

Arguably argumentative: A formal approach to the argumentative theory of reason

No Author Given

No Institute Given

Abstract.

1 Introduction

The idea that thought entails existence, or at least presupposes it, is a conceptual pillar of analytic philosophy. It is hard to disagree that it has a certain appeal: "I think, so I am, so I might as well go on thinking". For the sceptic, however, it also begs the following question: what is the "I" in such a line of thought? The armchair philosopher might be too busy with his thoughts to worry about it, but in the field of social psychology, particularly in the tradition going back to the work of George Herbert Mead and the Chicago school, it has well established pride of place (Mead 1967).

The best known theories developed in this field all agree that any "I" is essentially a social construction; you exist as a thinker only because you interact with other thinkers. In particular, social interaction is partly constitutive of self, not merely emergent from it. Consequently, reason itself is emergent from social contact, and the view that individual rationality is the basis for rational interaction must be rejected. Rather, interaction and reasoning should be seen as mutually dependent notions, on the basis of which more subtle notions of rationality can be explored.

This idea is both convincing and powerful, and it is becoming increasingly important to many different fields of research, including economy, law, biology and artificial intelligence (Blume and Durlauf 2001; Dworkin 1986; Waal and Ferrari 2010; Benthem 2011; Ossowski 2013). In all these research areas, there is a trend towards viewing rationality as fundamentally embedded in a social context. Importantly, this context is seen as important not only because people are social and tend to interact, but also because *who* they are, *what* they want, and *why* they want it, tends to depend on how they engage with each other and their environment. Hence the individual – the *agent* in the context of formal models – is himself in need of more subtle analysis, in terms of the same structures that are used to describe important aspects of the economic, legal, environmental and computational world that contains him.

To accommodate this point of view across different domains, we are in need of better theoretical foundations, allowing us to investigate the relationship between reasoning and interaction, taking into account that they are co-dependent and co-evolving. In paper, we address this challenge, and we do so using formal

logic, drawing on tools and techniques developed in the context of multi-agent systems. The connection between various branches of social science and formal logic and computer science has received much attention in recent years, and it has led to a surge of interesting interdisciplinary research (Wooldridge 2009; Benthem 2008; Verbrugge 2009; Ditmarsch, van der Hoek, and Kooi 2007; Parikh 2001).

While much recent work in applied logic has been devoted to modelling agency and interaction, the standard starting point is still that agents reason in strict adherence to some common standard of correctness, specified by some given formal logic. Also, it is typically assumed that rational interaction emerges from the fact that agents are individually rational in some appropriate sense, for instance because they seek to maximize some given utility function. In this paper, we argue that in order to provide adequate formal foundations for rational interaction we need to depart from such reductionist assumptions. We point to the argumentative theory of reason, introduced in (Mercier and Sperber 2011), as an alternative approach, and we sketch a *formal* representation of basic elements of this theory. We show, in particular, how existing tools and techniques in contemporary logic allows us to formulate systems of dynamic logic for multi-agent argumentation which can be used to encode and explore key theoretical aspects, as well as facilitate modelling of concrete systems.

The structure of the paper is as follows. In Section 2 we present the necessary background on the argumentative theory of reason and the distinct notion of argumentation relied on in the theory of argumentation frameworks, as studied in artificial intelligence. We discuss the differences between these two notions of argumentation, and we argue that in order to use argumentation frameworks to arrive at logics for representing the argumentative theory, we need to conceptualize argumentation frameworks as subjective representations of semantic content, on the basis of which deliberation can take place. We argue that a fundamental question raised by the argumentative theory, which can then be analysed by formal logical tools, is the question of how argumentative deliberation works, and how it can sometimes create a common representation, a *consensus* among participants. We propose, in particular, that the argumentative theory implicitly relies on, and suggests further study of, *social* rationality constraint – imposed at the deliberative level, and formulated with respect to the outcome of deliberation. Moreover, we argue that these constraints are *not* reducible to corresponding notions of rationality that applies to individual reasoners, who are instead characterized by a distinct form of *argumentative rationality*, in that they seek primarily to maximize their influence, to win as many arguments as possible.

In Section 3 we introduce *deliberative Kripke frames*, a versatile formal semantics based on modal logic which gives us access to an abstract view of argumentative deliberation, well suited for further exploration of core theoretical aspects. We provide some examples of semantic modeling facilitated by this formalism, and we go on to present a simple modal language to reason about argumentative social processes. We then motivate what we believe to be the main

challenge for future work: how to characterize interesting notions of social rationality using theories in modal logic. We present a few preliminary suggestions in this regard, but argue that more work is needed to explore different theories, in languages of different complexity and expressive power.

In Section 4 we discuss the limitations of our own approach, and suggest directions for future work, and in Section 5 we conclude.

- 2 Argumentative agents: Towards a semantics for individual reasoning based on argumentation**
- 3 Argumentative deliberation: Towards formalization of rational interaction using dynamic argumentation logic**
- 4 Discussion and future work**
- 5 Conclusion**