Jumping to Conclusions

A Logico-Probabilistic Foundation for Defeasible Rule-Based Arguments

Bart Verheij

Artificial Intelligence, University of Groningen

Abstract. A theory of defeasible arguments is proposed that combines logical and probabilistic properties. This logico-probabilistic argumentation theory builds on two foundational theories of nonmonotonic reasoning and uncertainty: the study of nonmonotonic consequence relations (and the associated minimal model semantics) and probability theory. A key result is that, in the theory, qualitatively defined argument validity can be derived from a quantitative interpretation. The theory provides a synthetic perspective of arguments 'jumping to conclusions', rules with exceptions, and probabilities.

1 Introduction

Jumping to conclusions is a necessary and oft-used skill. We hear a voice on the phone, and conclude it's our father's. We smell foul coffee, and conclude it's from that dreaded machine down the hall. We find a note on the kitchen table, and conclude that our son has gone out. But sometimes we jump too far. It's not our father, but his brother. It's not coffee from *that* machine, but from a similar one on the next floor. And we find our son in his room at home, playing his favorite computer game, as his message was yesterday's.

In this paper, a mathematical theory of jumping to conclusions is developed. The theory's starting point is that jumping to conclusions is allowed ('valid', using a heavily laden term) when the conclusions do not lead us too far from the premises. Or, to be a bit more precise, when *the case made* on the basis of the premises, is close to those premises. Here 'the case made' is defined as the conjunction of premises and conclusions. In other words, we can jump to certain conclusions, if adding them to the premises is not a jump too far.

For instance, when a witness says she saw the suspect at the crime scene (w), we 'jump to' the conclusion that the suspect was indeed there (s) by making the case $w \wedge s$. A valid jump will be written $\varphi \mathrel{|\!\sim} \psi$, an invalid jump as $\varphi \mathrel{|\!\not\sim} \psi$. That the corresponding case made is sufficiently close to the premises is written as $\varphi \sim \varphi \wedge \psi$, that it is too far a jump as $\varphi \succ \varphi \wedge \psi$.

Our theory formalizes *ampliative reasoning*, as it has been called by Peirce, that is: reasoning that goes beyond the premises. Toulmin used the term *substantial reasoning* for the same concept, and considered reasoning only interesting when it adds information to what is already in the premises.

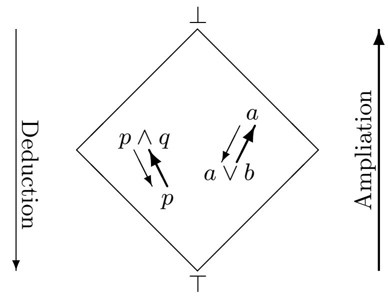


Fig. 1. Deduction and ampliation in a Boolean lattice

In Section 2, ampliative inference is studied in relation to nonmonotonic inference. In Section 3, a connection is made to the mathematical (pre)order relations, as a first step towards the quantitative interpretation, given in Section 4. It is shown that well-behaved ampliative inference can be derived from an 'argument value' function on the language with properties close to those of a standard probability function. In Section 5, a connection to argumentation theory is made by considering arguments with local structure in terms of premises, rules, and exceptions. It is shown that the global validity of an ampliative argument can be determined by the application of non-excluded rules.

2 Ampliative arguments

The propositions that occur in arguments are expressed in a classical language L with BNF specification $\varphi ::= \top \mid \bot \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \leftrightarrow \psi$, and the associated classical, deductive, monotonic consequence relation, denoted $\vdash$. As in the study of nonmonotonic inference, we write $\varphi \mathrel{|\!\sim} \psi$ to denote the validity of an ampliative argument from premises φ to conclusions ψ (plural to allow us to speak of separate premises/conclusions that are in an argument joined by conjunction). The proposition $\varphi \wedge \psi$ is the *case made* by the argument.

In the Boolean lattice associated with our language L, ampliative and deductive arguments have opposite directions (Figure 1, with $\top$ at the bottom and $\bot$ on top). In Figure 1, we can jump from p to $p \wedge q$ and from $a \vee b$ to a; written as $p \mathrel{|\!\sim} p \wedge q$ and $a \vee b \mathrel{|\!\sim} a$.

The following qualitative properties, well-known in the general study of nonmonotonic inference relations [2, 11, 12], are our starting point. For all propositions φ, φ', ψ, ψ', $\chi \in L$:

(LE) If $\varphi \mathrel{|\!\sim} \psi$, $\vdash \varphi \leftrightarrow \varphi'$ and $\vdash \psi \leftrightarrow \psi'$, then $\varphi' \mathrel{|\!\sim} \psi'$.
(Ant) If $\varphi \mathrel{|\!\sim} \psi$, then $\varphi \mathrel{|\!\sim} \varphi \wedge \psi$.
(PR) If $\varphi \mathrel{|\!\sim} \varphi \wedge \psi$, then $\varphi \mathrel{|\!\sim} \psi$.
(R) $\varphi \mathrel{|\!\sim} \varphi$.
(RW) If $\varphi \mathrel{|\!\sim} \psi \wedge \chi$, then $\varphi \mathrel{|\!\sim} \psi$.
(CCM) If $\varphi \mathrel{|\!\sim} \psi \wedge \chi$, then $\varphi \wedge \psi \mathrel{|\!\sim} \chi$.

(CCT) If $\varphi \mathrel{|\!\sim} \psi$ and $\varphi \wedge \psi \mathrel{|\!\sim} \chi$, then $\varphi \mathrel{|\!\sim} \psi \wedge \chi$.

(LE), for Logical Equivalence, expresses that in a valid ampliative argument the premises and the conclusions can be replaced by a classical equivalent (in the sense of $\vdash$). (Ant), for Antededence, expresses that when we can jump from certain premises to a conclusion, we can also jump to the case made by the argument (recall: the conjunction of conclusion and premises). Since (Ant) holds for ampliative arguments, every argument $\varphi \mathrel{|\!\sim} \psi$ has an associated 'ampliation', i.e., an argument of which the conclusion deductively implies the premises, namely $\varphi \mathrel{|\!\sim} \varphi \wedge \psi$. (PR), for Premise Reduction, says that we can also jump to a conclusion that classically follows from the case made by a valid argument. As the converse of (Ant), it is technically useful below. (R), for Reflexivity, expresses the validity of the limiting case of jumping from premises to themselves. (RW), for Right Weakening, expresses that when the premises justify a composite conclusion also the intermediate conclusions are justified. It strengthens (PR) (given (LE)). (CCM), for Conjunctive Cautious Monotony, expresses that we can still jump to the case made by a valid argument when an intermediate conclusion is added to the argument's premises. (CCT), for Conjunctive Cumulative Transitivity, is a variation of the Cumulative Transitivity property (CT, also known as Cut) extensively studied in the literature (which has $\varphi \mathrel{|\!\sim} \chi$ instead of $\varphi \mathrel{|\!\sim} \psi \wedge \chi$ as a consequent). The variation may seem minor, but is essential in the absence of the (And) property (If $\varphi \mathrel{|\!\sim} \psi$ and $\varphi \mathrel{|\!\sim} \chi$, then $\varphi \mathrel{|\!\sim} \psi \wedge \chi$). Assuming (Ant), (CCT) expresses the validity of chaining jumps from φ via $\varphi \wedge \psi$ to $\varphi \wedge \psi \wedge \chi$.

The relation $\sim$ associated with $\mathrel{|\!\sim}$ singles out those arguments that have the case made by the argument as conclusion, i.e., have a conclusion that logically implies the premises.

Definition 1. *For* $\mathrel{|\!\sim} \subseteq L \times L$, *we define:*

$\varphi \sim \psi := \psi \vdash \varphi$ *and* $\varphi \mathrel{|\!\sim} \psi$.

We now show that the properties of $\mathrel{|\!\sim}$ have close counterparts in terms of $\sim$. Beware: notwithstanding the suggestive notation, $\sim$ need not be symmetric.

(LEAmpl) If $\varphi \sim \psi$, $\vdash \varphi \leftrightarrow \varphi'$ and $\vdash \psi \leftrightarrow \psi'$, then $\varphi' \sim \psi'$.
(Ampl) If $\varphi \sim \psi$, then $\psi \vdash \varphi$.
(Eq) If $\vdash \varphi \leftrightarrow \psi$, then $\varphi \sim \psi$.
(Int) If $\chi \vdash \psi$, $\psi \vdash \varphi$ and $\varphi \sim \chi$, then $\varphi \sim \psi$ and $\psi \sim \chi$.
(Tr) If $\varphi \sim \psi$ and $\psi \sim \chi$, then $\varphi \sim \chi$.

(LEAmpl) expresses Logical Equivalence, this time for $\sim$. By (Ampl), for Ampliation, $\sim$ is an ampliation relation: when $\varphi \sim \psi$, ψ goes logically beyond φ. Also, by (Ampl), $\sim$ is not normally symmetric (which would come at the price of reducing $\sim$ to classical equivalence). (Eq), for Equivalence, says that a proposition's ampliations include all propositions logically equivalent to the proposition. (Int), for Interpolation, says that an ampliation can be split at a proposition that is logically in between. (Tr), for Transitivity, says that the ampliation relation is transitive.

The listed properties for $\mathrel{|\!\sim}$ and $\sim$ have close formal connections:

Proposition 1. *Let $\mathrel{|\!\sim}$ be an inference relation obeying (LE), (Ant) and (PR), and $\sim$ the associated ampliation relation. Then the following hold:*

1. *$\sim$ obeys (LEAmpl) and (Ampl).*
2. *$\varphi \mathrel{|\!\sim} \psi$ if and only if $\varphi \sim \varphi \wedge \psi$.*
3. *$\mathrel{|\!\sim}$ obeys (R) if and only if $\sim$ obeys (Eq).*
4. *$\mathrel{|\!\sim}$ obeys (RW) and (CCM) if and only if $\sim$ obeys (Int).*
5. *$\mathrel{|\!\sim}$ obeys (CCT) if and only if $\sim$ obeys (Tr).*

When the conditions (LE), (Ant) and (PR) of the proposition obtain, so that 1–5 follow, we speak of an *ampliative inference* relation $\mathrel{|\!\sim}$ and a corresponding *ampliation* relation $\sim$. When all properties listed obtain, we speak of *qualitative ampliative inference* and *qualitative ampliation*. The inference relation can be recovered from the ampliation relation (part 2 of the proposition).

An example of qualitative ampliative inference is the following, for which $p \mathrel{|\!\sim} q$, but $p \wedge e \mathrel{|\!\not\sim} q$, so e is an exception blocking the jump from p to q:

$\varphi \mathrel{|\!\sim} \psi$ if and only if ($p \wedge q \vdash \varphi \wedge \psi$ and $\varphi \vdash p$) or ($p \wedge e \vdash \varphi \wedge \psi$ and $\varphi \vdash p \wedge e$) or ($\phi \vdash \psi$).

A second example shows that we can sometimes jump to incompatible conclusions, e.g., in the case of two reasonable, but inconsistent decisions (as in a choice situation). Here we can jump from p to either q, or to $\neg q$, but not to their conjunction:

$\varphi \mathrel{|\!\sim} \psi$ if and only if ($p \wedge q \vdash \varphi \wedge \psi$ and $\varphi \vdash p$) or ($p \wedge \neg q \vdash \varphi \wedge \psi$ and $\varphi \vdash p$) or ($\phi \vdash \psi$).

This example of non-qualitative ampliative inference illustrates the role of (CCT) and (Tr):

$\varphi \mathrel{|\!\sim} \psi$ if and only if ($p \wedge q \vdash \varphi \wedge \psi$ and $\varphi \vdash p$) or ($p \wedge q \wedge r \vdash \varphi \wedge \psi$ and $\varphi \vdash p \wedge q$) or ($\varphi \vdash \psi$).

The relation $\mathrel{|\!\sim}$ obeys all properties listed above, but not (CCT)/(Tr) since it is not possible to chain the arguments $p \mathrel{|\!\sim} p \wedge q$ and $p \wedge q \mathrel{|\!\sim} p \wedge q \wedge r$ since $p \mathrel{|\!\not\sim} p \wedge q \wedge r$. Cf. the following figure.

$$\begin{tikzcd} \boxed{p} \arrow[r] \arrow[rr, bend right, "\times"'] & p \wedge q \arrow[r] & p \wedge q \wedge r \end{tikzcd} \qquad (1)$$

3 Ampliation and the ordering of propositions

In Section 4, we establish a quantitative interpretation of well-behaved ampliative inference in terms of a numeric 'argument value' function on the language. By this function, the propositions of the language can be ordered, in a way that is similar to a probability function that assigns decreasing probabilities to more specific propositions. As a first step, we consider certain (pre)order relations associated with ampliation. For instance, the ampliation relation itself is a preorder, since Reflexivity ($\varphi \sim \varphi$) holds by (LEAmpl) and (Eq), and Transitivity is part of the definition. Since Antisymmetry obtains for L's logical equivalence classes, $\sim$ is even a partial order on logical equivalence classes. But Symmetry (If $\varphi \sim \psi$, then $\psi \sim \varphi$) fails in general, so $\sim$ is not an equivalence relation.

The following definitions are needed.

Definition 2. *For an ampliative inference relation $\mid\!\sim$ and a corresponding ampliation relation $\sim$, we define:*

$\varphi \succ \psi := \psi \vdash \varphi$ *and* $\varphi \nsim \psi$
$\varphi \succsim \psi := \varphi \sim \psi$ *or* $\varphi \succ \psi$
$\varphi \prec \psi := \psi \succ \varphi$
$\varphi \precsim \psi := \psi \succsim \varphi$

As a result, $\varphi \succsim \psi$ is equivalent to $\psi \vdash \varphi$. Also, given $\psi \vdash \varphi$, $\varphi \mid\!\sim \psi$ is equivalent to $\varphi \sim \psi$, and $\varphi \mid\!\not\sim \psi$ is equivalent to $\varphi \succ \psi$. Furthermore, for all φ and ψ, exactly one of $\varphi \sim \psi$, $\varphi \succ \psi$, and $\psi \nvdash \varphi$ holds.

The following result says that, if we stay within an ampliative chain, $\sim$ and $\succ$ are well-defined 'up to $\sim$', although care is needed by the failure of Symmetry.

Proposition 2. *Let $\mid\!\sim$ be a qualitative ampliative inference relation. Assume $\chi \vdash \psi \vdash \varphi$.*

1. *When $\varphi \sim \psi$, the following hold:*
   (a) $\psi \sim \chi$ if and only if $\varphi \sim \chi$.
   (b) $\psi \succ \chi$ if and only if $\varphi \succ \chi$.
2. *When $\psi \sim \chi$, the following hold:*
   (a) $\varphi \sim \psi$ if and only if $\varphi \sim \chi$.
   (b) $\varphi \succ \psi$ if and only if $\varphi \succ \chi$.

If $\sim$ and $\succ$ are to be represented by numeric values, their properties should be sufficiently well-behaved *over the whole language*, and not just within ampliative chains. This requires a further property that expresses a kind of conservativeness across ampliation chains. Consider a case in which $\varphi \mid\!\sim \psi$ and $\psi \mid\!\sim \chi$ (cf. the following picture). The arrows indicate ampliations (in the upward direction of the arrow), the dotted lines deductions (in the downward direction).

$$\begin{tikzcd}
& \varphi \wedge \psi \arrow[dr, dash, dashed, "\precsim"'] & & \psi \wedge \chi \arrow[dr, dash, dashed, "\precsim"'] & \\
\varphi \arrow[ur, "\sim"'] & & \psi \arrow[ur, "\sim"'] & & \chi
\end{tikzcd} \tag{2}$$

The two ampliations occur in (possibly) different chains, one containing φ and $\varphi \wedge \psi$, the other ψ and $\psi \wedge \chi$. When the relations are numerically derived, one would expect that they can be chained. For instance, $\varphi \sim \varphi \wedge \psi$ and $\varphi \wedge \psi \precsim \psi$ would give φ "$\precsim$" ψ, and similarly ψ "$\precsim$" χ, hence φ "$\precsim$" χ.

Now consider what happens if also $\chi \mid\!\sim \varphi$ (as in the figure below).

$$
\begin{tikzcd}[column sep=small]
& \varphi\wedge\psi \arrow[dr, dash, "\precsim"] & & \psi\wedge\chi \arrow[dr, dash, dashed, "\precsim"] & & \chi\wedge\varphi \arrow[dr, dash, dashed, "\precsim"] & & \\
\varphi \arrow[ur, "\sim"'] & & \psi \arrow[ur, "\sim"'] & & \chi \arrow[ur, "\sim"'] & & \varphi & \dots
\end{tikzcd}
\qquad (3)
$$

In this situation, we conclude φ "$\precsim$" $\chi\wedge\varphi$. Since also $\chi\wedge\varphi$ "$\precsim$" φ (by $\chi\wedge\varphi \precsim \varphi$, i.e., $\chi\wedge\varphi \vdash \varphi$), we conclude φ "$\sim$" $\chi\wedge\varphi$. But φ and $\chi\wedge\varphi$ are in the same chain, so we could have $\varphi \not\sim \chi\wedge\varphi$, in disagreement with φ "$\sim$" $\chi\wedge\varphi$. What we need is that a sequence of $\sim$ and $\precsim$ statements as in the figure is *conservative* on ampliation chains. All this can be made precise using the following definition.

Definition 3. *For an ampliative inference relation $\mid\!\sim$ and a corresponding ampliation relation $\sim$, we define:*

$\precsim^*$ *is the transitive closure of* $\mid\!\sim$
$\varphi \sim^* \psi := \varphi \precsim^* \psi$ *and* $\psi \precsim^* \varphi$
$\varphi \prec^* \psi := \varphi \precsim^* \psi$ *and not* $\psi \precsim^* \varphi$
$\varphi \succsim^* \psi := \psi \precsim^* \varphi$
$\varphi \succ^* \psi := \psi \prec^* \varphi$

$\precsim^*$ extends $\precsim$ and $\sim^*$ extends $\sim$, but $\prec^*$ does not always extend $\prec$. (Cf. the counterexample to Conservativeness below.) By (R), $\precsim^*$ is reflexive, hence $\sim^*$ is also reflexive. As $\precsim^*$ is transitive, it is a preorder, and $\sim^*$ an equivalence relation. $\precsim^*$ is well-defined on $\sim^*$-equivalence classes:

Proposition 3. *Let $\mid\!\sim$ be a qualitative ampliative inference relation. Then the following holds:*

If $\varphi \precsim^* \psi$, $\varphi \sim^* \varphi'$ and $\psi \sim^* \psi'$, then $\varphi' \precsim^* \psi'$.

Conservativeness is expressed as follows. The $\mid\!\sim$ version is known as Loop.

(L) If $\varphi_0 \mid\!\sim \varphi_1$, $\varphi_1 \mid\!\sim \varphi_2$, ... and $\varphi_n \mid\!\sim \varphi_0$, then $\varphi_0 \mid\!\sim \varphi_n$.
(C) If $\varphi \precsim^* \psi$ and $\psi \vdash \varphi$, then $\varphi \sim \psi$.

Assuming qualitative ampliative inference, the $\mid\!\sim$ and $\sim$ versions are equivalent:

Proposition 4. *Assume an ampliative inference relation $\mid\!\sim$ and a corresponding ampliation relation $\sim$. Then $\mid\!\sim$ obeys (L) if and only if $\sim$ obeys (C).*

Here is an example of a consequence relation for qualitative ampliative arguments that does not obey (L)/(C):

$\varphi \mid\!\sim \psi$ if and only if ($p\wedge q \vdash \varphi\wedge\psi$ and $\varphi \vdash p$) or ($q\wedge r \vdash \varphi\wedge\psi$ and $\varphi \vdash q$) or ($p\wedge r \vdash \varphi\wedge\psi$ and $\varphi \vdash r$) or ($\varphi \vdash \psi$).

For this consequence relation, we have $p \mid\!\sim q \mid\!\sim r \mid\!\sim p$, while $p \mid\!\not\sim r$.

4 Quantifiable qualitative ampliative arguments

The next step towards a quantified interpretation of qualitative ampliative arguments requires the notion of a numeric order of magnitude of a proposition.

Definition 4. *For a relation $\mid\!\sim \subseteq L \times L$, we define the* order *of φ, notation $O(\varphi)$, as the maximal length n of a sequence $\varphi_0, \ldots, \varphi_n$ with $\varphi = \varphi_0$ and, for all $i \in \{0, \ldots, n-1\}$, $\varphi_i \succ^* \varphi_{i+1}$ (if such a finite maximal length exists). When every proposition has a finite order, we say that $\mid\!\sim$* has finite orders*. When also there is a maximum order, we say that $\mid\!\sim$* has bounded finite orders*.*

When the property (L)/(C) holds, there cannot be a $\prec^*$-loop. So, for a language with a finite number of elementary propositions, (L)/(C) implies finite orders.

Restricting to bounded finite orders, qualitative ampliative inference can be defined in terms of the order function.

Theorem 1. *Let $\mid\!\sim$ express qualitative ampliative inference with bounded finite orders. Then there is an integer-valued function $O : L \to \mathbb{N}$ such that*

$\varphi \sim \psi$ if and only if $O(\varphi) = O(\psi)$ and $\psi \vdash \varphi$
$\varphi \succ \psi$ if and only if $O(\varphi) > O(\psi)$ and $\psi \vdash \varphi$

for which the following properties hold:

1. *$O(\bot) \leq O(\varphi) \leq O(\top)$.*
2. *$O(\varphi) = 0$ if and only if $\varphi \mid\!\sim \bot$.*
3. *$O(\varphi) \geq max(O(\varphi \wedge \psi), O(\varphi \wedge \neg\psi))$.*
4. *If $\psi \vdash \varphi$, then $O(\varphi) \geq O(\psi)$.*
5. *$\varphi \mid\!\sim \psi$ if and only if $O(\varphi) = O(\varphi \wedge \psi)$.*
6. *$\varphi \not\mid\!\sim \psi$ if and only if $O(\varphi) > O(\varphi \wedge \psi)$.*

In the proof, $O(\varphi)$ is taken to be the order of φ. The properties in the theorem are already quite close to the properties of probability theory. But the order of a proposition behaves like an order of magnitude, reflected by the max-relation in part 3, in contrast with probability theory's addition (which also has equality instead of $\geq$).

We next show that (assuming a further finitizing restriction) the max-relation can be replaced by addition. The order function is replaced by an 'argument value' function, in such a way that the order function behaves like a kind of rounded logarithm of the value function. We need definitions of maximally specific conclusions and of minimally specific exceptions.

Definition 5. *A* maximal conclusion *(or* extension*) of a proposition φ is a proposition ψ with $\varphi \mid\!\sim \psi$ and $\psi \vdash \varphi$ that is maximally specific in the sense of $\vdash$, i.e., for all χ, if $\psi \nvdash \chi$ and $\chi \vdash \psi$, then $\varphi \not\mid\!\sim \chi$. When every conclusion of a proposition can be amplified to a maximal conclusion, we say that $\mid\!\sim$ has* finitely expressible maximal conclusions*.*

A minimal non-conclusion *of a proposition φ is a proposition ψ with $\varphi \not\mathrel{|\!\sim} \psi$ and $\psi \vdash \varphi$ that is minimally specific in the sense of $\vdash$, i.e., for all χ, if $\psi \vdash \chi$ and $\chi \nvdash \psi$, then $\varphi \mathrel{|\!\sim} \chi$. When every non-conclusion of a proposition ($\vdash$-implied by the non-conclusion) is an amplification of a minimal non-conclusion, we say that $\mathrel{|\!\sim}$ has* finitely expressible minimal non-conclusions*.*

A minimal exception *of a proposition φ is a minimal non-conclusion ψ of φ such that there does not exist a maximal conclusion χ of φ with $\psi \vdash \chi$.*

To simplify technicalities (involving infinite disjunctions and conjunctions corresponding to infinite unions and intersections of sets), we restrict ourselves to inference relations with finitely expressible maximal conclusions and minimal non-conclusions.

The following theorem contains the inductive definition of an integer-valued function $v : L \to \mathbb{N}$ that characterizes qualitative ampliative inference. The induction is based on the property that minimal non-conclusions of a proposition have an order that is strictly lower than the proposition.

Theorem 2. *Let $\mathrel{|\!\sim}$ express qualitative ampliative inference, have bounded finite orders, have finitely expressible maximal conclusions and minimal non-conclusions, have a bounded finite number of maximal conclusions per proposition, and a bounded finite number of minimal exceptions per proposition. Let C denote the minimal upper bound of the number of maximal conclusions. Let O be the order function on propositions and $c(\varphi)$ the number of maximal conclusions of a proposition φ. We inductively define the* argument value *function $v : L \to \mathbb{N}$:*

$$v(\varphi) := \begin{cases} 0 & \text{if } O(\varphi) = 0; \\ 1 & \text{if } O(\varphi) = 1; \\ c(\varphi).v(O(\varphi)) + \sum\{v(\psi) \mid \psi \text{ minimal exception of } \varphi\} & \text{if } O(\varphi) > 1, \end{cases}$$

where $v(n) = (C^2 + 1)^{n-1}$ for $n > 0$. Then:

1. *$v(\bot) \leq v(\varphi) \leq v(\top)$.*
2. *$v(\varphi) = 0$ if and only if $\varphi \mathrel{|\!\sim} \bot$.*
3. *$v(\varphi) \geq v(\varphi \wedge \psi) + v(\varphi \wedge \neg\psi)$.*
4. *If $\psi \vdash \varphi$, then $v(\varphi) \geq v(\psi)$.*
5. *$\varphi \mathrel{|\!\sim} \psi$ if and only if and $\varphi \mathrel{|\!\sim} \bot$ or $\frac{v(\varphi \wedge \psi)}{v(\varphi)} > \frac{1}{C+1}$.*
6. *$\varphi \not\mathrel{|\!\sim} \psi$ if and only if $\varphi \not\mathrel{|\!\sim} \bot$ and $\frac{v(\varphi \wedge \psi)}{v(\varphi)} \leq \frac{1}{C+1}$.*

Note the special role of $\bot$, related to 0 having a logarithm of minus infinity.

As promised, the integer value function v of the theorem behaves almost like a probability function (bearing the standard connection between propositions and sets in mind). There is still one telling difference: whereas the sum of the probability of disjoint sets is equal to the probability of their union (in accordance with Kolmogorov's standard axioms), the sum of the argument values of mutually inconsistent propositions may not be *equal* to their disjunction, but can also be lower. There is a natural interpretation for this technical difference. Whereas probability theory counts cases of which all properties are available, in our setting of ampliative argumentation, the argument values count cases in which there can be unknown properties. Logically speaking, probability counts worlds (complete interpretations), whereas ampliative argumentation counts states expressed by propositions (partial interpretations). The role of partiality is to be expected, as ampliation is a way of adding information to the partial information expressed by the premises (cf. the examples at the start of the paper). The amplified information will again be partial.

In light of the close analogy between probabilities and the argument values used in the theorem, we will write $v(\psi|\varphi) := v(\varphi \wedge \psi)/v(\varphi)$. The argument value version of Bayes' theorem is an immediate consequence: $v(\psi|\varphi) = v(\varphi|\psi)v(\psi)/v(\varphi)$.

Theorem 2 shows that the 'conditional argument value' $v(\psi|\varphi)$ can be interpreted as the strength of the argument. When the strength is above the threshold $1/(C+1)$, the argument is valid, when the strength is below (or equal to) the threshold it is invalid.

5 Structured arguments

Until now, by our focus on the inference relation $\mid\!\sim$, we have used the classical model of arguments as unstructured premise-conclusion pairs. As such, we have established a theory of the global validity of arguments. However, in contemporary formal theory of defeasible arguments (see [16] for an overview), arguments have additional structure since they are constructed using premises, rules, exceptions, defeaters, etc. A central problem addressed in this kind of work is how the local structure of an argument, and of the arguments attacking the argument, determines the global validity of the argument. Today, Dung's seminal abstract perspective [5] plays a key role in determining such global validity (or *argumentative warrant*). Dung proposed different kinds of 'semantics' — preferred, complete, grounded and stable —, each determining a different kind of argumentative warrant. Several other semantics have been proposed, e.g., the stage and semi-stable semantics, both in [17] (though the semi-stable with another name), and the ideal and cf2 semantics; cf. the overview [1].

When we restrict ourselves to qualitative ampliative inference, as in this paper, determining the global validity of structured arguments becomes relatively simple: an argument is globally valid if the case it makes is constructed using rules that are not excluded by an exception. This is possible because we do not assume that all sets of rules, exceptions, defeaters, etc. can occur; the input from which arguments are constructed is constrained by the properties of a global theory (expressed in the inference relation). As a result, arguments are only constructed from 'well-behaved' sets rules and exceptions.

Syntactically, structured arguments are sequences of premises (propositions in L), rules of the form $\varphi \Rightarrow \psi$, and exceptions of the form $\neg(\varphi \Rightarrow \psi)$. An exception $\neg(\varphi' \Rightarrow \psi)$ *excludes* a rule $\varphi \Rightarrow \psi$ when $\varphi' \vdash \varphi$. When $\varphi \mid\!\sim \psi$, $\varphi \Rightarrow \psi$

is an $\mid\!\sim$-rule; when $\varphi \mid\!\not\sim \psi$, $\neg(\varphi \Rightarrow \psi)$ is an $\mid\!\sim$-exception. We will assume a language L with a finite number of elementary propositions and a qualitative ampliative inference relation $\mid\!\sim$.

Definition 6. *Let $\mid\!\sim$ express qualitative ampliative inference. We inductively define valid arguments α from premises $P(\alpha) \in L$ making the case $C(\alpha) \in L$ as follows.*

1. *The empty argument [ ] is a valid argument from $\top$ making the case $\top$. It has no applicable rules and no applicable exceptions.*
2. *When an argument $\alpha = [\gamma_0, \ldots, \gamma_n]$ is valid, it can be extended to a valid argument α' by adding a premise φ provided that, for each applicable rule $\psi \Rightarrow \chi$ occurring in α such that $C(\alpha) \wedge \varphi \mid\!\not\sim \chi$, an $\mid\!\sim$-exception overruling the rule is also added (after the new premise). Each thus excluded rule is not applicable in α'. Also each rule that comes after an excluded rule is not applicable in α'. α' is an argument from $P(\alpha) \wedge \varphi$ making the case $P(\alpha) \wedge \varphi \wedge \psi_0 \wedge \ldots \wedge \psi_k$, where $\psi_0, \ldots, \psi_k$ are the consequents of the applicable rules of α'.*
3. *When an argument $\alpha = [\gamma_0, \ldots, \gamma_n]$ is valid, it can be extended to a valid argument α' by adding a $\mid\!\sim$-rule $\varphi \Rightarrow \psi$ provided that $C(\alpha) \vdash \varphi$ and $C(\alpha) \mid\!\sim \psi$. That rule is an additional applicable rule of α'. The extended argument's case is $C(\alpha) \wedge \psi$.*

Theorem 3. *Let $\mid\!\sim$ express qualitative ampliative inference. Then $\varphi \mid\!\sim \psi$ if and only if there is a valid argument, structured as in Definition 6, from φ making the case $\varphi \wedge \psi$.*

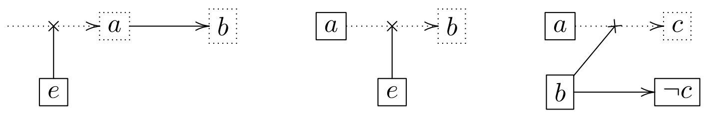
By this theorem, we can use the same term 'argument' for an argument in the sense of a global judgment of warrant and for an argument in the sense of a valid structured argument, as defined in Definition 6.

Assume for instance that the argument $[p, p \Rightarrow q, q \Rightarrow r, e, \neg(e \wedge q \Rightarrow r)]$ is valid, where all rules and exceptions are taken from a qualitative ampliative inference relation $\mid\!\sim$. This argument's construction consists of three steps. Initially there is only the premise p, then the rule $p \Rightarrow q$ is applied, then a second rule $q \Rightarrow r$, but that rule is immediately excluded by the exception e. The argument is an argument from $p \wedge e$ making the case $p \wedge q \wedge e$.

As said, the construction of valid arguments is more straightforward than in other proposals because the properties of qualitative ampliative inference restrict which rules and exceptions can form an argument. Consider for instance this argument, a slight variant of a puzzle studied by Pollock [14]: $[p, p \Rightarrow q, q \Rightarrow r, \neg(p \wedge r \Rightarrow q)]$. After the application of the second rule, the first one is excluded, making the application of the second rule impossible; contradiction. In the present proposal, this example is not a valid argument of a qualitative ampliative inference relation $\mid\!\sim$. For if the two rules can both be validly applied, we would have that $p \mid\!\sim q \wedge r$, hence by (CCM) and (RW) (and (LE)) $p \wedge r \mid\!\sim q$, so $\neg(p \wedge r \Rightarrow q)$ is not an $\mid\!\sim$-exception.

What about the expressiveness of the present approach? Isn't it too restrictive? Let's consider the three kinds of argument attack that are most prominent in state-of-the-art argumentation formalizations (e.g., [15]): *assumption-based attack*, aimed at the defeasible assumptions of an argument, *undercutting attack*, aimed at the connection between a reason and its conclusion, and *rebutting attack*, in which a reason for the opposite of the argument's conclusion is given.

Consider the following three arguments, one of each type:



(4)

The argument $[\Rightarrow a, a \Rightarrow b, e, \neg(e \Rightarrow a)]$, depicted on the left, in which the defeasible assumption a is attacked by e, is an example of assumption-based attack. The argument from a to b is based on the defeasible assumption a (formalized by the rule $\Rightarrow a$ with empty antecedent). The assumption is attacked by the exception e. When the argument is valid (with respect to a qualitative inference relation $|\!\sim$), the rules and exceptions in the argument correspond to $|\!\sim a, a \,|\!\sim b$ and $e \,|\!\not\sim a$. The gradual construction of the argument starts with the assumption. At this stage of construction, the argument's validity corresponds to $|\!\sim a$ (cf. Theorem 3). Then the rule $a \Rightarrow b$ is applied. The argument's validity now corresponds to $|\!\sim a \wedge b$. Finally, the premise e is added, that leads to the addition of an exception $\neg(e \Rightarrow a)$ excluding the assumption $\Rightarrow a$. The argument's validity now corresponds to $e \,|\!\not\sim a$ and $e \,|\!\sim e$.

Similarly for the middle, undercutting argument $[a, a \Rightarrow b, e, \neg(e \wedge a \Rightarrow b)]$, in which the reason a for b is undercut by e. Its rule and exception correspond to $a \,|\!\sim b$ and $a \wedge e \,|\!\not\sim b$. Its gradual construction corresponds to the sequence $a \,|\!\sim a, a \,|\!\sim a \wedge b, e \wedge a \,|\!\not\sim b, e \wedge a \,|\!\sim e \wedge a$.

The argument on the right, $[a, a \Rightarrow c, b, \neg(a \wedge b \Rightarrow c), b \Rightarrow \neg c]$, an example of a reason b for $\neg c$ that rebuts the reason a for c, has rules corresponding to $a \,|\!\sim c$ and $b \,|\!\sim \neg c$ and an exception corresponding to $a \wedge b \,|\!\not\sim c$. Its gradual construction corresponds to $a \,|\!\sim a, a \,|\!\sim a \wedge c, a \wedge b \,|\!\not\sim c, a \wedge b \,|\!\sim \neg c$.

6 Summary and concluding remarks

In this paper, the foundations of argumentation theory have been reconsidered using the formal properties of nonmonotonic consequence as a starting point [2, 11, 12]. Well-known properties of a nonmonotonic consequence relation $|\!\sim$ were reinterpreted in terms of a so-called ampliation relation $\sim$. By this reinterpretation, ampliative inference could be connected to properties studied in the mathematical theory of (pre)orders. This allowed the development of a quantitative interpretation of qualitative ampliative inference, in two versions.

The first quantitative interpretation was in terms of a numeric function measuring an 'order of magnitude' of a proposition (if you like: an 'order of normality'), with larger orders being more general (with the order of $\top$ as maximum) and lower orders more specific (with the order of $\perp$ as minimum). The properties of this order-of-magnitude function remind of possibility logic [4] by its use of a maximum property (as opposed to the additivity of probability), with a key difference being the use of an inequality. The order-of-magnitude function helps to explain why preferential logic (which can be obtained from our qualitative ampliative inference by adding the (And) and (Or) rules) can be defined both in terms of limits (Geffner & Pearl [6]) and in terms of minimal models (Kraus, Lehmann, Magidor [11]): a premise and its conclusions have the same order of magnitude, with a difference that vanishes 'in the limit'.

From the order function, a second quantitative interpretation of qualitative ampliative inference has been derived in terms of an 'argument value' function with properties that closely resemble the Kolmogorov probability axioms [7]. This second 'logico-probabilistic' interpretation of ampliative arguments uses numeric argument values that behave like conditional probabilities and that measure argument strength. When the strength of an argument is above a threshold, the argument is valid. The Kolmogorov additivity axiom is replaced by an inequality, that has the natural interpretation that the cases that are the basis of the numeric distribution (e.g., by counting observations or by otherwise weighting them) contain partial information, and not information about the 'whole world'. This partiality of information is a cornerstone of ampliative inference, that can by our formal proposal be interpreted as *jumping to conclusions* when staying sufficiently close in value to the premises.

One interpretation of the case made by an argument, defined here as the conjunction of premises and conclusions, is as an explanation supported by the argument's premises. As a result, the approach in this paper is related to theories of abductive inference to the best explanation [10]. This is no coincidence since one inspiration for the present approach was the intuition that argument-based and explanation-based approaches could be formally integrated (cf. Bex's hybrid argumentative-narrative theory of legal evidential reasoning [3]).

Our approach can be regarded as a bridge between logic and probability. While [8] consider the (And) rule as a watershed between qualitative and quantitative interpretations of nonmonotonic reasoning, we have proposed a quantitative interpretation of qualitative ampliative arguments that is independent of (And). Here, the (CCT) property, that expresses a kind of defeasible rule application and is probabilistically not valid, plays a key role in distinguishing quantitative and qualitative argumentation.

Bayesian networks [9, 13] also combine logic and probability. They are however often causally interpreted and need additional tools for the modeling of an agent's decision making, e.g., utilities. By the combination of the decision-oriented notions of defeasible argument, rules with exceptions and argument strength, the here proposed theory provides a fresh perspective on intelligent agents jumping to conclusions in order to interpret their world and act in it.

The distinguishing role of defeasible rule application and the associated property (CCT) shows on formal grounds that the rules and exceptions underlying qualitative ampliative inference cannot be derived from statistical correlations alone. Defeasible rules need to be tested by applying them, thereby generating hypotheses. Sometimes rules with their associated exceptions will lead to hypotheses that are not falsified too often. When this happens, such rules can be regarded as ‘knowledge’, in the sense that they describe accurate patterns.

Acknowledgements

The author would like to thank Allard Tamminga and Rineke Verbrugge for their valuable comments.

References

1. Baroni, P., Caminada, M., Giacomin, M.: Review: an introduction to argumentation semantics. The Knowledge Engineering Review 26(4), 365–410 (2011)
2. van Benthem, J.: Foundations of conditional logic. Journal of Philosophical Logic pp. 303–349 (1984)
3. Bex, F., Van Koppen, P., Prakken, H., Verheij, B.: A hybrid formal theory of arguments, stories and criminal evidence. Artificial Intelligence and Law pp. 1–30 (2010)
4. Dubois, D., Prade, H.: Possibility theory, probability theory and multiple-valued logics: A clarification. Annals of Mathematics and Artificial Intelligence 32(1), 35–66 (2001)
5. Dung, P.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence 77, 321–357 (1995)
6. Geffner, H., Pearl, J.: Conditional entailment: Bridging two approaches to default reasoning. Artificial Intelligence 53(2-3), 209–244 (1992)
7. Hájek, A.: Interpretations of probability. In: Zalta, E. (ed.) The Stanford Encyclopedia of Philosophy. Stanford University (2011)
8. Hawthorne, J., Makinson, D.: The quantitative/qualitative watershed for rules of uncertain inference. Studia Logica 86(2), 247–297 (2007)
9. Jensen, F., Nielsen, T.: Bayesian networks and decision graphs. Springer, Berlin (2007)
10. Josephson, J., Josephson, S.: Abductive Inference: Computation, Philosophy, Technology. Cambridge University Press, Cambridge (1996)
11. Kraus, S., Lehmann, D., Magidor, M.: Nonmonotonic reasoning, preferential models and cumulative logics. Artificial Intelligence 44, 167–207 (1990)
12. Makinson, D.: General patterns in nonmonotonic reasoning. In: Gabbay, D., Hogger, C., Robinson, J. (eds.) Handbook of Logic in Artificial Intelligence and Logic Programming. Volume 3. Nonmonotonic Reasoning and Uncertain Reasoning, pp. 35–110. Clarendon Press, Oxford (1994)
13. Pearl, J.: Causality. Cambridge University Press, Cambridge (2000)
14. Pollock, J.: Cognitive Carpentry: A Blueprint for How to Build a Person. The MIT Press, Cambridge (Massachusetts) (1995)
15. Prakken, H.: An abstract framework for argumentation with structured arguments. Argument and Computation 1(2), 93–124 (2010)
16. Rahwan, I., Simari, G. (eds.): Argumentation in Artificial Intelligence. Springer, Dordrecht (2009)
17. Verheij, B.: Two approaches to dialectical argumentation: admissible sets and argumentation stages. In: Meyer, J.J., van der Gaag, L. (eds.) Proceedings of NAIC’96, pp. 357–368. Universiteit Utrecht, Utrecht (1996)