Transparent Rational Decisions by Argumentation

Francesca Toni, Robert Craven, Xiuyi Fan
Imperial College London, London, United Kingdom
{ft,robert.craven,xf309}@imperial.ac.uk

May 7, 2013

Abstract

There is a well-documented indication that several applications would benefit from the *transparency* afforded by *argumentation* to support decision-making where standard decision theory is not useful, e.g. in healthcare. However, to date, research on argumentation-based decision making has only been partially successful in realising its promise. In our view this is predominantly due to its lack of theoretical validation in the form of *rationality* properties and its disregard for the interplay between *individual rationality and social good* when used in collaborative settings. We discuss how to address these challenges for the promise of argumentation-based decision-making to be fully realised as a principled mechanism for transparent and rational decision-making.

Context

Argumentation, initially studied in philosophy and law [20, 16], has been researched extensively in AI and Computer Science in the last two decades [2, 18]. Simply stated, argumentation focuses on interactions where parties plead for and against some conclusion. In its most abstract form [6], an argumentation framework consists simply of a set of *arguments* and a binary relation representing the *attacks* between the arguments, and corresponds to a directed graph $(args, attacks)$, where $args$ is the set of arguments and $attacks \subseteq args \times args$. By instantiating the notion of arguments and the attack relations, different 'structured' argumentation frameworks can be obtained, predominantly based on logic, within which arguments have a concrete structure and attacks are determined on the basis of this structure. For example, in assumption-based argumentation [7], arguments are obtained recursively from a given set of logical *rules* and are supported by these rules and *assumptions*. Also, an argument attacks another if the former supports a claim conflicting with some assumption in the latter,

where conflicts are given in terms of the notion of *contrary* of assumptions. Rules, assumptions and contraries are defined in terms of an underlying *logical language*.

Within argumentation framework (of abstract or structured form) attacks represent conflicts between arguments, which in turn provide abstractions of conflicts between conflicting views of an individual or clashes between views of different individuals. Argumentation frameworks are equipped with methods (typically referred to as 'semantics') for determining winning arguments, and thus resolve the underlying conflicts.

Argumentation provides a powerful mechanism for dealing with incomplete, possibly inconsistent information and for the resolution of conflicts and differences of opinion amongst different parties. Further, it is useful for justifying outcomes. Thus, argumentation can, in principle, support several aspects of *decision-making* [19], either by individual entities performing critical thinking (needing to evaluate pros and cons of conflicting decisions) or by multiple entities dialectically engaged to come to mutually agreed decisions (needing to assess the validity of information the entities become aware of and resolve conflicts), especially when decisions need to be transparently justified (e.g. in healthcare).

Argumentation-based decision-making is a form of *qualitative decision theory* (e.g. see [5]), understood as an alternative to classical *quantitative decision theory* when a decision problem cannot be easily or usefully formulated in standard decision-theoretic terms using decision tables, utility functions and probability distributions. Argumentation has been proposed for decision-making *under certainty* (where the outcomes of decisions are known to the decision maker), e.g. in [12, 13], *strict uncertainty* (where the outcomes of decisions are uncertain and no probabilistic information is available), e.g. in [1] and [9], and also for decision *under risk* (where some probabilistic information is known), e.g. in [15, 8]. In all these settings argumentation is used in the context of *individual decision-making*. In addition, argumentation has been advocated to support *collaborative decision-making* (e.g. in [11]), where different decision-makers interact to reach jointly agreed decisions or separate but compatible individual decisions.

Challenges

Argumentation-based decision-making is empowered by highly *transparent* methods. Indeed, argumentation affords a transparent depiction of the detailed logical structure of arguments and conflicts between them, and argumentation semantics provide a transparent way to resolve these conflicts. However, argumentation-based decision-making predominantly lacks means of formally evaluating decisions as *rational*. Moreover, when decision-making is collaborative amongst different parties and stakeholders, the interplay between individually and socially rational decisions needs to be fully understood and strategic behaviour of individual parties towards individual good and away from social good needs to be controlled.

Transparent versus rational decisions

Standard decision theory is well studied as a mechanism for *rational choice*, but has paid little attention to the structure and content of reasons brought to bear to support or argue against given choices, as pointed out, for example, in [4]. Indeed, in standard decision theory, the outcomes of choices are typically assigned values which are then treated as embodying a cardinal or ordinal preference relation, with decision rules identifying good choices according to various decision criteria and under differing conditions of circumstantial knowledge (certainty, strict uncertainty, risk). We believe that a 'normative' analysis of argumentation-based decision-making, in terms of rationality criteria similar (in spirit at least) to those studied in standard decision-theory, is essential for argumentation-based decision-making to become a convincing, widely usable transparent as well as rational mechanism and have, as a result, a fruitful impact on applications. This normative analysis may beneficially draw also from the vast literature in AI on formal theories of defeasible reasoning in the presence of various forms of priorities and postulates regulating the outcomes of this form of reasoning (e.g. see [3, 10]. Indeed, the support that argumentation gives to a decision is in general defeasible: if a stronger, conflicting argument is present, then a weaker argument may be rejected.

Individual versus collaborative decision-making

Mechanism design in economics [14] is a well-developed framework for studying (and identifying conditions for avoiding) strategic behaviour of interacting individuals with conflicting preferences, with some early promise in the context of argumentation [17]. We believe that collaborative argumentation-based decision-making would benefit (if not require) a mechanism design analysis and formulation to guarantee the social rationality of decisions and control the strategic behaviour of individually rational stakeholders. We also believe that a novel normative perspective for argumentation-based dialogues between agents is needed to ascertain the rationality of collaborative decisions and strategic aspects of dialogues.

An example in healthcare

Consider the case of two physicians taking a decision about a cancer patient on whether to prescribe drug a or b, in the presence of two clinical trials, one of which recommends cancer drug a over b on the grounds that, whilst there is no significant difference in their effects on disease-free survival, a brings markedly lower risks of alopecia. The second clinical trial shows a very slightly higher probability of disease-free survival with b compared to a, and no difference in the likelihood of alopecia. However, the first

trial involved 40 patients who were followed for 5 years; the second trial involved 100 patients for 20 years. Here, there is a partial order over outcomes (perhaps involving some aggregation of orders over different attributes), but there is also an ordering—derived, in this instance, from the differing longevity and size of the trials—over the strength of the reasons afforded to choose drug a or b. Assume also that one of the physicians has a personal preference for drug a, and that the other has no knowledge of the second clinical trial. This decision-making problem requires for the physicians to share information (possibly against their personal preferences) and come to an agreed rational decision taking into account all factors, and transparent in being explainable to each other and the patient.

References

[1] L. Amgoud and H. Prade. Making decisions from weighted arguments. In *Decision theory and multi-agent planning*. Springer, 2006.

[2] T. Bench-Capon and P. E. Dunne. Argumentation in artificial intelligence. *Art. Int.*, 171, 2007.

[3] G. Brewka and T. Eiter. Prioritizing default logic. In S. Hölldobler, editor, *Intellectics and Computational Logic*, volume 19 of *Applied Logic Series*, pages 27–45. Kluwer, 2000.

[4] F. Dietrich and C. List. A Reason-Based Theory of Rational Choice. *Noûs*, 2011.

[5] D. Dubois, H. Fargier, and P. Perny. Qualitative dec. theory with preference relations and comparative uncertainty: An axiomatic approach. *Art. Int.*, 148, 2003.

[6] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Art. Int.*, 77(2), 1995.

[7] P. M. Dung, R. A. Kowalski, and F. Toni. Assumption-based argumentation. In *Argumentation in AI*. Springer, 2009.

[8] P. M. Dung and P. M. Thang. Towards probabilistic argumentation for jury-based dispute resolution. In *Proc. COMMA*, 2010.

[9] P. M. Dung, P. M. Thang, and F. Toni. Towards argumentation-based contract negotiation. In *Proc. COMMA*, 2008.

[10] J. Hansen. Prioritized conditional imperatives: problems and a new proposal. *Autonomous Agents and Multi-Agent Systems*, 17(1):11–35, 2008.

[11] N. I. Karacapilidis and D. Papadias. Computer supported argumentation and collaborative decision making: the HERMES system. *Inf. Syst.*, 26(4), 2001.

[12] P.-A. Matt, F. Toni, T. Stournaras, and D. Dimitrelos. Argumentation-based agents for eprocurement. In *Proc. AAMAS*, 2008.

[13] P.-A. Matt, F. Toni, and J. Vaccari. Dominant decisions by argumentation agents. In *Proc. ArgMAS*, 2009.

[14] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[15] S. Parsons. Normative argumentation and qualitative probability. In *Proc. ECSQARU-FAPR*, 1997.

[16] C. Perelman. *Justice, Law and Argument*. Reidel, Dordrecht, 1980.

[17] I. Rahwan and K. Larson. Logical mechanism design. *Knowledge Eng. Review*, 26(1), 2011.

[18] I. Rahwan and G. R. Simari, editors. *Argumentation in AI*. Springer, 2009.

[19] R. D. Rieke, M. O. Sillars, and T. R. Peterson. *Argumentation and critical decision making*. Pearson, 2005.

[20] S. Toulmin. *The uses of arguments*. Cambridge University Press, 1958.