

INTRODUCTION

Trust, argumentation and technology

Fabio Paglieri*

Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, Via S. Martino della Battaglia 44, Rome 00185, Italy

(Received 6 April 2014; final version received 6 April 2014)

Trust and argumentation have both been the focus of much scholarly interest. While each topic has originated its own multidisciplinary literature (for an overview, see [Castelfranchi & Falcone, 2010](#); [van Eemeren et al., 2014](#)), more recently they have started to be pursued in conjunction with each other. There are several reasons why looking at trust and argumentation together is considered especially fruitful. First, both phenomena share a common function: dealing with change and uncertainty in complex social environments. This is why they both constitute key pillars in the emerging and rapidly growing new field of agreement technologies ([Ossowski, 2013](#)). But their connections run deeper than that: on the one hand, the degree of (justified) trust we have in the source of an argument should affect how that argument is reconstructed and assessed (see for instance [Paglieri et al., 2014](#)); on the other hand, argumentation can be a powerful tool to reason about trust, making sure such trust is well-founded (this is the approach advocated in [Tang, Cai, McBurney, Sklar, & Parsons, 2012](#), for instance). To put it simply, trust can assist in argument assessment, and argumentation can support trust reasoning. In light of the growing relevance of argumentation in computer science (witness this journal, as well as [Rahwan & Simari, 2009](#); [Reed & Norman, 2004](#)), it is not surprising that its interplay with trust is attracting so much attention.

Albeit technological innovation and in particular the Internet have made trust and argumentation pressing concerns for anyone interested in fostering a better worldwide communication ecology, their connection is more fundamental than that, since it revolves around a basic epistemological problem: whether a certain attitude, either practical (a goal) or epistemic (a belief), can be properly justified by an act of trust, or if it requires the kind of careful scrutiny demanded by argumentation. This in turn invites a further question: is this a true dilemma? Are trust and argumentation friends or foes? For a long time, and mostly in the philosophical literature, trust has tended to be seen as an abdication of critical scrutiny, and thus at odds with argumentation, which is the full exercise of that scrutiny. In epistemology, the problem of testimony and its justification (for a discussion, see [Faulkner, 2011](#); [Origgi, 2004](#)) hinges on whether trusting something on the say-so of a source can ever be justified, and if so how – an issue dating back at least to David Hume and Thomas Reid in the eighteenth century. Similarly, appeal to the authority of a source in argumentation has long been regarded as fallacious (*ad verecundiam*), as was the case also for the practice of attacking an argument based on the shaky credentials of the arguer (*ad hominem*). With the revival of interest in fallacy theory initiated, among others, by [Hamblin \(1970\)](#), a more nuanced approach has become prominent, and now most scholars accept that arguments for or against trusting a source are not necessarily fallacious, depending on a variety of factors (thoroughly discussed in [Walton, 1998, 2008](#); [Woods, 2013](#)). More generally, it has become clear that trust, far from being a blind attitude,

*Email: fabio.paglieri@istc.cnr.it

is in fact an exercise of critical judgment, which is precisely why argumentation can be used to reason about it, as noted above.

Against the backdrop of the long, intertwined history of these two notions, this special issue aims to showcase recent breakthroughs in the study of trust and argumentation, with special emphasis on their mutual interaction in relation to technology. Andrew Koster, in ‘Trust and argumentation in multi-agent systems’, offers an integrative review of computational trust and argumentation, especially in relation to agent-based models and technologies. He stresses how these two areas tackle different facets of reasoning in uncertain and dynamical social contexts. A special merit of his contribution is to offer an updated and balanced view on both positive results and ongoing difficulties of the various approaches proposed to integrate trust and argumentation.

In ‘On a razor’s edge: evaluating arguments from expert opinion’, Douglas Walton discusses trust as a method for evaluating arguments based on the opinion of an alleged expert. His approach uses the argumentation scheme for argument from expert opinion along with its matching set of critical questions, showing how to use it in three formal computational argumentation models to analyse and evaluate instances of such arguments. In his conclusions, Walton stresses that, from an argumentation point of view, it is better to critically question arguments from expert opinion than to accept or reject them based solely on trust. Here, the qualification ‘solely’ plays a central role: far from relegating once again trust among noxious forms of epistemic laziness, Walton’s method provides us with the analytical tools to establish when trust in the opinion of experts is justified, and when it is not.

A similar strategy is advocated and extended to other argument schemes by Simon Parsons, Katie Atkinson, Zimi Li, Peter McBurney, Elizabeth Sklar, Munindar Singh, Karen Haigh, Karl Levitt and Jeff Rowe. In ‘Argument schemes for reasoning about trust’, they emphasise the relevance of trust in any decentralised system as a mechanism by which an agent can deal with the inherent uncertainty regarding the behaviours of other parties, as well as the uncertainty in the information it shares with those parties. Building on recent efforts to use argumentation to reason about trust (Tang et al., 2012), the authors provide a set of argument schemes, abstract patterns of reasoning that apply in multiple situations, geared towards trust and inspired to Walton’s seminal work in this area (for a recent overview, see Walton, Reed & Macagno, 2008). In particular, Parsons and collaborators describe schemes in which one agent, A, can establish arguments for trusting another agent, B, directly, as well as schemes that A can use to construct arguments for trusting C, where C is trusted by B. Finally, for both types of schemes, a set of critical questions is offered that identify the situations in which these schemes can fail.

Leila Amgoud and Robert Demolombe are interested in looking at trust as the foundation of more secure interactions in agent-based software applications. In ‘An argumentation-based approach for reasoning about trust in information sources’, they focus on trust in information sources, building on previous work (Amgoud & Cayrol, 2002; Demolombe, 2004) and proposing an argumentation-based model for reasoning about agent’s beliefs and trust. They articulate six basic forms of trust in information sources and present a formal representation for each of them in a modal logic language: then, starting from a beliefs base which contains such formulas and others, they show how to build arguments in favour of each form of trust and discuss how these arguments may interact. This model thus plays two key roles: (i) it allows reasoning about trust in information sources and (ii) it enables agents to critically handle information received from other sources.

The last contribution in this issue puts a different spin on the relationship between trust in arguers and the assessment of their arguments. In ‘Trust, relevance and arguments’, Paglieri and Castelfranchi extend their previous work on trust in the relevance of information sources (2012) to the analysis of argumentative relevance, suggesting that trust plays a central role in it as well. They begin by distinguishing two types of argumentative relevance: internal relevance, i.e. the extent by which a premise has a bearing on its purported conclusion, and external relevance, i.e. a

measure of how much a whole argument is pertinent to the matter under discussion, in the broader dialogical context where it is proposed. Then they argue that judgments of internal relevance heavily rely on trust, and that such trust, although occasionally misplaced (e.g. in some so-called fallacies of relevance), is nonetheless often justified by either epistemic or pragmatic considerations. They conclude by sketching potential methods to formally model trust in argumentative relevance, some of which based on the same approach championed by Amgoud and Demolombe in their paper, and briefly discussing the technological implications of this line of research.

Taken together, the papers in this special issue provide a partial but thought-provoking bird's-eye view of the rich interplay between trust and argumentation, and of their mutual influence on how the future technological landscape is shaping up. The true implications of these phenomena, however, are bound to be even deeper than those hinted at by these contributions: for instance, the impact of e-participation on politics will make trust and argumentation dramatically important in determining whether the ICT revolution will ultimately empower or undermine the democratic system. The jury is still out, and unfettered optimism is no longer an option: Krastev (2011) has convincingly showed how the continuous monitoring of politicians' views and deeds by the public (reciprocated by the politicians' obsessions with poll results) is not necessarily an instrument of transparency, but rather an indicator of utter distrust in public servants, and a pernicious source of bad-but-popular decisions, made by politicians with the sole aim of appeasing their voters. More generally, the political and economical impact of 'trust crises' is apparent to even the most distracted observer, as is the fact that they often coincide with periods of deterioration in the argumentative quality of public debate, both among politicians and in the general public. Technology is not neutral with respect to these dynamics. It certainly grants citizens unprecedented access to public debate: but it can also facilitate *either* poor reasoning and generalised distrust, *or* good argumentation and a climate of (vigilant) trust. Argumentation and trust scholars need to join forces to make sure that the latter scenario comes to pass.

Acknowledgements

I am grateful to the members of the Goal-Oriented Agents Lab (GOAL) and the Trust Theory and Technology group (T3) at the ISTC-CNR, and in particular to Rino Falcone, for supporting my editorial work on this issue, and to all reviewers for providing excellent comments and thorough criticisms on the papers collected here. My gratitude goes also to the editorial team of the journal, and in particular to Floriana Grasso and Chris Reed, for accepting this issue for publication and for their constant help throughout the editorial process. As for the staff at Taylor & Francis, their kind and competent assistance at all stages was invaluable and much appreciated. Finally, none of this would have been possible without the excellent contribution of all the authors: to them go my deepest thanks and appreciation.

Funding

This work was supported by the PON research project 'Interoperable Cloud Platforms for Smart Governments' (PRISMA), founded by the Italian Ministry for Education, University and Research (MIUR), and by the European Network for Social Intelligence (SINTELNET, <http://www.sintelnet.eu/>).

References

- Amgoud, L., & Cayrol, C. (2002). A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, 34, 197–216.

- Castelfranchi, C., & Falcone, R. (2010). *Trust theory: A socio-cognitive and computational model*. London: Wiley.
- Demolombe, R. (2004). Reasoning about trust: A formal logical framework. In C. Jensen, S. Poslad, & T. Dimitrakos (Eds.), *Trust management: Proceedings of iTrust 2004* (pp. 291–303). Berlin: Springer.
- van Eemeren, F.H., Garssen, B., Krabbe, E.C.W., Snoeck Henkemans, A.F., Verheij, B., & Wagemans, J.H.M. (2014). *Handbook of argumentation theory*. Berlin: Springer.
- Faulkner, P. (2011). *Knowledge on trust*. Oxford: Oxford University Press.
- Hamblin, C. (1970). *Fallacies*. London: Meuthen.
- Krastev, I. (2011). The age of populism: Reflections on the self-enmity of democracy. *European View*, 10, 11–16.
- Origgi, G. (2004). Is trust an epistemological notion? *Episteme*, 1(1), 61–72.
- Ossowski, S. (Ed.) (2013). *Agreement technologies*. Berlin: Springer.
- Paglieri, F., & Castelfranchi, C. (2012). Trust in relevance. In S. Ossowski, G. Vouros, & F. Toni (Eds.), *Proceedings of agreement technologies 2012* (pp. 332–346). Tilburg: CEUR-WS.org.
- Paglieri, F., Castelfranchi, C., da Costa Pereira, C., Falcone, R., Tettamanzi, A., & Villata, S. (2014). Trusting the message and the messenger: Feedback dynamics from information quality to source evaluation. *Computational and Mathematical Organization Theory*. doi:10.1007/s10588-013-9166-x
- Rahwan, I., & Simari, G. (Eds.) (2009). *Argumentation in artificial intelligence*. Berlin: Springer.
- Reed, C., & Norman, T. (Eds.) (2004). *Argumentation machines. New frontiers in argument and computation*. Dordrecht: Kluwer.
- Tang, Y., Cai, K., McBurney, P., Sklar, E., & Parsons, S. (2012). Using argumentation to reason about trust and belief. *Journal of Logic and Computation*, 22(5), 979–1018.
- Walton, D. (1998). *Ad hominem arguments*. Tuscaloosa, AL: University of Alabama Press.
- Walton, D. (2008). *Witness testimony evidence: Argumentation, artificial intelligence and law*. Cambridge: Cambridge University Press.
- Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge: Cambridge University Press.
- Woods, J. (2013). *Errors of reasoning. Naturalizing the logic of inference*. London: College.