This paper will appear in F. Zenker (Ed.), *Bayesian Argumentation.* Berlin, Heidelberg: Synthese Library (Springer).

On argument strength

Niki Pfeifer
(niki.pfeifer@lrz.uni-muenchen.de)
Munich Center for Mathematical Philosophy
Ludwig-Maximilians-Universität München
Geschwister-Scholl-Platz 1, 80539 Munich, Germany

Abstract

Everyday life reasoning and argumentation is defeasible and uncertain. I present a probability logic framework to rationally reconstruct everyday life reasoning and argumentation. Coherence in the sense of De Finetti is used as the basic rationality norm. I discuss two basic classes of approaches to construct measures of argument strength. The first class imposes a probabilistic relation between the premises and the conclusion. The second class imposes a deductive relation. I argue for the second class, as the first class is problematic if the arguments involve conditionals. I present a measure of argument strength that allows for dealing explicitly with uncertain conditionals in the premise set.

Probabilistic approaches to argumentation became popular in various fields including argumentation theory (e.g., Hahn & Oaksford, 2006), formal epistemology (e.g., Pfeifer, 2007, 2008), the psychology of reasoning (e.g., Hahn & Oaksford, 2007), and computer science (e.g., Haenni, 2009).

Probabilistic approaches allow for dealing with the uncertainty and defeasibility of everyday life arguments. This paper presents a procedure to formalize everyday life arguments in probability logical terms and to measure its strength.

"Argument" denotes an ordered triple consisting of (i) a (possibly empty) premise set, (ii) a conclusion indicator (usually denoted by "therefore" or "hence"), and (iii) a conclusion. As an example consider the following argument $\mathcal{A}$:

(1) If Tweety is a bird, then Tweety can fly.
(2) Tweety is a bird.
(3) Therefore, Tweety can fly.

In terms of the propositional calculus, $\mathcal{A}$ can be represented by $\mathcal{A}_1$:

(1) $B \supset F$
(2) B
(3) $\therefore F$

where "B" denotes "Tweety is a bird.", "F" denotes "Tweety can fly.", "$\therefore$" denotes the conclusion indicator, and "$\supset$" denotes the material conditional. The material conditional ($A \supset B$) is false if the antecedent (A) is true and the consequent (B) is false, and true otherwise.[1]

Argument $\mathcal{A}_1$ is an instance of the logically valid *modus ponens*. An argument is logically valid if, and only if, it is impossible that all premises

[1]Note that the propositional-logically atomic formulae B and F in argument $\mathcal{A}_1$ can be represented in predicate logic by $Bird(Tweety)$ and $Can\_Fly(Tweety)$, respectively. Moreover, F may be represented even more fine-grained in modal logical terms by $\diamond F$, where "$\diamond$" denotes a possibility operator. However, for the sake of sketching a theory of argument strength, it is sufficient to formalize atomic propositions by propositional variables.

are true and the conclusion is false. In everyday life, however, premises are often uncertain and conditionals allow for exceptions. Not all birds fly: penguins, for example, are birds that do not fly. Also the second premise may be uncertain: Tweety could be a non-flying bird or not even a bird. This uncertainty and defeasibility cannot be properly expressed in the language of the propositional calculus. Nevertheless, as long as there is no evidence that Tweety is a bird that cannot fly (e.g., that Tweety is a penguin), the conclusion of $\mathcal{A}$ is rational.

Probability logic allows for dealing with exception and uncertainty (e.g., Adams, 1975; Hailperin, 1996; Coletti & Scozzafava, 2002). It provides tools to reconstruct the rationality of reasoning and argumentation in the context of arguments like $\mathcal{A}_1$. Among the various approaches to probability logic, I advocate *coherence based probability logic* for formalizing everyday life arguments (Pfeifer & Kleiter, 2006a, 2009). Coherence based probability logic combines coherence based probability theory with propositional logic. It received strong empirical support in a series of experiments on the basic nonmonotonic reasoning System P (Pfeifer & Kleiter, 2003, 2005, 2006b), the paradoxes of the material conditional (Pfeifer & Kleiter, 2011), the conditional syllogisms (Pfeifer & Kleiter, 2007), and on how people interpret (Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011) and negate conditionals (Pfeifer, 2012).

Coherence based probability theory was originated by De Finetti (1980, 1974). Among others, it has been further developed by Walley (1991), Lad (1996), Biazzo and Gilio (2000), and Coletti and Scozzafava (2002). In the framework of coherence, probabilities are (subjective) *degrees of belief* and not objective quantities. It seems natural that different people may assign different degrees of belief to the premises of one and the same argument.

This does not mean, however, that everything is subjective and therefore no general rationality norms are available. *Coherence* requires to avoid bets that lead to sure loss, which in turn guarantees that the axioms of probability theory are satisfied.[2] Another characteristic feature of coherence is that conditional probability, $P(B \mid A)$, is a *primitive* notion. Consequently, the probability value is assigned *directly* to the conditional event, $B \mid A$, as a whole. This contrasts with the standard approaches to probability, where conditional probability ($P(B \mid A)$) is defined by the fraction of the joint and the marginal probability ($P(A \wedge B) \,/\, \Pr(A)$). The probability axioms are formulated for conditional probabilities and not for absolute probabilities (the latter is done in the standard approach to probability and is problematic if $P(A) = 0$). Coherence based probability logic tells us how to propagate the uncertainty of the premises to the conclusion. As an example consider a probability logical version of the above argument, $\mathcal{A}_2$:

(1) $P(F|B) = x$
(2) $P(B) = y$
(3) $\therefore\ xy \leq P(F) \leq xy + 1 - y$

where xy and $xy + 1 - y$ are the tightest coherent lower and upper probability bounds, respectively, of the conclusion. $\mathcal{A}_2$ is an instance of the probabilistic modus ponens (see, e.g., Pfeifer & Kleiter, 2006a). If premise (1) had been replaced by the probability of the material conditional, then the tightest coherent lower and upper probability bounds of the conclusion would have been different ones. However, paradoxes and experimental results suggest that uncertain conditionals should not be represented by the

[2] I argued elsewhere (Pfeifer, 2008) that violation of coherence is a necessary condition for an argument to be fallacious.

probability of the material conditional ($P(A \supset B)$), but rather by the conditional probability ($P(B|A)$; Pfeifer & Kleiter, 2010, 2011).

The consequence relation between the premises and the conclusion is deductive in the framework of coherence based probability logic. The probabilities of the premises are transmitted deductively to the conclusion. Depending on the logical and probabilistic structure of the argument, the best possible coherent probability bounds of the conclusion can be a *precise* (point) probability value or an imprecise (interval) probability. Interval probabilities are constrained by a lower and an upper probability bound (see the conclusion of $\mathcal{A}_2$). In the worst case, the unit interval is a coherent assessment of the probability of the conclusion. In this case the argument form is probabilistically non-informative: zero and one are the tightest coherent probability bounds (Pfeifer & Kleiter, 2009, 2006a).

The tightest coherent probability bounds of the conclusion provide useful building blocks for a measure of argument strength. Averages of the tightest coherent lower and upper probabilities of the conclusion given some threshold probabilities of the premises allow for measuring the strength of *argument forms* (like the modus ponens; see Pfeifer & Kleiter, 2006a). In the following I focus on measuring the strength of *concrete arguments* (like argument $\mathcal{A}$).

There are at least two alternative ways to construct measures of argument strength: one presupposes a *deductive* consequence relation, whereas the other one presupposes an *uncertain* consequence relation. As explained above, coherence based probability logic involves a deductive consequence relation. Theories of confirmation assume that there is an uncertain relation between the evidence and the hypothesis. "Theories of confirmation may be cast in the terminology of argument strength, because $P_1 \ldots P_n$ confirm C

$S_d(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{C}\|\mathcal{P}) - P(\mathcal{C})$	(Carnap, 1962)
$S_s(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{C}\|\mathcal{P}) - P(\mathcal{C}\|\neg\mathcal{P})$	(Christensen, 1999)
$S_m(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{P}\|\mathcal{C}) - P(\mathcal{P})$	(Mortimer, 1988)
$S_n(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{P}\|\mathcal{C}) - P(\mathcal{P}\|\neg\mathcal{C})$	(Nozick, 1981)
$S_c(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{P} \wedge \mathcal{C}) - P(\mathcal{P}) \times P(\mathcal{C})$	(Carnap, 1962)
$S_r(\mathcal{P},\mathcal{C})$	$=$	$\frac{P(\mathcal{C}\|\mathcal{P})}{P(\mathcal{C})} - 1$	(Finch, 1960)
$S_g(\mathcal{P},\mathcal{C})$	$=$	$1 - \frac{P(\neg\mathcal{C}\|\mathcal{P})}{P(\neg\mathcal{C})}$	(Rips, 2001)
$S_l(\mathcal{P},\mathcal{C})$	$=$	$\frac{P(\mathcal{P}\|\mathcal{C}) - P(\mathcal{P}\|\neg\mathcal{C})}{P(\mathcal{P}\|\mathcal{C}) + P(\mathcal{P}\|\neg\mathcal{C})}$	(Kemeny & Oppenheim, 1952)

Table 1: Measures of confirmation presented in the literature (adapted from Crupi et al., 2007).

only to the extend that $P_1 \dots P_n / C$ is a strong argument." (Osherson, Smith, Wilkie, López, & Shafir, 1990, p. 185). Table 1 casts a number of prominent measures of confirmation in terms of argument strength.

The underlying intuition of measures of confirmation is that premise set $\mathcal{P}$ *confirms* conclusion $\mathcal{C}$, if the conditional probability of the conclusion given the premises is higher than the absolute probability of the conclusion, $P(\mathcal{C}|\mathcal{P}) > P(\mathcal{C})$. $\mathcal{P}$ *disconfirms* $\mathcal{C}$, if $P(\mathcal{C}|\mathcal{P}) < P(\mathcal{C})$. If $\mathcal{C}$ is stochastically independent of $\mathcal{P}$, i.e. $P(\mathcal{C}|\mathcal{P}) = P(\mathcal{C})$, then the premises are *neutral* w.r.t. the confirmation of the conclusion. As pointed out by Fitelson (1999), these three conditions do not impose restrictions on the choice of the measures in Table 1, i.e., they are satisfied in the context of the listed measures.

Measures of confirmation may be appropriate for measuring the strength of arguments if we do not want to formalize explicitly the structure of the premise set. However, if the premise set includes conditionals (like argument $\mathcal{A}$), then these measures require a theory of how to combine conditionals and how to conditionalize on conditionals. Consider, for example argument $\mathcal{A}$ and the general requirement that a strong argument should satisfy the inequality $P(\mathcal{C}|\mathcal{P}) > P(\mathcal{C})$. It is easy to instantiate the conclusion of $\mathcal{A}$: $P(B|\mathcal{P}) > P(B)$. There are at least two options to instantiate the premise set $\mathcal{P}$. Both options depend on how the conditional in premise 1 is interpreted.

The first option consists in the interpretation of the conditional in terms of a conditional event, $B|A$. In this case at least two problems need to be solved. The first one is the combination of the conditional premise(s) with the other premise(s): "$(B|A) \,\&\, A$" is not defined.[3] The second problem concerns the conditionalization on conditionals: the meaning of "$P(B \mid (B|A)\ldots)$" needs to be explicated. This is a deep problem and an uncontroversial general theory is still missing (for a proposal of how to conditionalize on conditionals see, e.g., Douven, 2012).

The second option consists in the interpretation of the conditional in terms of the material conditional, $A \supset B$. Here, it is straightforward to combine the material conditionals and to conditionalize on the material conditional. Argument $\mathcal{A}$ is instantiated in the general requirement of strong arguments as follows: $P(B \mid A \wedge (A \supset B)) > P(B)$. However, coherence requires that $P(B \mid A \wedge (A \supset B)) = 1$. Thus, the inequality is trivially satisfied (if $P(\mathcal{C}) < 1$). It is counterintuitive that any instance—including those with low premise probabilities—of $\mathcal{A}$ are strong arguments. Therefore, measures of confirmation are not appropriate measures of argument strength if

[3]Since the conditional event is non-propositional, it cannot be combined by classical logical conjunction. Conditional events *can* be combined by so-called "quasi-conjunctions" (Adams, 1975, p. 46 f). As Adams notes, however, quasi-conjunctions lack some important logical features of conjunctions.

we want to explicitly formalize arguments that include conditionals.

I will now turn to a measure of argument strength and show how it allows for formalizing arguments that involve conditionals. The crucial idea is that (i) the precision of a strong argument is high and that (ii) the location of the coherent probability (interval) is close to 1 (Pfeifer, 2007). The imprecision is measured by the size of the tightest coherent probability bounds of the conclusion. Let z' and z'' denote the tightest coherent lower and upper bounds, respectively, of an argument $\mathcal{A}_x$. The imprecision of $\mathcal{A}_x$ is measured by $z'' - z'$. Consequently, the *precision* of $\mathcal{A}_x$ is measured by $1 - (z'' - z')$. The location of the coherent conclusion probability is measured by the arithmetic mean of the tightest coherent probability bounds, $\frac{z' + z''}{2}$. The argument strength s of $\mathcal{A}_x$ is equal to the product of the precision and the location of the tightest coherent probability bounds of the conclusion:

$$s(\mathcal{A}_x) \;=\; [1 - (z'' - z')] \times \frac{z' + z''}{2},$$

where $0 \leq s(\mathcal{A}_x) \leq 1$, since $0 \leq z' \leq z'' \leq 1$. The values 0 and 1 denote the weakest and the strongest value, respectively.

As an example of the evaluation procedure of the strength of an argument, consider the following instance of argument $\mathcal{A}_2$:

(1) $P(F|B) = .8$

(2) $P(B) = .9$

(3) $\therefore\; .72 \leq P(F) \leq .82$

The strength of this argument is .69. In the special case where the premises are certain (i.e., probabilities equal to 1) the strength of the argument obtains its maximum value 1.

Figure 1 presents the behavior of the measure in general. According to the measure, the argument strength increases if the location of the tightest

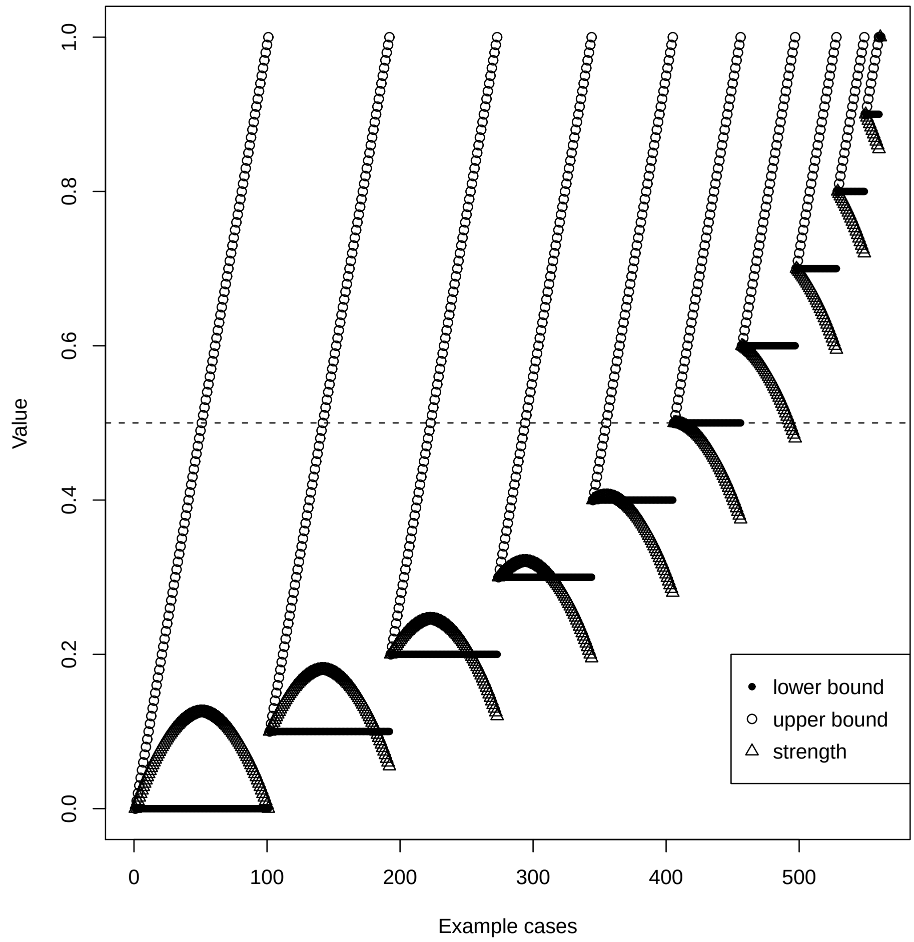

Figure 1: Let z' denote the tightest coherent lower and z'' denote the tightest coherent upper bound of an argument $\mathcal{A}$. The argument strength of $\mathcal{A}$ is equal to $[1-(z''-z')] \times \dfrac{z'+z''}{2}$. The strength of $\mathcal{A}$ increases if the precision of the conclusion is high and the location of the tightest coherent probability interval is close to 1.

coherent bounds of the conclusion approaches 1. The argument strength decreases if the imprecision increases. Moreover, an argument is weak if the conclusion probability is low. Maximum imprecision implies minimum argument strength. It follows that all probabilistically non-informative arguments are also weak arguments (with $s = 0$). Figure 2 shows the behavior of the measure for coherent lower conclusion probabilities of at least .5. If the conclusion probability is at least .5, then the argument strength varies between .375 and .500. The higher the precision the higher the strength of the argument.

The proposed measure contrasts with the traditional measures of confirmation presented in Table 1. The consequence relation remains deductive, while measures of confirmation assume an uncertain relation between the premises and the conclusion. Using probability logic to formalize arguments is advantageous as it does justice to the logical structure: premise sets that include conditionals can be represented explicitly. If a measure of argument strength requires to calculate the conditional probability of the conclusion given some combination of the premises, $P(\text{conclusion} \mid \text{premise set})$, then severe problems arise of how to connect premises containing conditionals with each other and how to conditionalize on conditionals. In the proposed measure this problem is avoided, as probability logic tells us how to infer the tightest coherent probability bounds of the conclusion from the premises, which are in turn exploited for calculating the argument strength.

The proposed measure s has not only attractive theoretical consequences (as explained above), it also implies at least two psychologically plausible hypothesis. People judge arguments as strong, if the premises imply high conclusion probabilities (i) and if the conclusion probability

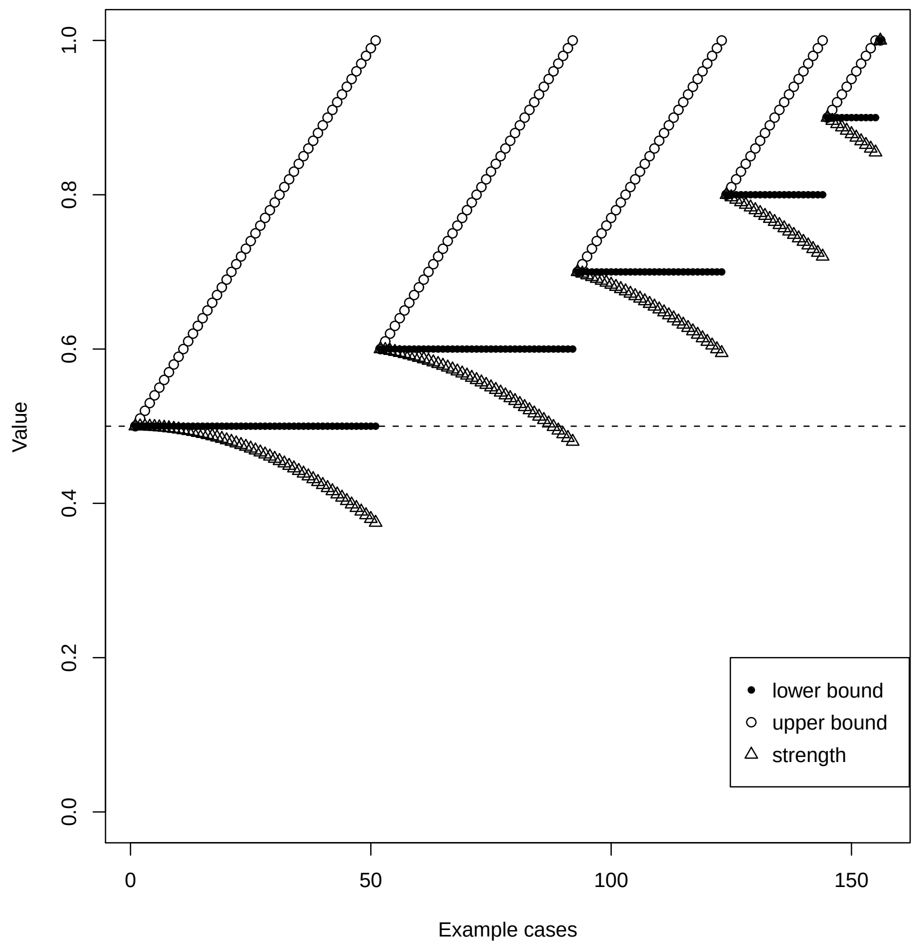

Figure 2: Detail of Figure 1, showing the behavior of measure s for coherent lower conclusion probabilities of at least .5.

is—at the same time—precise (ii). The empirical test of these hypothesis is a challenge for future research.

Acknowledgments

This work is financially supported by the *Alexander von Humboldt Foundation*, the *German Research Foundation* project PF 740/2-1 "Rational reasoning with conditionals and probabilities. Logical foundations and empirical evaluation" (Project leader: Niki Pfeifer; Project within the DFG Priority Program SPP 1516 "New Frameworks of Rationality"), and the *Austrian Science Fund* project P20209 "Mental probability logic" (Project leader: Niki Pfeifer).

References

Adams, E. W. (1975). *The logic of conditionals*. Dordrecht: Reidel.

Biazzo, V., & Gilio, A. (2000). A generalization of the fundamental theorem of De Finetti for imprecise conditional probability assessments. *International Journal of Approximate Reasoning*, *24*(2-3), 251-272.

Carnap, R. (1962). *Logical foundations of probability* (2nd ed.). Chicago: University of Chicago Press.

Christensen, D. (1999). Measuring confirmation. *Journal of Philosophy*, *96*, 437–461.

Coletti, G., & Scozzafava, R. (2002). *Probabilistic logic in a coherent setting*. Dordrecht: Kluwer.

Crupi, V., Tentori, K., & Gonzales, M. (2007). On Bayesian measures of confirmation. *Philosophy of Science*, *74*, 229-252.

De Finetti, B. (1974). *Theory of probability* (Vols. 1, 2). Chichester: John Wiley & Sons. (Original work published 1970)

De Finetti, B. (1980). Foresight: Its logical laws, its subjective sources (1937). In H. J. Kyburg & H. E. Smokler (Eds.), *Studies in subjective probability* (p. 55-118). Huntington, New York: Robert E. Krieger Publishing Company.

Douven, I. (2012). Learning conditional information. *Mind & Language*, *27*(3), 239–263.

Finch, H. A. (1960). Confirming power of observations metricized for decisions among hypotheses. *Philosophy of Science*, *27*, 293–207 (part I), 391–404 (part II).

Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science*, *66*, 362–378.

Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts towards the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 635–648.

Haenni, R. (2009). Probabilistic argumentation. *Journal of Applied Logic*, 155–176.

Hahn, U., & Oaksford, M. (2006). A normative theory of argument strength. *Informal Logic*, *26*, 1-22.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, *114*(3), 704-732.

Hailperin, T. (1996). *Sentential probability logic. Origins, development, current status, and technical applications*. Bethlehem: Lehigh University Press.

Kemeny, J., & Oppenheim, P. (1952). Degrees of factual support. *Philosophy* *of Science*, *19*, 307–324.

Lad, F. (1996). *Operational subjective statistical methods: A mathematical, philosophical, and historical introduction*. New York: Wiley.

Mortimer, H. (1988). *The logic of induction*. Paramus, NJ: Prentice Hall.

Nozick, R. (1981). *Philosophical explanations*. Oxford: Clarendon.

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*(2), 185–200.

Pfeifer, N. (2007). Rational argumentation under uncertainty. In G. Kreuzbauer, N. Gratzl, & E. Hiebl (Eds.), *Persuasion und Wissenschaft: Aktuelle Fragestellungen von Rhetorik und Argumentationstheorie* (p. 181-191). Wien: LIT.

Pfeifer, N. (2008). A probability logical interpretation of fallacies. In G. Kreuzbauer, N. Gratzl, & E. Hiebl (Eds.), *Rhetorische Wissenschaft: Rede und Argumentation in Theorie und Praxis* (pp. 225–244). Wien: LIT.

Pfeifer, N. (2012). Experiments on Aristotle's Thesis: Towards an experimental philosophy of conditionals. *The Monist*, *95*(2), 223–240.

Pfeifer, N., & Kleiter, G. D. (2003). Nonmonotonicity and human probabilistic reasoning. In *Proceedings of the 6th workshop on uncertainty processing* (p. 221-234). Hejnice: September 24–27$^{\text{th}}$, 2003.

Pfeifer, N., & Kleiter, G. D. (2005). Coherence and nonmonotonicity in human reasoning. *Synthese*, *146*(1-2), 93-109.

Pfeifer, N., & Kleiter, G. D. (2006a). Inference in conditional probability logic. *Kybernetika*, *42*, 391-404.

Pfeifer, N., & Kleiter, G. D. (2006b). Is human reasoning about nonmonotonic conditionals probabilistically coherent? In *Proceedings of the 7th workshop on uncertainty processing* (p. 138-150). Mikulov: September 16–20$^{\text{th}}$, 2006.

Pfeifer, N., & Kleiter, G. D. (2007). Human reasoning with imprecise probabilities: Modus ponens and Denying the antecedent. In G. De Cooman, J. Vejnarová, & M. Zaffalon (Eds.), *5th International Symposium on Imprecise Probability: Theories and Applications* (p. 347-356). Prague, Czech Republic: SIPTA.

Pfeifer, N., & Kleiter, G. D. (2009). Framing human inference by coherence based probability logic. *Journal of Applied Logic*, *7*(2), 206–217.

Pfeifer, N., & Kleiter, G. D. (2010). The conditional in mental probability logic. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thought* (p. 153-173). Oxford: Oxford University Press.

Pfeifer, N., & Kleiter, G. D. (2011). Uncertain deductive reasoning. In K. Manktelow, D. E. Over, & S. Elqayam (Eds.), *The science of reason: A Festschrift for Jonathan St B.T. Evans* (p. 145-166). Hove: Psychology Press.

Rips, L. J. (2001). Two kinds of reasoning. *Psychological Science*, *12*(2), 129-134.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. London: Chapman and Hall.