Argumentation and rules with exceptions

Bart VERHEIJ
Artificial Intelligence, University of Groningen

Abstract. Models of argumentation often take a given set of rules or conditionals as a starting point. Arguments to support or attack a position are then built from these rules. In this paper, an attempt is made to develop constraints on rules and their exceptions in such a way that they correspond exactly to arguments that successfully support their conclusions. The constraints take the form of properties of nonmonotonic consequence relations, similar to the ones that have been studied for cumulative inference.

1. Introduction[1]

Consider the following information: p, p is a prima facie reason for q, q is a prima facie reason for r, r is an exception that undercuts the support of q by p. Formally: $p, p \Rightarrow q, q \Rightarrow r, r \Rightarrow \neg(p \Rightarrow q)$ (see the argument depicted on the left in Figure 1). The example is essentially the important and well-known example used by Pollock (1995, p. 119 [6]) in his groundbreaking work on defeasible argumentation. The issue with this theory can be summarized as follows:

> Either p successfully supports q, or it does not. Assume that it does. Then r is also successfully supported, as q supports r and that link is not disputed. But if r is successfully supported, the assumption that p successfully supports q is contradicted since r attacks the link between p and q. So assume the second possibility that p does not successfully support q. But only an attack by r can break the link between p and q, so r should be successfully supported. But it isn't, as the successful support of r requires the successful support of q based on p, which contradicts the assumption. Paradox!

There are two possible kinds of responses. A first kind is to reconsider the way in which we construct and evaluate arguments. This is the route taken by Pollock: he adapts his approach to the determination of defeat status. But there is a second kind of response, namely that the input information is in some sense flawed. For instance, the input theory could be 'inconsistent' or 'incomplete' in a sense that is relevant in the context of argumentation with pros and cons. It is this second kind of response that is pursued in the present paper. We will establish constraints for rules and their exceptions in such a way that they correspond closely to argumentation. The constraints will take the form of logical properties of a consequence relation associated with argumentative input information. With respect to non-monotonic logic more generally, a similar type of response has been followed in the important work that led to the theory of cumulative

[1]This technical report extends the paper of the same title presented at the COMMA 2010 conference. The main differences consist in the inclusion of proofs and the correction of an error (see Example 4.3).

inference (Kraus et al. 1990 [4]; see also the overview by Makinson 1994 [5]). Until now, the connection of nonmonotonic consequence relations with argumentation seems to only have been touched upon (but see Bochman 2005 [2]).

In this paper, the relation between the nonmonotonic consequence relations and arguments that successfully support their conclusion is studied. The following equivalence will be formally elaborated:

> ϕ defeasibly implies ψ ($\phi \mid\!\sim \psi$) if and only if there is an argument from ϕ that successfully supports the conclusion ψ.

For instance, when there is a witness testimony (t), of a witness claiming that the suspect was at the crime scene (p), an argument for the conclusion that the suspect committed the crime (c) can be expressed as $[t, t \Rightarrow p, p \Rightarrow c]$. If the conclusion c is successfully supported by the argument, c is a defeasible consequence of t, denoted $t \mid\!\sim c$. When the witness is lying (l), there is an attacking argument $[l]$, since lying makes a witness testimony unreliable ($l \Rightarrow \neg(t \Rightarrow p)$). As a result, the extended argument $[t, t \Rightarrow p, p \Rightarrow c, e]$ does not successfully support the conclusion c. If there are no other arguments, c is not a defeasible consequence of t and l together ($t \wedge l \not\mid\!\sim c$).

A useful role in the analysis is played by what is here called the *case made* by an argument, that will be defined as the conjunction of all claims made by the argument. For instance, the case made by the argument $[t, t \Rightarrow p, p \Rightarrow c]$ is $t \wedge p \wedge c$.

In our proposal, the normal situation is that, given certain premises, exactly one set of conclusions follows. However, in the context of argumentation it can also occur that more than one position is defensible; a choice still can be made, but the current premises do not decide the choice. Consider for example two witness testimonies (t_1 and t_2), the first witness claiming that the suspect was at the crime scene (p), the second that he was not ($\neg p$). If now both testimonies are of equal strength, no choice between $t_1 \wedge t_2 \wedge p$ and $t_1 \wedge t_2 \wedge \neg p$ can be made. In the formalism proposed here, it is a matter of the input information (not of logic) how this is addressed. It is possible, firstly, to make no choice (neither $t_1 \wedge t_2 \wedge p$ nor $t_1 \wedge t_2 \wedge \neg p$ is defensible), secondly, to not decide about p (making $t_1 \wedge t_2$ defensible, and leaving the status of p and $\neg p$ open) and, thirdly, to consider both options $t_1 \wedge t_2 \wedge p$ and $t_1 \wedge t_2 \wedge \neg p$ defensible. Each of these styles are reasonable under different circumstances (perhaps reflecting different proof standards), so each can be modeled. Of course there is also the possibility that one option is chosen, but that represents a different situation: then one of the testimonies is stronger. When for instance t_1 is stronger, $t_1 \wedge t_2 \wedge p$ is defensible, and $t_1 \wedge t_2 \wedge \neg p$ is not.

Our strategy is as follows. The start is a formalization of rules and the exceptions to them (section 2). Then the notions of argument, argument attack and defensibility are defined (section 3). The core of the paper is section 4 in which argumentation on the basis of rules with exceptions is studied in terms of the properties of nonmonotonic consequence relations. Then follows a discussion of the results (section 5), and the conclusion (section 6).

2. Rules with exceptions

As logical object language, we use a language L for standard truth-functional propositional calculus with connectives $\neg$, $\wedge$, $\vee$, $\leftrightarrow$ and $\top$, and its associated monotonic con-

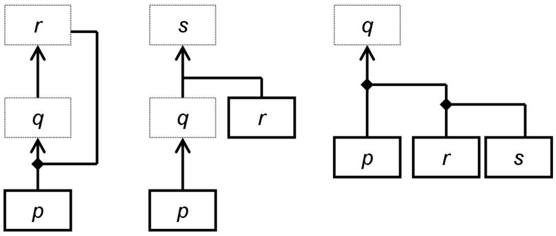

Figure 1. Some arguments

sequence relation, denoted $\vdash$. The language L is the language in which the premises, conclusions and exceptions that occur in an argument are expressed. Statements are connected in an argument on the basis of the inference and exception rules from which arguments are built. Such input information from which arguments are built and subsequently evaluated is formalized in the concept of a *(defeasible) rule system*.

Definition 2.1. A *(defeasible) rule system* is a triple $\mathcal{R} = (L, R, X)$, where R is a set of expressions of the form $\phi \Rightarrow \psi$, where ϕ and ψ are elements of L, and X is a set of expressions of the form $\neg(\phi \Rightarrow \psi)$, where ϕ and ψ are elements of L. The elements of R are the *inference rules* of the system $\mathcal{R}$, the elements of X the *exception rules*.

Inference rules are the *warrants* or *licenses* that allow one to draw (possibly defeasible) conclusions, while exception rules express the *prohibition* to draw a conclusion. Exception rules are here treated on a par with inference rules: both are part of the input information. The conditional connective $\Rightarrow$ is not used in L, so an inference rule $\phi \Rightarrow \psi$ is not an element of L, nor is an exception rule $\neg(\phi \Rightarrow \psi)$. An example of an inference rule is — following the famous example by Pollock — that when an object looks red, it can be inferred that it is red. Formally: $l \Rightarrow r$. An example of an exception rule excluding this inference rule is that when an object looks red and is illuminated by a red light, it cannot be inferred that it is red. Formally: $\neg(l \wedge i \Rightarrow r)$.

In the following, rule systems will be *logical*, in the sense that sentences can be replaced by logically equivalent sentences (from the perspective of the language L and its consequence relation $\vdash$): When $\vdash \phi \leftrightarrow \phi'$ and $\vdash \psi \leftrightarrow \psi'$ it holds that $\phi \Rightarrow \psi \in R$ if and only if $\phi' \Rightarrow \psi' \in R$, and that $\neg(\phi \Rightarrow \psi) \in X$ if and only if $\neg(\phi' \Rightarrow \psi') \in X$.

Exception rules can express exceptions to inference rules. We use the following terminology.

Definition 2.2. Let $\mathcal{R} = (L, R, X)$ be a rule system. A sentence ϵ expresses an *exception to a rule* $\phi \Rightarrow \psi$ in R if $\neg(\phi \wedge \epsilon \Rightarrow \psi) \in X$. A sentence δ *reinstates* the inference from ϕ to ψ excluded by the exception ϵ if $\phi \wedge \epsilon \wedge \delta \Rightarrow \psi \in R$.

Notation: An exception rule $\neg(\phi \wedge \epsilon \Rightarrow \psi)$ is also written $\epsilon \Rightarrow \neg(\phi \Rightarrow \psi)$. An inference rule $\phi \wedge \epsilon \wedge \delta \Rightarrow \psi$ reinstating an inference rule $\phi \Rightarrow \psi$ excluded by the exception rule $\neg(\phi \wedge \epsilon \Rightarrow \psi)$ is also written $\delta \Rightarrow \neg(\epsilon \Rightarrow \neg(\phi \Rightarrow \psi))$. By this notation (which reflects the intended meaning), the inference and exception rules used in the argument on the left in Figure 1 can be represented as $p \Rightarrow q, q \Rightarrow r, r \Rightarrow \neg(p \Rightarrow q)$, as we did in the introduction. The argument on the right uses the rules $p \Rightarrow q$, $r \Rightarrow \neg(p \Rightarrow q)$ and $s \Rightarrow \neg(r \Rightarrow \neg(p \Rightarrow q))$. By this notation, the arrows in the figure correspond exactly to (nested) conditional sentences (cf. the relation between argument diagrams and the logical language of DefLog developed by Verheij 2003 [10]).

3. Arguments, argument attack and defensible cases

Given a rule system, arguments are the result of chaining inference rules, and counterarguments are the result of exception rules. Our arguments are *dialectical* in the sense that they can contain both pros and cons.

Definition 3.1. Let $\mathcal{R} = (L, R, X)$ be a rule system. Then the set of *arguments* is inductively defined as follows:

1. The empty list [] is an argument from $\top$ making the case $\top$.
2. If $[\alpha_0, \ldots, \alpha_n]$ (with each $\alpha_i \in L \cup R \cup X$) is an argument from ϕ making the case ψ (with $\phi \in L$ and $\psi \in L$), then
   (a) $[\alpha_0, \ldots, \alpha_n, \phi']$ with $\phi' \in L$ is an argument from $\phi \wedge \phi'$ making the case $\psi \wedge \phi'$.
   (b) $[\alpha_0, \ldots, \alpha_n, \phi' \Rightarrow \psi']$ with $\phi' \Rightarrow \psi' \in R$ and $\psi \vdash \phi'$ is an argument from ϕ making the case $\psi \wedge \psi'$.
   (c) $[\alpha_0, \ldots, \alpha_n, \neg(\phi' \Rightarrow \psi')]$ with $\neg(\phi' \Rightarrow \psi') \in X$ and $\psi \vdash \phi'$ is an argument from ϕ making the case ψ.

An argument $[\alpha_0, \ldots, \alpha_n]$ from ϕ making the case ψ has each $\chi \in L$ with $\phi \vdash \chi$ as a *premise*, and each $\chi \in L$ with $\psi \vdash \chi$ as a *conclusion*.

The case made by an argument can be thought of as the 'overall position' supported by the argument; it is the conjunction of all claims made in the argument. Adding a premise or applying an inference rule can extend the case made by an argument, but adding an exception rule does not. For instance, the argument $[p, p \Rightarrow q, q \Rightarrow r]$ is an argument from p making the case $p \wedge q \wedge r$. The argument $[p, p \Rightarrow q, q \Rightarrow r, r \Rightarrow \neg(p \Rightarrow q)]$ (the one on the left in Figure 1), which is equivalent to $[p, p \Rightarrow q, q \Rightarrow r, \neg(p \wedge r \Rightarrow q)]$, is an argument from p making the case $p \wedge q \wedge r$. The argument in the middle of the figure, $[p, p \Rightarrow q, r, q \wedge r \Rightarrow s]$, is an argument from $p \wedge r$ with s as one of its conclusions. It is an argument making the case $p \wedge q \wedge r \wedge s$. The argument on the right is $[p, p \Rightarrow q, r, r \Rightarrow \neg(p \Rightarrow q), s, s \Rightarrow \neg(r \Rightarrow \neg(p \Rightarrow q))]$, or equivalently, $[p, p \Rightarrow q, r, \neg(p \wedge r \Rightarrow q), s, p \wedge r \wedge s \Rightarrow q]$, which is an argument from $p \wedge r \wedge s$ with q as a conclusion. The argument makes the case $p \wedge q \wedge r \wedge s$.

In the following, we will make a logicality assumption, as follows. When A is an argument from ϕ to ψ, while $\vdash \phi \leftrightarrow \phi'$ and $\vdash \psi \leftrightarrow \psi'$, then we will also say that A is an argument from ϕ' to ψ'.

Argument attack is defined in terms of exception rules: argument attack occurs when an argument supports an exception to a conclusion of another argument. Coherent arguments do not attack themselves, and defensible arguments attack their attackers.

Definition 3.2. Let $\mathcal{R} = (L, R, X)$ be a rule system and A an argument from ϕ with a conclusion ψ. Then an argument A' *attacks* A if A' has a conclusion ψ' such that $\neg(\phi \wedge \psi' \Rightarrow \psi) \in X$. An argument is *coherent* if it does not attack itself; otherwise *incoherent*. An argument A from ϕ *defends its case* if A is coherent and attacks all coherent arguments from ϕ that attack A.[2]

An attacking argument can contain the exception rule needed for the attack, but does not have to. So when $\neg(p_1 \wedge e \Rightarrow q)$ (which is the same as $e \Rightarrow \neg(p \Rightarrow q)$) is an exception rule, both $[p_2, p_2 \Rightarrow e]$ and $[p_2, p_2 \Rightarrow e, \neg(p_1 \wedge e \Rightarrow q)]$ attack $[p_1, p_1 \Rightarrow q]$. When $e \Rightarrow \neg(p \Rightarrow q) \in X$, the argument $[e]$ attacks the argument $[p, p \Rightarrow q]$. When $\neg(p \wedge r \Rightarrow q) \in X$, the argument $[p, p \Rightarrow q, q \Rightarrow r]$ is incoherent. To be defensible, an argument needs to attack only those attacking arguments that start from the same premises. For instance, if e is an exception to $p \Rightarrow q$ (and there are no other exceptions in the system), then $[p, p \Rightarrow q]$ is defensible, while $[p, p \Rightarrow q, e]$ is not, as the latter does not defend itself against the attack by the argument $[p, e]$.

It can occur that different cases are defensible, while they do not go together. For instance, one can perhaps both defend a weekend trip to Paris or one to London, but not both (cf. the so-called Nixon diamond in non-monotonic logic). A formal example in the present setting is the following. Consider the rule system with $p \Rightarrow q$, $p \Rightarrow r$, $r \Rightarrow \neg(p \Rightarrow q)$ and $q \Rightarrow \neg(p \Rightarrow r)$ as only (inference and exception) rules. Then the argument $[p, p \Rightarrow q]$ is coherent. The only argument from the same premises attacking $[p, p \Rightarrow q]$ is $[p, p \Rightarrow r]$, and this attack is attacked. So $[p, p \Rightarrow q]$ defends its case. In this system a second argument is defensible, namely by the argument $[p, p \Rightarrow r]$. However, the two cases cannot be defended simultaneously, as they exclude each other. This is reflected by the fact that the argument $[p, p \Rightarrow q, p \Rightarrow r]$ attacks itself.

4. Argumentation on the basis of rules with exceptions as nonmonotonic inference

Now we can start our study of connections between argumentation and rules with exceptions from the perspective of non-monotonic consequence relations. Consider the following properties of consequence relations:

1. (Logical equivalence)
   If $\phi \mid\!\sim \psi$, $\vdash \phi \leftrightarrow \phi'$ and $\vdash \psi \leftrightarrow \psi'$, then $\phi' \mid\!\sim \psi'$.
2. (Restricted reflexivity)
   If $\phi \mid\!\sim \psi$, then $\phi \mid\!\sim \phi$.
3. (Antecedence)
   If $\phi \mid\!\sim \psi$, then $\phi \mid\!\sim \phi \wedge \psi$.
4. (Right weakening)
   If $\phi \mid\!\sim \psi$ and $\psi \vdash \chi$, then $\phi \mid\!\sim \chi$.
5. (Conjunctive cautious monotony)
   If $\phi \mid\!\sim \psi \wedge \chi$, then $\phi \wedge \psi \mid\!\sim \chi$.

These properties characterize the arguments that do not attack themselves, in the sense of the following two theorems.

Theorem 4.1. *Let $\mathcal{R} = (L, R, X)$ be a rule system and let $\phi \mid\!\sim_c \psi$ denote that there is a coherent argument from ϕ with a conclusion ψ. Then $\mid\!\sim_c$ obeys the properties (1) to (5) above.*

[2]In the COMMA 2010 version of this text, it was omitted that only *coherent* arguments need to be defended against.

Proof. (1): Let A be a coherent argument from ϕ with a conclusion ψ, and ω the case made by A. Then, by Definition 3.1, A has ψ' as a conclusion. By the assumption of logicality of arguments, A is an argument from ϕ'. (The coherence of A is not used.)

(2): Let A be a coherent argument from ϕ with a conclusion ψ, and ω the case made by A. Then it follows by induction on the definition of arguments (Definition 3.1) that $\omega \vdash \phi$. Hence ϕ is a conclusion of A. (The coherence of A is not used.)

(3): Let A be a coherent argument from ϕ with a conclusion ψ. Then, as under the proof of (2), it has ϕ as one of its conclusions. By definition 3.1, the conjunction of two conclusions of an argument is also a conclusion of the argument. Hence, A has $\phi \wedge \psi$ as a conclusion. (The coherence of A is not used.)

(4): Let A be a coherent argument from ϕ with a conclusion ψ, and ω the case it makes. Then $\omega \vdash \psi$, and because $\psi \vdash \chi$ also $\omega \vdash \chi$. In other words, χ is a conclusion of A. (The coherence of A is not used.)

(5): Let $A = [\alpha_0, \ldots, \alpha_n]$ be a coherent argument from ϕ with a conclusion $\psi \wedge \chi$. Then $A' = [\phi \wedge \psi, \alpha_0, \ldots, \alpha_n]$ is an argument from $\phi \wedge \psi$ with a conclusion χ. (Here the logicality assumption for arguments is used.) Assume that A' is incoherent. Then it has conclusions ψ' and ψ'' such that $\neg(\phi \wedge \psi \wedge \psi' \Rightarrow \psi'') \in X$. But $\psi \wedge \psi'$ and ψ'' are conclusions of A, so then A would attack itself; contradiction.

□

Theorem 4.2. *Let* $\mid\!\sim$ *be a consequence relation obeying the properties (1) to (5) above, and* $\mathcal{R}_{\mid\sim}$ *the associated rule system defined by* $R_{\mid\sim} := \{\phi \Rightarrow \psi \mid \phi \mid\!\sim \psi\}$*, and* $X_{\mid\sim} := \{\neg(\phi \Rightarrow \psi) \mid \phi \not\mid\!\sim \psi\}$*. Then the following are equivalent:*

1. $\phi \mid\!\sim \psi$
2. $[\phi, \phi \Rightarrow \psi]$ *is a coherent* $\mathcal{R}_{\mid\sim}$*-argument.*
3. *There is an* $\mathcal{R}_{\mid\sim}$*-argument from* ϕ *with a conclusion* ψ *that is coherent.*

Proof. $1 \Rightarrow 2$: Since $\phi \mid\!\sim \psi$, $\phi \Rightarrow \psi$ is an element of R, and $[\phi, \phi \Rightarrow \psi]$ is an $\mathcal{R}_{\mid\sim}$-argument from ϕ making the case $\phi \wedge \psi$. Assume it is not coherent. Then it has conclusions ψ' and ψ'' such that $\neg(\phi \wedge \psi' \Rightarrow \psi'') \in X_{\mid\sim}$. By $\phi \mid\!\sim \psi$, (Antecedence) gives $\phi \mid\!\sim \phi \wedge \psi$. Since $\phi \wedge \psi \vdash \psi'$, (Logical equivalence) and (Conjunctive cautious monotony) lead to $\phi \wedge \psi' \mid\!\sim \phi \wedge \psi$. But also $\phi \wedge \psi \vdash \psi''$, so by (Right weakening), $\phi \wedge \psi' \mid\!\sim \psi''$. This contradicts $\neg(\phi \wedge \psi' \Rightarrow \psi'') \in X_{\mid\sim}$.

$2 \Rightarrow 3$: $[\phi, \phi \Rightarrow \psi]$ is an argument from ϕ making the case $\phi \wedge \psi$, so it has ψ as a conclusion.

$3 \Rightarrow 1$: Let A be an $\mathcal{R}_{\mid\sim}$-argument as under 3. Assume $\phi \not\mid\!\sim \psi$. As a result, $\neg(\phi \Rightarrow \psi) \in X_{\mid\sim}$. But then A attacks itself, as ϕ and ψ are conclusions of A. Contradiction. □

When properties (1) to (5) hold, we speak of *coherent argumentation*.

A coherent argument does not always defend itself. Consider for instance the rule system consisting of the inference rules $a \Rightarrow b$ and $a \Rightarrow c$ and the exception rule $\neg(a \wedge c \Rightarrow b)$. Then the argument $[a, a \Rightarrow b]$ from a making the case $a \wedge b$ is coherent, but does not defend itself against the attack by $[a, a \Rightarrow c]$.

The following additional property expresses defensibility:

6. (Mutual attack)
   If $\phi \mid\!\sim \psi$, $\phi \mid\!\sim \chi$ and $\phi \wedge \psi \not\mid\!\sim \chi$, then $\phi \wedge \chi \not\mid\!\sim \psi$.

Example 4.3. Consider $\mathcal{R} = (L, R, X)$ with $R = \{a \Rightarrow b, a \Rightarrow c, a \wedge b \Rightarrow c, a \wedge c \Rightarrow d\}$ and $X = \{\neg(a \wedge c \wedge d \Rightarrow b)\}$. Let $\phi \mathrel{|\!\sim}_d \psi$ denote that there is a defensible argument from ϕ with a conclusion ψ. Then:

1. $a \mathrel{|\!\sim}_d b$ because the argument $[a, a \Rightarrow b]$ is not attacked and is hence defensible.
2. $a \mathrel{|\!\sim}_d c$ because the argument $[a, a \Rightarrow c]$ is not attacked and is hence defensible.
3. $a \wedge b \mathrel{|\!\sim}_d c$ because the argument $[a \wedge b, a \Rightarrow c]$ is not attacked and is hence defensible.
4. $a \wedge c \mathrel{|\!\not\sim}_d b$ because there is no defensible argument from $a \wedge c$ with conclusion b: the argument $[a \wedge c, a \wedge c \Rightarrow d]$ attacks all arguments that use $a \Rightarrow b$, the rule that is needed to derive b, and there is no defense against this attack.

This example shows that $\mathrel{|\!\sim}_d$ does not in general obey (Mutual attack) (in contrast with what was erroneously reported in the COMMA 2010 conference paper).

Theorem 4.4. *Let* $\mathrel{|\!\sim}$ *be a consequence relation obeying the properties (1) to (6) above, and* $\mathcal{R}_{\mathrel{|\!\sim}}$ *the associated rule system defined by* $R_{\mathrel{|\!\sim}} := \{\phi \Rightarrow \psi \mid \phi \mathrel{|\!\sim} \psi\}$, *and* $X_{\mathrel{|\!\sim}} := \{\neg(\phi \Rightarrow \psi) \mid \phi \mathrel{|\!\not\sim} \psi\}$. *Then the following are equivalent:*

1. $\phi \mathrel{|\!\sim} \psi$
2. $[\phi, \phi \Rightarrow \psi]$ *is a defensible* $\mathcal{R}_{\mathrel{|\!\sim}}$*-argument.*
3. *There is an* $\mathcal{R}_{\mathrel{|\!\sim}}$*-argument from* ϕ *with a conclusion* ψ *that is defensible.*

Proof. $1 \Rightarrow 2$: Since $\phi \mathrel{|\!\sim} \psi$, $\phi \Rightarrow \psi$ is an element of $R_{\mathrel{|\!\sim}}$, and $A = [\phi, \phi \Rightarrow \psi]$ is an $\mathcal{R}_{\mathrel{|\!\sim}}$-argument from ϕ making the case $\phi \wedge \psi$. A's coherence follows as in the proof of Theorem 4.2. Let A' be a coherent argument from ϕ that attacks A. So A has a conclusion χ and A' a conclusion χ', such that $\neg(\phi \wedge \chi' \Rightarrow \chi) \in X_{\mathrel{|\!\sim}}$. Hence $\phi \wedge \chi' \mathrel{|\!\not\sim} \chi$. Also, by Theorem 4.2, $\phi \mathrel{|\!\sim} \chi'$. By (Right weakening), also $\phi \mathrel{|\!\sim} \chi$. Therefore by (Mutual attack), $\phi \wedge \chi \mathrel{|\!\not\sim} \chi'$. This implies that $\neg(\phi \wedge \chi \Rightarrow \chi') \in X_{\mathrel{|\!\sim}}$, so A attacks A'.

$2 \Rightarrow 3$: $[\phi, \phi \Rightarrow \psi]$ is an argument from ϕ making the case $\phi \wedge \psi$, so it has ψ as a conclusion.

$3 \Rightarrow 1$: Let A be an $\mathcal{R}_{\mathrel{|\!\sim}}$-argument as under 3. Assume $\phi \mathrel{|\!\not\sim} \psi$. As a result, $\neg(\phi \Rightarrow \psi) \in X_{\mathrel{|\!\sim}}$. But then A attacks itself, as ϕ and ψ are conclusions of A. Contradiction, as defensible arguments are coherent. □

When properties (1) to (6) hold, we speak of *defensible argumentation*.

Can an argument always be extended by an inference rule to which there is no exception? The answer is no. Consider for instance the rule system consisting of the inference rules $a \Rightarrow b$ and $a \wedge b \Rightarrow c$ and the exception rule $\neg(a \Rightarrow b \wedge c)$. Then the argument $[a, a \Rightarrow b]$ making the case $a \wedge b$ is coherent and the inference rule $a \wedge b \Rightarrow c$ has a satisfied antecedent and no exception given $a \wedge b$. Still $A = [a, a \Rightarrow b, a \wedge b \Rightarrow c]$ is not coherent as it attacks itself. One can say that the argument is defeated by 'sequential weakening' (cf. Verheij 1996 [9]). The following property forbids such defeat by sequential weakening.

7. (Conjunctive cumulative transitivity, Conjunctive cut)
   If $\phi \mathrel{|\!\sim} \psi$ and $\phi \wedge \psi \mathrel{|\!\sim} \chi$, then $\phi \mathrel{|\!\sim} \psi \wedge \chi$.

Given the other properties, this property is equivalent to (Conclusions are compatible or mutually exclusive) 'If $\phi \mathrel{|\!\sim} \psi$ and $\phi \mathrel{|\!\sim} \chi$, then either $\phi \mathrel{|\!\sim} \psi \wedge \chi$ or ($\phi \wedge \psi \mathrel{|\!\not\sim} \chi$ and $\phi \wedge \chi \mathrel{|\!\not\sim} \psi$)'.

When there is no defeat by sequential weakening, coherent argumentation implies defensible argumentation, as can be seen using the following proposition.

Proposition 4.5. *If a consequence relation obeys (Logical equivalence), (Conjunctive cautious monotony) and (Conjunctive cumulative transitivity, Conjunctive cut), then it obeys (Mutual attack).*

Proof. Assume $\phi \mid\sim \psi$, $\phi \mid\sim \chi$ and $\phi \wedge \chi \mid\sim \psi$. Then by (Conjunctive cautious transitivity, Conjunctive cut), $\phi \mid\sim \chi \wedge \psi$. Hence by (Logical equivalence), $\phi \mid\sim \psi \wedge \chi$. Now (Conjunctive cautious monotony) gives $\phi \wedge \psi \mid\sim \chi$. □

When properties (1) to (7) hold, we speak of *reason-based argumentation*.

5. Discussion

We have looked at argumentation and rules with exceptions from the perspective of non-monotonic consequence relations.

The first kind of inference that we distinguished was *coherent argumentation*, obeying the properties (1) to (5). It is related to the system that Bochman (2001, p. 140 [1]) refers to as basic inference and that goes back to van Benthem (1984 [7]). Coherent argumentation is somewhat more minimal than basic inference. For basic inference, (Restricted reflexivity) also obtains, but in the unrestricted form (Reflexivity) $\phi \mid\sim \phi$. We keep the restriction to allow the possibility of incoherent premises, i.e., premises from which no conclusions can be drawn (the 'no extensions' possibility in non-monotonic logic). Basic inference also has the property (Deduction) 'If $\phi \wedge \psi \mid\sim \chi$, then $\phi \mid\sim \psi \rightarrow \chi$', which does not in general obtain for our systems. The property (Conjunctive cautious monotony) of basic inference holds for coherent argumentation. In our approach, this property reflects that an argument does not attack itself.

The second kind of inference was *defensible argumentation*, obeying the properties (1) to (6). Dung's notion of admissibility (1995 [3]) for abstract argumentation frameworks is close in spirit to our notion of a defensible argument: an admissible set of arguments in the sense of Dung is one that does not attack itself and that attacks its attackers. However, whereas Dung abstracts from arguments structure and argument attack is given, here arguments are constructed using (defeasible) inference rules, and argument attack is the result of exception rules. Also, for the most constrained kind of inference that we have considered, reason-based argumentation, the argument attack relation is particularly well-behaved: arguments are either compatible or mutually attacking.

The third kind of inference that we distinguished is *reason-based argumentation*. Of the three systems proposed, it is closest to cumulative inference as studied by Kraus et al. (1990 [4]). Like basic inference mentioned above, cumulative inference has full (Reflexivity), but a more important difference is that it has (And) 'If $\phi \mid\sim \psi$ and $\phi \mid\sim \chi$, then $\phi \mid\sim \psi \wedge \chi$'. By (And), in cumulative inference, conclusions can always be drawn simultaneously, while in our systems one conclusion can exclude another. An example is the rule system consisting of the inference rules $a \Rightarrow b$ and $a \Rightarrow c$ and the exception rules $\neg(a \wedge b \Rightarrow c)$ and $\neg(a \wedge c \Rightarrow b)$. Then $a \mid\sim b$ and $a \mid\sim c$ (using defensible arguments), but $a \not\mid\sim b \wedge c$ (there is even no coherent argument). Cumulative inference also has a stronger version of our (Cumulative cautious monotony), namely

(Cautious monotony) 'If $\phi \mid\sim \psi$ and $\phi \mid\sim \chi$, then $\phi \wedge \psi \mid\sim \chi$'. When (And) holds, the difference disappears. The property (Cut) 'If $\phi \mid\sim \psi$ and $\phi \wedge \psi \mid\sim \chi$, then $\phi \mid\sim \chi$' of cumulative inference holds for reason-based argumentation, but in the stronger form of (Conjunctive cumulative transitivity, Conjunctive cut). Since (And) does not obtain for reason-based argumentation, (Cut) is a bit too weak for what is needed; our (Conjunctive cumulative transitivity, Conjunctive cut) does the job. It seems to be a new proposal. As a consequence of the fact that our constraints are not the same as those for cumulative inference, the preferential model semantics of cumulative inference does not apply to reason-based argumentation. However, since the constraints of reason-based argumentation inference do obtain for cumulative inference, our results connecting rules with exceptions and argumentation work also for cumulative inference. In other words, one could say that our approach provides an 'argumentation interpretation' of cumulative inference.

Veltman (1996 [8]) has studied rules with exceptions semantically. He does so in the context of update semantics. In his system, rules can have what he calls 'nonaccidental exceptions', i.e., exceptions that obtain conditionally. He focuses on contradicting exceptions and does not treat Pollock's undercutting defeaters.

6. Conclusion

In this paper, argumentation has been formalized in such a way that the arguments that are constructed using rules with exceptions can be studied from the perspective of nonmonotonic consequence relations. Properties have been given that characterize when arguments are coherent (i.e., not self-attacking) and when they are defensible. Three kinds of inference have been distinguished in terms of the properties of consequence relations: coherent argumentation, in which arguments do not attack themselves, defensible argumentation, in which arguments attack all their attackers, and reason-based argumentation, in which arguments can always be extended when there is a non-excluded rule with satisfied antecedent. In reason-based argumentation, there is no 'defeat by sequential weakening' (cf. Verheij 1996 [9]).

References

[1] A. Bochman. *A logical theory of nonmonotonic inference and belief change*. Springer, Berlin, 2001.

[2] A. Bochman. Propositional argumentation and causal reasoning. In *International Joint Conference on Artificial Intelligence*, volume 19, pages 388–393, 2005.

[3] 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.

[4] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

[5] D. Makinson. General patterns in nonmonotonic reasoning. In D.M. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming. Volume 3. Nonmonotonic Reasoning and Uncertain Reasoning*, pages 35–110. Clarendon Press, Oxford, 1994.

[6] J.L. Pollock. *Cognitive Carpentry: A Blueprint for How to Build a Person*. The MIT Press, Cambridge (Massachusetts), 1995.

[7] J. van Benthem. Foundations of conditional logic. *Journal of Philosophical Logic*, pages 303–349, 1984.

[8] F. Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996.

[9] B. Verheij. *Rules, Reasons, Arguments. Formal studies of argumentation and defeat.* Dissertation Universiteit Maastricht, Maastricht, 1996.

[10] B. Verheij. Artificial argument assistants for defeasible argumentation. *Artificial Intelligence*, 150(1–2):291–324, 2003.