

Computing consensus: A logic for reasoning about deliberative processes based on argumentation

S & T

January 11, 2014

Abstract

Human reasoning does not take place in a vacuum, and human intelligence, whatever its exact nature might be, always manifests in an environment.

1 Background

Normative and descriptive. What we have, and what we are trying to add.

2 Tangent thoughts

The process of reasoning that agents perform in these environments have aspects in several fields.

The argumentative theory of reason presented by Mercier & Sperber in [3] discusses things which are interesting for our motion.

Why the hell do we cite this? Well, I suppose we have to... In artificial intelligence there is a growing interest in the issue, as witnessed for example by [4].

From logic [5].

Economy [1].

Law [2].

Biology *Missing reference to de Waal Primates and Philosophers: How Morality Evolved...*

3 The Ultimatum Game

This is a

4 Logic

In the classical understanding of logic as a normative ideal of reasoning, the function (**so to speak/footnote**) of reasoning is to emulate this ideal as close as possible. The best reasoner, then, is the reasoner which most “frequently” makes true claims/inferences/conclusions. If truth equates with utility, the rational agent (in the sense of being utility-maximizing) is the agent which is ideal in this mentioned sense. (Always makes true or valid inferences/conclusions.)

With the twist of Mercier & Sperber we get the notion that the perfect agent is the agent which always wins arguments. In a way, we can understand the thesis as keeping the notion of rationality as utility-maximizing, but we need a new profile for the agent’s utility.

Using argumentation theory as an underlying semantics for this process, we can show how arguments develop. Do we permit the agents to disagree on the *meaning* of arguments? This is not a common position in formal argumentation theory (**yet!**), but we will conduct this study permitting it as in this case, the analysis subsumes the alternative. We have shown in [White] how we can reason about the development of a state of an argument when the agents are permitted to disagree on the meaning of arguments using formal methods. However, in that analysis, we only discuss what is possible developments and pay no attention to the fact that different agents may have different *preferences*.

4.1 Twist of preference profile

So we are supposing that we have grounded the possible developments of an argument in [White] and we now need to make an ordering of the outcomes relative to the agents which reflects the *preferences of the agents*. When the rational reasoner is the reasoner which maximizes utility and utility is one of either i) the classical utility of obtaining truth, or ii) the novel suggestion of utility in winning arguments.

In obtaining *truth* we need a reference point. Our [White] logic does not permit us reason about proposition which are not modally bounded by the black connective. Propositions then always describe or are grounded in the developed state. Let us then introduce some reference model \models_T . This model doesn’t need to be a classical boolean propositional logic, but for simplicity of the exposition we will often assume this. We can easily incorporate this into the deliberative logic, but we will keep it outside for now.

When does the truth win in the development of an argument? Some alternative characterizations. Let ϕ be a non-modal proposition/argument.

$$M_T \models_T \phi \text{ and } M, q \models \blacksquare\phi \quad (1)$$

$$M_T \models_T \phi \text{ and } M, q \models \Diamond\blacksquare\phi \quad (2)$$

$$M_T \models_T \phi \text{ and } M, q \models \Diamond^*\blacksquare\phi \quad (3)$$

$$M_T \models_T \phi \text{ and } M, q \models \Diamond\Box\blacksquare\phi \quad (4)$$

$$M_T \models_T \phi \text{ and } M, q \models \Diamond^*\Box^*\blacksquare\phi \quad (5)$$

When we're interested "winning the debate", we may characterize success by

$$M, q \models \blacksquare\phi \quad (6)$$

$$M, q \models \Diamond^*(\Box\Diamond)^*\blacksquare\phi \quad (7)$$

$$M, q \models \Diamond^*\Box\Diamond\blacksquare\phi \quad (8)$$

What can we say about which preferences these two utilities induce? What are the differences and similarities between the two theories? And where, in this formalism, do we find confirmation bias?

5 Preferences

Along the lines of the assumption of logical omniscience in epistemic logic, we make an assumption of infinite argumentative opinions. Each agent has (and has access to) an argumentation framework covering the set of all arguments. This will generally be infinite. In the deliberative logic [White] we generally start the debate in some empty or finite argumentative scenario; that is that when the evaluation of possible developments are being considered, the agents have made no or a finite number of utterances. This is not strictly necessary, but simplifies/enables certain procedures like model checking. Either way, in a truthful debate we have some arguments which have been made $S \subseteq \Pi$ and some state of affairs E , $\bigcap_{a \in \mathcal{A}} A_a \subseteq E \subseteq \bigcup_{a \in \mathcal{A}} A_a$. Two such points may have different utility to an agent. Lets say we have points (S, E) and (S', E') .

6 Conclusion and Future work

7 TODO

- Add reference [White] : Computing consensus: A logic for reasoning about deliberative processes based on argumentation

References

- [1] L. Blume and S. Durlauf. The interaction-based approach to socioeconomic behavior. In S. Durlauf and P. Young, editors, *Social Dynamics*. Brookings Institutions Press, 2001.
- [2] R. Dworkin. *Laws Empire*. Harvard University Press, 1986.
- [3] H. Mercier and D. Sperber. Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34:57–74, 3 2011.
- [4] S. Ossowski, editor. *Agreement Technologies*, volume 8 of *Law, Governance and Technology Series*. Springer, 2013.

- [5] J. van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2011.