A formal analysis of the notion of preference between deductive arguments

A recent branch of studies within the field of epistemic logic has been the analysis of justifications that an epistemic agent possesses to support one belief or another regarding a specific topic [1, 3, 2, 4, 5]. Some of the epistemic effects caused by dynamic processes involving justifications, such as evidence elimination caused by communications, have also been studied [11]. Nevertheless, there is something that has not received enough attention yet: how epistemic agents might prefer certain justifications to others, in order to have better pieces of evidence to support a particular belief. It seems clear that the manner we prefer one justification to another for our beliefs depends on a variety of factors. In an ongoing work, we have developed a formal analysis of the preference between a particular kind of justifications: *deductive arguments*.

More concretely, let a be an epistemic agent, let φ be a proposition and let t,s be a couple of (possibly different) deductive arguments: *(i)* Which of the arguments (t or s) is preferred by a if she intends to support φ? The motivation for this question and the applications of our answer are diverse and belong to different fields. For example, knowing which argument is better for a given agent would imply knowing which argument will be selected by her in the context of an argumentative dialogue. It would also offer an approach to face the problem of how an agent establishes her beliefs solidly enough to convert them into knowledge. Regarding formal argumentation theory, an appropriate answer to *(i)* represents a formal theory of argument evaluation.

According to our approach, the preference between deductive arguments is reducible to other notions. In particular, it can be analysed attending to the following list: the syntactic shape of the arguments, the epistemic attitudes of the agent with respect to the arguments, the sources from which arguments are taken from or construed, and lastly the quantity and complexity of information contained in both arguments. We have built a system that is able to capture this analysis using tools from epistemic logic, justification logic and logics for belief dependence; and that allows us to reason systematically about particular problems that may arise in this context. The mentioned system has some natural connections with some published research in abstract argumentation frameworks (see for example [9, 10]).

Before sketching the analysis we should make clear which kind of situations we are intending to model. First, we will focus on cases where the goal of argument evaluation is rational inquiry and where only some initial beliefs, but no knowledge, are available. That is to say, cases where an agent assesses a couple of arguments in order to decide which of them is better for pursuing knowledge and truth but where no initial knowledge is accessible to her. This assumption implies at least two consequences. On the one hand, we suppose that the beliefs of the agent are the result of some previous epistemic processes, for instance, observations; therefore pure irrational beliefs are excluded. On the other hand, the sources from which she builds up her arguments have been previously contrasted somehow. Second, no rhetorical features are taken into account; the agent will evaluate her arguments regardless of the intention of convincing anyone else. However, we think that it may be possible to capture some rhetorical features using the same logical apparatus that will be used in our analysis. Third, in the range of cases we are studying no time bounds are considered; agents will take as much time as they need –but always a finite amount– for their evaluations. Fourth, the notion of preference that we are modelling is claim-related. In other words, we are seeking the best argument for a given

proposition. This excludes direct applications of the current analysis to the problem of deciding which proposition of a given set is better supported from the point of view of the agent. Finally, the resulting theory has a clearly normative character, we do not talk about what epistemic agents actually do when they evaluate their deductive arguments, but rather what they should do in order to have the best epistemic standards.

The intuitive picture behind our analysis is that, when comparing two arguments t and s for supporting a given proposition φ, an epistemic agent makes them go through a test consisting of several filters. As the result of this process, the agent will weakly prefer t to s (prefer t as much as s) if t has passed at least as many filter as s. The remaining preference concepts (strict preference, indifference and incomparability) are defined as usual in the preference logic literature [6, 7].

The first filter is purely syntactic and we call it *syntactic accuracy*. If a tries to justify φ using argument t, then it seems rather obvious that the first thing she should require of t is to have φ as its conclusion. In this sense, no semantic notion (true, validity or any other related) is taken into account at the beggining of the argument evaluation, since a is just checking whether t and s have the proper syntactic shape. This filter is implemented in our system using the formal notion of admissibility, taken from [4, 5].

In the second filter, we focus on the *epistemic attitudes* of the agent towards the premises of the argument. It seems clear that a will prefer arguments whose premises are believed by her to arguments whose not. For the time being, we have just studied the case where only the beliefs of the agent are considered. Although extending this study with (various types of) knowledge, for example using plausibility models, looks like a promising idea for future developments. Be as it may, if this filter is passed too, what else could she do for deciding which argument is better? The answer is provided by the third filter.

In the third filter the agent turns her attention to her *best advisor* on the topic. Once her beliefs have been found to be insufficient to decide which argument is better, she seeks help in the other agents of the system. Notice that *the best advisor* is understood here in a broad sense (it could be a human being, a computer or an encyclopedia), i.e., the best advisor means the best source for the agent regarding the topic. This idea is taken from [8]; but some changes are made in order to fulfil our theoretical intentions. Unfortunately for our agent, it might be the case that her best advisor is not sufficiently informed to tell her what to do; since she (the advisor) could also consider t and s equally good. However, we could still establish some criteria to make her decide.

Finally, in the fourth and fifth filter; the agent checks the *simplicity* of the arguments (and she prefers, of course, the simplest). The difference between these two requirements is that the former is related to the to the manner in which the information is structured in the arguments the latter is related to the quantity of information contained in the arguments. These two filters could be understood as a new version of Ockham's razor. It is not straightforward which of the requirements should be considered more important, since we could think of counter-examples for the two hierarchies. However, we decided to take the complexity of information as more important than the quantity of information and we offer several reasons for doing so.

References

[1] Sergei Artemov. The logic of justification. *The Review of Symbolic Logic*, 1(4):477–513, 2008.

[2] Sergei Artemov and Melvin Fitting. Justification logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016.

[3] Sergei Artemov and Elena Nogina. Introducing justification into epistemic logic. *Journal of Logic and Computation*, 15(6):1059–1073, 2005.

[4] Alexandru Baltag, Bryan Renne, and Sonja Smets. The logic of justified belief change, soft evidence and defeasible knowledge. *Logic, Language, Information and Computation*, pages 168–190, 2012.

[5] Alexandru Baltag, Bryan Renne, and Sonja Smets. The logic of justified belief, explicit knowledge, and conclusive evidence. *Annals of Pure and Applied Logic*, 165(1):49–81, 2014.

[6] Sven Ove Hansson. Preference logic. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of philosophical logic*, pages 319–393. Springer, 2001.

[7] Sven Ove Hansson and Till Grüne-Yanoff. Preferences. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2017 edition, 2017.

[8] Zhisheng Huang. Logics for belief dependence. In *International Workshop on Computer Science Logic*, pages 274–288. Springer, 1990.

[9] Souhila Kaci and Leendert van der Torre. Preference-based argumentation: Arguments supporting multiple values. *International Journal of Approximate Reasoning*, 48(3):730–751, 2008.

[10] Sanjay Modgil and Henry Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397, 2013.

[11] Bryan Renne. Multi-agent justification logic: Communication and evidence elimination. *Synthese*, 185:43–82, 2012.