled *Logical Tools for Modelling Legal Argument*. After my thesis defence I joined him for a year at Imperial College. Half of the time I worked with him on deontic logic, mainly on contrary-to-duty structures (though defeasibility played a role), the other half I continued my work on argumentation. My stay at Imperial was the most fascinating period of my academic life so far, not in the least because of my many meetings with Marek. I still think with great pleasure of the mixture of lucid analyses of research issues, insightful observations about academic life, gossip and, of course, the many hilarious stories. While the work I did with Marek on deontic logic resulted in some publications I am still proud of [63, 64, 65], my work on argumentation was less productive: basically I did not produce anything of interest during that year. Nevertheless, my stay at Imperial was still immensely useful, since I could be a witness to some exciting developments in the formal study of argumentation. I shared an office with Francesca Toni and Bob Kowalski's office was next door. When I arrived at Imperial, they were working on a workshop paper with the Russian visitor Andrei Bondarenko [15], which they later extended with Dung to [14], the official start of assumption-based argumentation. Towards the end of the year I met Dung, who came to Imperial for three months to work with Kowalski and Toni on their 1997 paper with Bondarenko. Just before Dung arrived, Francesca Toni gave me a copy of one of Dung's papers "just to get to know his work". It turned out to be a version of his now famous 1995 paper on abstract argumentation [25].

After my year in London I returned to Amsterdam for a postdoc fellowship. Marek and I continued working together on deontic logic for a few years, resulting in [64, 65]. However, during these years my research interested shifted more and more towards argumentation, and my three papers with Marek turned out to be my last publications on deontic logic (so far?). Nevertheless, as I discussed in Section 6, Marek has since then occasionally contributed to the field of argumentation, and for this reason I felt it was appropriate to write my contribution to this Festschrift on argumentation and to inform Marek of what has become of me after we parted.

Acknowledgements

In writing this paper I have profited from discussions with Martin Caminada and Sanjay Modgil. However, I take all responsibility for the opinions expressed in this paper. I thank all three reviewers for their many valuable suggestions and one of them in particular for triggering the discussion in footnote 10.

Bibliography

[1] L. Amgoud and Ph. Besnard. Bridging the gap between abstract argumentation systems and logic. In *Proceedings of the 3rd International Conference on Scalable Uncertainty (SUM'09)*, number 5785 in Springer Lecture Notes in AI, pages 12–27, Berlin, 2009. Springer Verlag.

[2] L. Amgoud and Ph. Besnard. A formal analysis of logic-based argumentation systems. In *Proceedings of the 4th International Conference on Scalable Uncertainty (SUM'10)*, number 6379 in Springer Lecture Notes in AI, pages 42–55, Berlin, 2010. Springer Verlag.

[3] L. Amgoud and C. Cayrol. On the acceptability of arguments in preference-based argumentation. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, pages 1–7, 1998.

[4] L. Amgoud and C. Cayrol. A model of reasoning based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, 34:197–215, 2002.

[5] L. Amgoud and S. Vesic. Two roles of preferences in argumentation frameworks. In *Proceedings of the 11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2011)*, number 6717 in Springer Lecture Notes in Computer Science, pages 86–97, Berlin, 1993. Springer Verlag.

[6] A. Artikis, M.J. Sergot, and J. Pitt. An executable specification of a formal argumentation protocol. *Artificial Intelligence*, 171:776–804, 2007.

[7] A.B. Baker and M.L. Ginsberg. A theorem prover for prioritized circumscription. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, pages 463–467, 1989.

[8] T.J.M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13:429–448, 2003.

[9] T.J.M. Bench-Capon and M.J. Sergot. Towards a rule-based representation of open texture in law. In C. Walter, editor, *Computing Power and Legal Language*, pages 39–60. Greenwood/Quorum Press, Westport, 1988.

[10] Ph. Besnard and A. Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128:203–235, 2001.

[11] Ph. Besnard and A. Hunter. *Elements of Argumentation*. MIT Press, Cambridge, MA, 2008.

[12] F.J. Bex. *Evidence for a Good Story. A Hybrid Theory of Arguments, Stories and Criminal Evidence*. Doctoral dissertation Faculty of Law, University of Groningen, 2009.

[13] F.J. Bex, H. Prakken, C. Reed, and D.N. Walton. Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law*, 12:125–165, 2003.

[14] A. Bondarenko, P.M. Dung, R.A. Kowalski, and F. Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93:63–101, 1997.

[15] A. Bondarenko, R.A. Kowalski, and F. Toni. An assumption-based fromework for nonmonotonic reasoning. In *Proceedings of the 2nd International Workshop on Logic Programming and Non-monotonic Reasoning*, pages 171–189, 1993.

[16] G. Brewka. Preferred subtheories: An extended logical framework for default reasoning. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence (IJCAI-89)*, pages 1043–1048, 1989.

[17] G. Brewka. *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge University Press, Cambridge, 1991.

[18] M. Caminada. On the issue of reinstatement in argumentation. In *Proceedings of the 11th European Conference on Logics in Artificial Intelligence (JELIA 2006)*, number 4160 in Springer Lecture Notes in AI, pages 111–123, Berlin, 2006. Springer Verlag.

[19] M. Caminada. On the issue of contraposition of defeasible rules. In Ph. Besnard, S. Doutre, and A. Hunter, editors, *Computational Models of Argument. Proceedings of COMMA 2008*, pages 109–115. IOS Press, Amsterdam etc, 2008.

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

[21] M. Caminada and Y. Wu. On the limitations of abstract argumentation. In *Proceedings of the 23rd Benelux Conference on Artificial Intelligence (BNAIC-11)*, Gent, Belgium, 2011.

[22] C. Cayrol. On the relation between argumentation and non-monotonic coherence-based entailment. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95)*, pages 1443–1448, 1995.

[23] S. Coste-Marquis, C. Devred, and P. Marquis. Constrained argumentation frameworks. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Tenth International Conference (KR-06)*, pages 112–122. AAAI Press, 2006.

[24] P.M. Dung. An argumentation semantics for logic programming with explicit negation. In *Proceedings of the Tenth Logic Programming Conference*, pages 616–630, Cambridge, MA, 1993. MIT Press.

[25] 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.

[26] P.M. Dung, R.A. Kowalski, and F. Toni. Assumption-based argumentation. In I. Rahwan and G.R. Simari, editors, *Argumentation in Artificial Intelligence*, pages 199–218. Springer, Berlin, 2009.

[27] P.M. Dung, P. Mancarella, and F. Toni. Computing ideal sceptical argumentation. *Artificial Intelligence*, 171:642–674, 2007.

[28] P.E. Dunne, P. McBurney, S. Parsons, and M. Wooldridge. Weighted argument systems: basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175:457–486, 2011.

[29] A.J. Garcia and G.R. Simari. Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming*, 4:95–138, 2004.

[30] M.L. Ginsberg. AI and nonmonotonic reasoning. In D. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 1–33. Clarendon Press, Oxford, 1994.

[31] T.F. Gordon. The Pleadings Game: formalizing procedural justice. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*, pages 10–19, New York, 1993. ACM Press.

[32] N. Gorogiannis and A. Hunter. Instantiating abstract argumentation with classical-logic arguments: postulates and properties. *Artificial Intelligence*, 175:1479–1497, 2011.

[33] J.C. Hage. A theory of legal reasoning and a logic to match. *Artificial Intelligence and Law*, 4:199–273, 1996.

[34] S. Hanks and D. McDermott. Default reasoning, nonmonotonic logics and the frame problem. In *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI-86)*, pages 328–333, 1986.

[35] J. Horty. Some direct theories of nonmonotonic inheritance. In D. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 111–187. Clarendon Press, Oxford, 1994.

[36] J. Horty, R.H. Thomason, and D.S. Touretzky. A skeptical theory of inheritance in nonmonotonic semantic networks. *Artificial Intelligence*, 42:311–348, 1990.

[37] A. Hunter. Reasoning about the appropriateness of proponents for arguments. In *Proceedings of the 23rd National Conference on Artificial Intelligence (AAAI-08)*, pages 89–94, 2008.

[38] D. Israel. What's wrong with non-monotonic logic? In *Proceedings of the First National Conference on Artificial Intelligence (AAAI-80)*, pages 99–101, 1980.

[39] H. Jakobovits. *On the Theory of Argumentation Frameworks*. Doctoral dissertation Free University Brussels, 2000.

[40] H. Jakobovits and D. Vermeir. Robust semantics for argumentation frameworks. *Journal of Logic and Computation*, 9:215–261, 1999.

[41] B. Jefferys, L.A. Kelley, M.J. Sergot, J. Fox, and M.J.E. Sternberg. Capturing expert knowledge with argumentation: a case study in bioinformatics. *Bioinformatics*, 22:924–933, 2006.

[42] A.C. Kakas, R.A. Kowalski, and F. Toni. Abductive logic programming. *Journal of Logic and Computation*, 2:719–770, 1992.

[43] F. Lin and Y. Shoham. Argument systems. A uniform basis for nonmonotonic reasoning. In *Principles of Knowledge Representation and Reasoning: Proceedings of the First International Conference*, pages 245–255, San Mateo, CA, 1989. Morgan Kaufmann Publishers.

[44] R.P. Loui. Defeat among arguments: a system of defeasible inference. *Computational Intelligence*, 2:100–106, 1987.

[45] S. Modgil. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, 173:901–934, 2009.

[46] S. Modgil and H. Prakken. Reasoning about preferences in structured extended argumentation frameworks. In P. Baroni, F. Cerutti, M. Giacomin, and G.R. Simari, editors, *Computational Models of Argument. Proceedings of COMMA 2010*, pages 347–358. IOS Press, Amsterdam etc, 2010.

[47] S. Modgil and H. Prakken. Revisiting preferences and argumentation. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-11*, pages 1021–1026, 2011.

[48] D. Nute. Defeasible logic. In D. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 253–395. Clarendon Press, Oxford, 1994.

[49] S. Parsons, M. Wooldridge, and L. Amgoud. Properties and complexity of some formal inter-agent dialogues. *Journal of Logic and Computation*, 13, 2003. 347-376.

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

[51] J.L. Pollock. Justification and defeat. *Artificial Intelligence*, 67:377–408, 1994.

[52] J.L. Pollock. *Cognitive Carpentry. A Blueprint for How to Build a Person*. MIT Press, Cambridge, MA, 1995.

[53] J.L. Pollock. A recursive semantics for defeasible reasoning. In I. Rahwan and G.R. Simari, editors, *Argumentation in Artificial Intelligence*, pages 173–197. Springer, Berlin, 2009.

[54] D.L. Poole. A logical framework for default reasoning. *Artificial Intelligence*, 36:27–47, 1988.

[55] H. Prakken. An argumentation framework in default logic. *Annals of Mathematics and Artificial Intelligence*, 9:91–132, 1993.

[56] H. Prakken. *Logical Tools for Modelling Legal Argument*. Doctoral dissertation Free University Amsterdam, 1993.

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

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

[59] H. Prakken. On the nature of argument schemes. In C. Reed and C. Tindale, editors, *Dialectics, Dialogue and Argumentation. An Examination of Douglas Walton's Theories of Reasoning and Argument*, pages 167–185. College Publications, London, 2010.

[60] H. Prakken. Reconstructing Popov v. Hayashi in a framework for argumentation with structured arguments and Dungean semantics. *The Knowledge Engineering Review*, 2011. To appear. Available at `http://www.cs.uu.nl/groups/IS/archive/henry/ker09.pdf`.

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

[62] H. Prakken and G. Sartor. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-classical Logics*, 7:25–75, 1997.

[63] H. Prakken and M.J.S. Sergot. Contrary-to-duty imperatives, defeasibility and violability. In *Proceedings of the Second International Workshop on Deontic Logic in Computer Science*, pages 296–318, Oslo, 1994. Tano.

[64] H. Prakken and M.J.S. Sergot. Contrary-to-duty obligations. *Studia Logica*, 57:91–115, 1996.

[65] H. Prakken and M.J.S. Sergot. Dyadic deontic logic and contrary-to-duty obligations. In D. Nute, editor, *Defeasible Deontic Logic*, volume 263 of *Synthese Library*, pages 223–262. Kluwer Academic Publishers, Dordrecht/Boston/London, 1997.

[66] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[67] N. Rescher. *Plausible Reasoning*. Van Gorcum, Assen, 1976.

[68] N. Rescher. *Dialectics: a Controversy-oriented Approach to the Theory of Knowledge*. State University of New York Press, Albany, N.Y., 1977.

[69] G. Sartor. *Legal Reasoning: a Cognitive Approach to the Law*. Springer Verlag, Berlin, 2005.

[70] G.R. Simari and R.P. Loui. A mathematical treatment of defeasible argumentation and its implementation. *Artificial Intelligence*, 53:125–157, 1992.

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

[72] F. Toni and M.J. Sergot. Argumentation and answer set programming. In M. Balduccini and T.C. Son, editors, *Logic Programming, Knowledge Representation and Nonmonotonic Reasoning: Essays Dedicated to Michael Gelfond on the Occasion of His 65th Birthday*, number 6565 in Springer Lecture Notes in Computer Science, pages 164–180. Springer Verlag, Berlin, 2010.

[73] B. Verheij. An integrated view on rules and principles. In R.W. van Kralingen et al., editor, *Legal Knowledge Based Systems. Foundations of Legal Knowledge Systems (Proceedings of JURIX-1996)*, pages 25–38, Tilburg, 1996. Tilburg University Press.

[74] B. Verheij. Two approaches to dialectical argumentation: admissible sets and argumentation stages. In *Proceedings of the Eighth Dutch Conference on Artificial Intelligence (NAIC-96)*, pages 357–368, Utrecht, The Netherlands, 1996.

[75] B. Verheij. Dialectical argumentation with argumentation schemes: an approach to legal logic. *Artificial Intelligence and Law*, 11:167–195, 2003.

[76] G.A.W. Vreeswijk. *Studies in Defeasible Argumentation*. Doctoral dissertation Free University Amsterdam, 1993.

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