Argument Strength via Probabilistic Interpretations of Logical Support and Attack Principles*

Niki Pfeifer[1] and Christian G. Fermüller[2]

[1] Department of Philosophy, University of Regensburg, Germany, niki.pfeifer@ur.de
[2] Institute of Logic and Computation, TU Wien, Austria, chrisf@logic.at

There are many approaches to argument strength studying attack or support relations, combining logic and probability (e.g., [2, 4, 8, 9, 11, 18]). We propose to investigate argument strength contextually: the strength of an argument is evaluated with respect to the strength of attacks and supports, respectively, of logically related argumentative claims. We first define quantitative attack and support principles pertaining to the logical form of claims. Then, we generalise such principles quantitatively by attaching weights to the support and attack relations. Finally, we evaluate these principles within coherence-based probability logic and discuss some empirical data to assess the psychological plausibility of the proposed approach.

We have studied systematic relations between logical form and attacks between claims in an argumentative framework [15]. Usually, arguments are conceived as premise ("support") and conclusion ("claim") pairs. In what follows we write "$A \twoheadrightarrow B$" ("$A \Longrightarrow B$", resp.) to denote that there is an argument claiming A that attacks (supports, resp.) an argument with claim B, i.e., we abstract from the premises and focus on claims. In [5] qualitative rationality principles for attack principles were presented, which constrain attacks according to the logical form of the claims. For example,

(1) If $F \twoheadrightarrow A$ or $F \twoheadrightarrow B$, then $F \twoheadrightarrow A \wedge B$.

In [15], we interpreted $F \twoheadrightarrow G$ by the probability $p(\neg G|F) \geq t$ for some threshold $.5 < t \leq 1$, where F is not a logical contradiction ($\perp$). Thus, (1) corresponds to:

$$\text{if } p(\neg A|F) \geq t \text{ or } p(\neg B|F) \geq t, \text{ then } p(\neg(A \wedge B)|F) \geq t.$$

Principles like (1) were generalised to quantitative (weighted) versions in [6]. The quantitative version of (1), for example, was formulated as

(1') If $F \overset{x}{\twoheadrightarrow} A$, $F \overset{y}{\twoheadrightarrow} B$, and $F \overset{z}{\twoheadrightarrow} A \wedge B$, then $z \geq \max(x, y)$.

In [15] we interpreted (1') probabilistically as follows:

$$p(\neg(A \wedge B)|F) \geq \max(p(\neg A|F), p(\neg B|F))\,,$$

* Niki Pfeifer is supported by the BMBF project 01UL1906X.

which is coherent, and hence justifies (1'). In this manner, we used coherence-based probability logic evaluate the plausibility of a selection of such principles for conjunction, disjunction, implication, and negation.

In our contribution, we introduce corresponding support principles, like

(2) If $F \Longrightarrow A$ and $F \Longrightarrow B$, then $F \Longrightarrow A \wedge B$.
(2') If $F \overset{x}{\Longrightarrow} A$, $F \overset{y}{\Longrightarrow} B$, and $F \overset{z}{\Longrightarrow} A \wedge B$, then $z \leq \min(x, y)$.

Under the probabilistic interpretation, (2) corresponds to:

$$p(A|F) \geq t \text{ and } p(B|F) \geq t \text{ implies } p(A \wedge B|F) \geq 2t - 1,$$

for some threshold $.5 < t \leq 1$. (2') corresponds to:

$$p(A \wedge B|F) \leq \min(p(A|F), p(B|F))\,.$$

Moreover, we obtain

$$(A \wedge B|F) \geq \max(0, p(A|F) + p(B|F) - 1)\,.$$

Thus, coherence-based probability logic justifies both (2) and (2'). As for the attack principles, we will now systematically study support principles involving the usual logical connectives. Moreover, we will investigate principles which combine support and attack relations, like

(3) If $F \Longrightarrow A$ and $F \twoheadrightarrow B$, then $F \twoheadrightarrow (A \supset B)$.
(3') If $F \overset{x}{\Longrightarrow} A$, $F \overset{y}{\twoheadrightarrow} B$, and $F \overset{z}{\twoheadrightarrow} (A \supset B)$, then $z \leq \min(x, y)$.

$p(A|F) \geq t$ and $p(\neg B|F) \geq t$ implies $p(\neg(A \supset B)|F) \geq 2t - 1$. Hence (3) is justified only if $t = 1$. (3') corresponds to:

$$p(\neg(A \supset B)|F) \leq \min(p(A|F), p(\neg B|F))\,,$$

which holds in general, and thus (3') is justified.

To contextualize our approach, note that Dung-style argumentation theory can be seen as referring to two quite different levels. On an abstract level, following [7], arguments are represented simply as nodes in a directed graph and edges between nodes represent attacks between arguments. On a concrete (instantiated) level, arguments are structured compounds of specific logically complex statements and, possibly, rules of different kinds (see, e.g., [1, 3, 10]). The logical attack principles, introduced in [5], that we study here neither operate on the level of abstract argumentation frameworks nor on the level of concrete fully instantiated, complex arguments. Our attack and support principles rather focus on the *logical form* of claims, i.e. on the outermost logical connective of the formula representing the claim of an argument. As mentioned above, we consider principles of the following kind: If an argument X attacks an argument that features a claim A, then X also attacks arguments with claim $A \wedge B$. Formally, we follow [5], and thus consider *semi-abstract argumentation frameworks*, which are just ordinary argumentation frameworks, where each node is annotated with a propositional formula featuring the claim of the represented argument. The expression “semi-abstract” is meant to signal that we are not interested in the possibly quite complex internal structure of concrete arguments; rather, we add information about the logical form of claims to the abstract argumentation frameworks.

Finally, we discuss experimental-psychological data to assess the descriptive validity of our approach. Previous data on direct tests of coherence-based probability logic suggests a high descriptive validity: most people infer coherent intervals in diverse task settings (e.g., [12–14, 16, 17]). However data on quantitative logical attack principles, where the tasks consisted in inferring strengths of attacks, suggest modest agreement between the predictions and the data [15]. Since support, in contrast to attack, is formulated positively, we hypothesise to obtain a higher agreement between predictions of our quantitative support principles and the data, compared to the corresponding attack tasks. We are currently developing an experiment on support principles as well as principles which combine support and attack relations and plan to report first experimental results in our talk.

References

1. Arieli, O., Straßer, C.: Sequent-based logical argumentation. Argument & Computation **6**(1), 73–99 (2015)
2. Baroni, P., Gabbay, D.M., Giacomin, M., van der Torre, L.: Handbook of formal argumentation. College Publications (2018)
3. Besnard, P., Hunter, A.: Elements of argumentation. MIT Press Cambridge (2008)
4. Cohen, A., Gottifredi, S., García, A.J., Simari, G.R.: A survey of different approaches to support in argumentation systems. The Knowledge Engineering Review **29**(5), 513 (2014)
5. Corsi, E.A., Fermüller, C.G.: Logical argumentation principles, sequents, and nondeterministic matrices. In: Baltag, A., Seligman, J., Yamada, T. (eds.) LORI 2017. LNCS, vol. 10455, pp. 422–437. Springer, Berlin (2017)
6. Corsi, E.A., Fermüller, C.G.: Connecting fuzzy logic and argumentation frames via logical attack principles. Soft Computing **23**, 2255–2270 (2019)
7. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artif. Intelligence **77**(9), 321–357 (1995)
8. Haenni, R.: Probabilistic argumentation. Journal of Applied Logic pp. 155–176 (2009)
9. Hunter, A.: A probabilistic approach to modelling uncertain logical arguments. International Journal of Approximate Reasoning **54**(1), 47–81 (2013)
10. Modgil, S., Prakken, H.: The ASPIC+ framework for structured argumentation: a tutorial. Argument & Computation **5**(1), 31–62 (2014)
11. Parsons, S.: Normative argumentation and qualitative probability. In: Gabbay, D.M., Kruse, R., Nonnengart, A., Ohlbach, H.J. (eds.) Qualitative and Quantitative Practical Reasoning. pp. 466–480. Springer, Berlin (1997)
12. Pfeifer, N.: Experiments on Aristotle’s Thesis: Towards an experimental philosophy of conditionals. The Monist **95**(2), 223–240 (2012)

13. Pfeifer, N.: Reasoning about uncertain conditionals. Studia Logica **102**(4), 849–866 (2014)
14. Pfeifer, N.: Probability logic. In: Knauff, M., Spohn, W. (eds.) Handbook of Rationality. The MIT Press, Cambridge, MA (in press)
15. Pfeifer, N., Fermüller, C.G.: Probabilistic interpretations of argumentative attacks: logical and experimental foundations. In: 11$^{\text{th}}$ Workshop on Uncertainty Processing (WUPES'18). pp. 141–152. MatfyzPress Publishing House, Prague (2018)
16. Pfeifer, N., Kleiter, G.D.: Framing human inference by coherence based probability logic. Journal of Applied Logic **7**(2), 206–217 (2009)
17. Pfeifer, N., Tulkki, L.: Conditionals, counterfactuals, and rational reasoning. An experimental study on basic principles. Minds and Machines **27**(1), 119–165 (2017)
18. Prakken, H., Horty, J.: An appreciation of John Pollock's work on the computational study of argument. Argument and Computation **3**(1), 1–19 (2012)