Commentary on Licato & Cooper's Evaluating Relevance in Analogical Arguments Through Warrant-based Reasoning

NICOLE CRUZ
University of New South Wales
n.cruz@unsw.edu.au

1. INTRODUCTION

This paper introduces a computer programme that sets procedural rules for analogical reasoning between two people. The rules are aimed at improving the quality of analogical argumentation by trying to ensure the relevance of argumentative moves. This is done by making explicit the features that connect the components of an analogy. In the programme, one person tries to defend an analogy while the other tries to undermine it by pointing to counterexamples. The defender can then revise the analogy to strengthen it against counterarguments in an iterative process. This process is overseen by a moderator who can resolve disputes about the acceptability of argumentative steps.

The authors state that their longer term goals in developing the programme are for it to be implementable computationally with currently available tools, for it to be usable by minimal trained reasoners with a minimal amount of external moderation, and for it to be used not only by people but also by artificially intelligent systems. At this point, the programme, called WG-A (for "Warrant Game – Analogy") has been tested with university students receiving less than one hour of training.

2. THEORETICAL BACKGROUND

The authors follow Bartha's (2010) characterisation of an analogy as a one-to-one correspondence between a source and a target domain, on the basis of which the reasoner concludes that a statement that holds in the source domain also holds in the target domain. For an analogy to be good in this context, there must be a clear connection between the statement in the source domain and the corresponding statement in the target domain, and there must not be a critical difference between these two statements.

The connection between the source and target domain, or more generally between the premises and conclusion of an argument, is referred to as the arguments' *warrant* (Toulmin, 2003). Depending on the type of argument, the warrant could consist in e.g.

logical/probabilistic structure, causal relations, or similarity based relations.

The authors argue that making the warrants of an argument explicit can be difficult, but is very useful because it helps determine features of argument strength such as what kind of attacks can be used against the argument, whether the premises logically imply the conclusion, or whether the conclusion is more plausible given the premises than when standing on its own (e.g. Oaksford & Chater, 2013; Oaksford & Hahn, 2007; Stern & Hartmann, 2018; Over & Cruz, 2019). They call reasoning that focuses explicitly on such warrants "warrant based reasoning".

3. IMPLEMENTATIONAL ASSUMPTIONS

When implementing analogical reasoning in the WG-A software, the authors make a number of decisions that differ from the theoretical approaches described above. Bartha (2010) characterises analogies as referring to an association that is generalised to a new domain. But in WG-A the association that people reason about is already generalised, and the task is to assess whether this generalisation is justified. Further, in Bartha the association to be generalised can be a predictive relation from *p* to *q*, a diagnostic or explanatory relation from *q* to *p*, a bidirectional relation between *p* and *q*, or an unspecified, correlational relation between *p* and *q*. In contrast, in WG-A the association is always predictive.

Moreover, the authors phrase the predictive relation from *p* to *q* as a conditional, *if p then q*, and reasoners are instructed to treat this conditional "almost as if it were" the material conditional of classical logic. The material conditional is false when *p* is true but *q* is false, and it is true in all other cases (i.e. when *p* and *q* are both true, and whenever *p* is false). This makes the material conditional *if p then q* equivalent to the disjunction *not-p or q*. The material conditional has come under strong critique as a representation of natural language conditionals in the last decades, first in philosophical logic and then in the psychology of reasoning, and there is very strong, consistent empirical evidence against it within psychology (for overviews see Edgington, 1995; Evans & Over, 2004; Oaksford & Chater, 2007). As an example of how material conditionals can lead us astray when applied to actual reasoning, consider the conditional "If I take up smoking, my health will improve." I might consider this conditional implausible and therefore not take up smoking. But if it is a material conditional, then as I become more and more determined not to take up smoking, it will become more and more likely that if I did take up smoking, my health will improve, for the material conditional is true whenever its antecedent is false.

It is not clear to me why the authors adopt the material conditional in their work, when there are a number of alternatives with arguably more realistic properties when applied to reasoning and argumentation. Examples of alternatives currently advocated in psychology, philosophical logic and linguistics are the probabilistic conditional (de Finetti, 1936/1995; Edgington, 1995; Coletti & Scozzafava, 2002; Jeffrey, 1991; Ramsey, 1926/1990), the strict conditional (Cariani & Rips, 2017; Kratzer, 1991), and the inferentialist or causal conditional (Cruz, Over, Oaksford, & Baratgin, 2015; Douven, 2015; Oberauer, Weidenfeld, & Fischer, 2007; Skovgaard-Olsen, Kellen, Krahl, & Klauer, 2017; Sloman & Lagnado, 2005).

More generally, basing the software on general, material conditionals whose truth or falsity has to be evaluated by coming up with or refuting counterexamples to them, makes it unclear to me to what extent the programme implements analogical reasoning, as opposed to conditional reasoning.

Further, the instructions given to participants, to assume that the association in the source domain (resp. the antecedent of the conditional) is certain, and that the conditional can be refuted by a single counterexample, seems to imply a binary approach to reasoning. In a binary approach it is possible to assess whether or not something follows logically from a given assumption. But it fails to scale up to the arguably more realistic situations in which we are uncertain about the information we start out with, and yet need to make decisions about which conclusions it is reasonable to infer from them with which degree of confidence (Adams, 1998). More recent probabilistic and Bayesian approaches to reasoning, and to cognitive science more generally, might be worth considering as potential generalisations of the binary case modelled by the authors (e.g. Gilio, Over, Pfeifer, & Sanfilippo, 2016; Hahn, 2011; Oaksford & Chater, 2020).

A minor question in connection with this is why the authors describe their approach as involving conditionals that are treated "as if" they were material conditionals, rather than treating them as material conditionals simpliciter. What is the gap that separates the "as if" situation from the actual situation of using material conditionals composed of?

4. AUTOMATION AND THE PROBLEM OF DEFINING GOOD (ANALOGICAL) REASONING

The authors state that they aim to automate the software in two respects: on the one hand, they aim to minimise the role of moderators, so that people could use the software for analogical discussion largely on their own. On the other hand, they aim to make the software usable

not only by people, but also by hypothetical future artificially intelligent agents.

What goals would automation in these two respects achieve? And how does this aim fit with the further aim of the authors for the software to work using currently available computational tools? Further, what would have to be in place for these forms of automation to be achievable?

The authors refer to deep neural networks, and to the large amount of data that can in principle be generated by users of the software. But if a data driven approach is followed, then how can the authors identify when the analogies developed by users are of high or low quality? How could they prevent the neural networks from just reproducing the biases and irrelevancies of human participants (Constantinou & Fenton, 2017)?

An answer to this latter question requires a definition of argument strength. The authors advocate a procedural definition, according to which an analogy becomes stronger, the more attacks it has withstood. Such a definition has the advantage of facilitating automatized attribution of argument quality. However, it seems unable to differentiate arguments that have survived criticism because the criticism was weak, from those that have survived criticism because their high quality makes them difficult to criticise. It also seems unable to ascertain the strength of an argument that has not yet been subjected to criticism. Further, this radical context relativity of argument strength appears at odds with the instructions given to participants to assess logical entailment relations.

When developing an account of argument strength, a further question relevant to automation processes is to what extent the strength of an argument can be determined without understanding its content. If at least some aspects of argument strength are content specific (Hahn & Oaksford, 2007), then a software programme would only be able to categorise arguments on the basis of their quality if it solved the frame problem (of determining what information is relevant to a task) and the Chinese room problem (of establishing a mapping from syntactic rules to semantic content). These long standing open issues would constitute vast challenges going far beyond the author's aim for the software to work using currently available tools.

5. THE NEED FOR A COMMON GOAL

The authors state that a central goal of the restrictions that the software places on the argumentative process is to prevent reasoners from "playing the system", using irrelevant or deceptive tactics to maximise their personal benefit, rather than working together to achieve the common goal of developing a high quality piece of analogical reasoning.

The proposal that such tactics can be reduced through procedural restrictions is very interesting, and could potentially be generalised to real world argumentative contexts not mediated by software. It may be interesting to enquire about empirical work conducted e.g. in the field of politics about the impact of following, or violating, procedural rules and conventions about respect and politeness in argumentative interactions. Further work in this area could also explore the limitations of procedural restrictions when taken on their own, and the need to promote the development of common goals and willingness to cooperate at a more general level, e.g. by pointing out the different factors, including mathematical ones, that generally make it a win-win endeavour (Page, 2007).

6. CONCLUSION

In this interesting paper, the authors present a software programme aimed at improving analogical reasoning by placing procedural restrictions on the moves allowed in an argumentative dialog. These rules aim to focus the reasoners' attention on the warrants of an analogy, i.e. on the type and strength of the link between its source and its target domain. By making this link explicit, as well as requesting reasons for different moves made and rewarding cooperative behaviour with additional turns in the dialog, the programme aims to promote constructive argumentation towards the common goal of formulating analogies that are relevant and robust in the face of criticism.

Potential avenues for further specification and development could be a generalisation of the types of argument modelled to situations in which agents have to draw reasonable conclusions from uncertain information, and consider alternatives to the material conditional that can capture this uncertainty. It might also be useful to explore limits and alternatives to procedural definitions of argument strength, as well as ways of promoting cooperative reasoning towards common goals beyond the placement of procedural restrictions. A further advantage of the latter may be to avoid a "homunculus problem" regarding who gets to be a moderator and is allowed to determine the criteria for valid argumentative moves.

REFERENCES

Adams, E. (1998). *A primer of probability logic.* Standford, US: CLSI publications.

Bartha, P. F. A. (2010). By parallel reasoning: The construction and evaluation of analogical arguments. Oxford University Press.

Cariani, F., & Rips, L. J. (2017). Conditionals, context, and the suppression effect. *Cognitive Science, 41*(3), 540-589.

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

Constantinou, A., & Fenton, N. (2017). Towards smart-data: Improving predictive accuracy in long-term football team performance. *Knowledge-Based Systems, 124,* 93-104.

Cruz, N., Over, D., Oaksford, M., & Baratgin, J. (2016). Centering and the meaning of conditionals. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1104– 1109). Austin, US: Cognitive Science Society.

De Finetti, B. (1936/1995): The logic of probability. *Philosophical Studies*, 77, 181–190.

Douven, I. (2015). How to account for the oddness of missing-link conditionals. *Synthese.* DOI: 10.1007/s11229-015-0756-7.

Edgington, D. (1995). On conditionals. *Mind*, 104*,* 235-329.

Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford, UK: Oxford University Press.

Gilio, A., Over, D. E., Pfeifer, N., & Sanfilippo, G. (2016). Centering and compound conditionals under coherence. In M. B. Ferraro, P. Giordani, B. Vantaggi, M. Gagolewski, M. A. Gil, P. Grzegorzewski, & O. Hryniewicz (Eds.), *Soft methods for data science* (pp. 253-260). Springer.

Hahn, U. (2011). The problem of circularity in evidence, argument, and explanation. *Perspectives on Psychological Science, 6*(2), 172-182.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, *114*, 646-678.

Jeffrey, R. C. (1991). Matter of fact conditionals. *Proceedings of the Aristotelian Society, Supplementary Volumes, 65*, 161-183.

Kratzer, A. (1991). Conditionals. In A. von Stechow, & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 651–656). Berlin: de Gruyter.

Oaksford, M, & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, UK: Oxford University Press.

Oaksford M., & Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Thinking & Reasoning, 19*, 346-379.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology, 71,* 12.1-12.26.

Oaksford, M., & Hahn, U. (2007). Induction, deduction, and argument strength in human reasoning and argumentation. In A. Feeney, & E. Heit (Eds.), *Inductive Reasoning* (pp. 269-301). Cambridge: Cambridge University Press.

Oberauer, K., Weidenfeld, A., & Fischer, K. (2007). What makes us believe a conditional? The roles of covariation and causality. *Thinking & Reasoning, 13*(4), 340-369.

Over, D., & Cruz, N. (2019). Philosophy and the psychology of conditional reasoning. In A. Aberdein, & M. Inglis (Eds.), *Advances in Experimental Philosophy of Logic and Mathematics*. London, UK: Bloomsbury Publishing Plc.

Page, S. E. (2007). The difference: How the power of diversity creates better groups, firms, schools, and societies. Princeton, New Jersey: Princeton University Press.

Ramsey, F. P. (1926/1990). Truth and probability. In D. H. Mellor (Ed.), *Philosophical papers* (pp. 52-94). Cambridge, UK: Cambridge University Press.

Skovgaard-Olsen, N., Kellen, D., Krahl, H., & Klauer, C. K. (2017). Relevance differentially affects the truth, acceptability, and probability evaluations of "and", "but", "therefore", and "if-then". *Thinking & Reasoning, 23*(4), 449-482.

Sloman, S. A., & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science, 29,* 5-39.

Stern, R., & Hartmann, S. (2018). Two sides of Modus Ponens. *The Journal of Philosophy, 115*(11), 605-621.

Toulmin, S. E. (2003). The uses of argument. Cambridge University Press.