

Formal Argumentation: A Human-centric Perspective

Federico Cerutti, Nava Tintarev, and Nir Oren

University of Aberdeen, UK

{n.tintarev, f.cerutti, n.oren}@abdn.ac.uk

Abstract. A reason why formal argumentation is one of the most popular paradigms for non-monotonic reasoning is its relatively straight-forward semantics. In fact, argumentation provides a natural explanation for the reasoning process, while its semantics are designed to be intuitively understandable to people who are not familiar with formal argumentation. In this paper we propose a roadmap for a systematic evaluation of argumentation theories as a bridge between automated reasoning and human-centred reasoning.

1 Overview

Argumentation has become a popular approach to non-monotonic reasoning, with applications ranging from law [1], to practical reasoning [2], to underpinning decision support systems [3]. Reasoning using formal argumentation takes place in several steps. First, facts are encoded in some logic within a knowledge base. Arguments are then constructed from the knowledge base, according to the rules of the formalism being used. Next, attacks among arguments are generated. Then the justification status of each argument is evaluated, and sets of collectively acceptable (w.r.t. a semantics) arguments are identified. Finally, the conclusions of arguments within acceptable sets are associated with potentially valid inferences according to the argumentation framework [4].

One clear desideratum of any reasoning system is that its semantics agree with human intuition. While it is often claimed that it is “intuitively obvious”, to our knowledge, this claim has only been tested in [5, 6]. In [5], subjects evaluate the status of arguments, and they generally agree with the extensions obtained according to abstract argumentation semantics. However, they do not consider arguments derived from a formal knowledge base. [6] uses natural language techniques to generate arguments and attacks, and evaluate them using argumentation semantics but without human subjects.

In order to address the human-centric perspective of formal argumentation, we must first determine to what extent automated argumentation reasoning correspond to human reasoning. Below we introduce a first experiment in this direction, and the envisaged research plan following from it.

2 The Experiment

We aim at validating the correspondence between human reasoning and conclusions obtained from a formal argumentation framework. To validate this hypothesis, we began with a simple set of logical arguments. We encoded them using Prakken and Sartor’s

argumentation system [1, Sect. 5], as it is inspired by human reasoning, and allows for arguments in favour of and against preferences. We then handcrafted natural and fluent texts as an interface to formal arguments. This enabled us to compare the outputs from the formal system with the positions and reasoning taken by experimental participants.

Using Mechanical Turk (<https://www.mturk.com/>), we surveyed 161 participants who successfully completed an English test. These were presented with a natural language description [7] of a set of arguments and a questionnaire. We considered two arguments supporting contradictory conclusions, together with a third argument expressing a preference between the first two arguments. Then, we introduced another argument which reinstated the original contradiction between arguments (via an attack against one of the two contradicting arguments or against the preference argument).

Each participant completes a single questionnaire for a specific set of arguments (i.e. a between-subjects experimental design). The questionnaire is the medium by which we evaluate participants' (1) perceived relevance and acceptance for the concept of "argument" (i.e. agreement on seeing a "natural language statement" as a "coherent position") and (2) opinions about arguments' justification statuses. Although the analysis of this experiment is ongoing, first results suggest that while participants largely follow the theory, the domain affects the perception and the justification status of arguments, as well as how preferences are applied. This provides evidence that a human-centric perspective of formal argumentation is a complex and intriguing topic, and it has the potential to reshape the research field of argumentation in artificial intelligence.

Completing the analysis of such an experiment is just the beginning. In fact, we will also evaluate additional formal argumentation systems. In addition, we will examine if and how the subjectivity of domain causes a divergence between the expected output of the system and what human decision-makers decide [8]. Finally, we will investigate how humans reason about preferences, and their alignment with argumentation in more detail, as well as more complex interactions between arguments.

References

1. Prakken, H., Sartor, G.: Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics* 7(1-2) (January 1997) 25–75
2. Atkinson, K., Bench-Capon, T.J.M.: Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artificial Intelligence* 171(10-15) (2007) 855–874
3. Cerutti, F.: Decision Support through Argumentation-Based Practical Reasoning. In: *Proc. IJCAI 2011*. (2011) 2786–2787
4. Prakken, H., Vreeswijk, G.A.W.: Logics for defeasible argumentation. In: *Handbook of Philosophical Logic*, 2nd Edition. Kluwer Academic Publishers, Dordrecht (2002) 219–318
5. Rahwan, I., Madakkatel, M.I., Bonnefon, J.F., Awan, R.N., Abdallah, S.: Behavioural Experiments for Assessing the Abstract Argumentation Semantics for Reinstatement. In: *Cog. Sci.* (2010)
6. Cabrio, E., Villata, S.: Natural Language Arguments: A Combined Approach. In: *Proc. ECAI*. (2012)
7. Cerutti, F., Tintarev, N., Oren, N.: Human-Argumentation Experiment Pilot 2013: Technical Material. <http://goo.gl/FvgC08> (2014)
8. Nawwab, F.S., Dunne, P.E., Bench-Capon, T.J.M.: Exploring the Role of Emotions in Rational Decision Making. In: *Proceedings of COMMA 2010*. IOS Press (2010)