Theoretical Computer Science 603 (2015) 132–145

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

A logic of plausible justifications ☆

L. Menasché Schechter

Department of Computer Science, Federal University of Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 31 January 2013
Accepted 13 June 2014
Available online 14 July 2015

Keywords:
Epistemic logic
Justifications
Plausibility models
Axiomatic system
Multi-agent systems

ABSTRACT

In this work, we combine features from Justification Logics and Logics of Plausibility-Based Beliefs to build a logic for Multi-Agent Systems where each agent can explicitly state his justification for believing in a given sentence. Our logic is a normal modal logic based on the standard Kripke semantics, where we provide a semantic definition for the evidence terms and define the notion of plausible evidence for an agent, based on plausibility relations in the model. As we deal with beliefs, justifications can be faulty and unreliable. In our logic, agents can disagree not only over whether a sentence is true or false, but also on whether some evidence is a valid justification for a sentence or not. After defining our logic and its semantics, we provide a strongly complete axiomatic system for it, show that it has the finite model property, analyze the complexity of its Model-Checking Problem and show that its Satisfiability Problem has the same complexity as the one from basic modal logics. Thus, this logic seems to be a good first step for the development of a dynamic logic that can model the processes of argumentation and debate in Multi-Agent Systems.

1. Introduction and motivation

Epistemic logics [2] are a particular kind of modal logics [3] where the modalities are used to describe epistemic notions such as *knowledge* and *belief* of agents. Traditional epistemic logics are expressive enough to describe knowledge and belief of multiple agents in a multi-agent system, including higher-order notions, such as the knowledge of one agent about the knowledge of another, and some notions of knowledge and belief that are related to groups of agents, such as "everybody in a group knows…" or "it is common knowledge in a group…".

Nevertheless, such epistemic logics have two important limitations. The first is that the knowledge or belief of an agent is static, i.e., it does not change over time. One of the reasons for this is that, in such logics, it is not possible to describe communication between the agents. The second limitation is that the knowledge modeled by such logics is *implicit*, which means that if the agent knows something, then he knows it for some reason that remains unspecified.

In order to deal with the first limitation, *Dynamic Epistemic Logics* [4] were developed. In these logics, we can describe acts of communication between the agents. Such acts consist of *truthful* announcements that are made by one of the agents (or an external observer) to the other agents (or a sub-group of them).

In works such as [5–8], this framework of dynamic logics was extended so that not only knowledge, but also beliefs (which, unlike knowledge, may turn out to be actually false) could evolve over time. The semantics of such logics of *dynamic beliefs* is based on *Plausibility Models*, where each agent has a plausibility order for the possible states of the model and he

☆ A preliminary version of this work was published in the proceedings of WoLLIC 2012 [1].

E-mail address: luisms@dcc.ufrj.br.

http://dx.doi.org/10.1016/j.tcs.2015.07.018
0304-3975/

believes in those sentences that are true in the most plausible states according to his plausibility order. The change of beliefs is then modeled as changes in the plausibility orders of the agents.

In order to deal with the second limitation, *Justification Logics* [9–11] were developed. In these logics, instead of formulas simply stating "Agent i knows φ", we have formulas that state "t is agent i's justification (or evidence) for knowing φ". Thus, these are logics of *explicit* knowledge, where every knowledge that an agent has is accompanied by an explicit justification for it. This is why Justification Logics are also called *Logics of Explicit Knowledge* or *Logics of Evidence-Based Knowledge*.

In the processes of argumentation and debate, be it an internal debate or a public debate where each agent tries to convince an external observer of his particular point of view, it is unrealistic to say that all of the announcements are *truthful*. The realistic assumption is that the announcements are merely *sincere*, i.e., each agent believes in what he announces. However, in order to convince others of their belief, an agent should not only state what he believes in, but also *why* he believes in it. So, the appropriate logic to model these processes would be a *dynamic logic of evidence-based beliefs*. The approach that we propose in order to build such a logic is to define its semantics on Plausibility Models while enriching its language with explicit evidence terms, a feature that is inspired by the language of Justification Logics. The work [12] presents another approach to build a dynamic logic of evidence-based beliefs, but the two approaches present significant differences, as we discuss in more detail in the end of this section.

Keeping in mind that we want to model *beliefs*, it is important to notice that, in the logic we propose, justifications can be faulty and unreliable. Using Plausibility Models, we give a notion of plausible evidence, or plausible justification, for an agent. So, if an agent has a plausible evidence for a sentence, then he will believe in that sentence, but, as the evidence can possibly be faulty, this belief has the possibility to be false.

In this work, we take a first step in order to build such a logic for the description of the processes of argumentation and debate. We build a normal modal logic (for the definition of normal modal logics, [3] can be consulted) where we can describe the plausibility of evidences for all the different agents, give a sound and strongly complete axiomatic system for this logic, show that it has the finite model property and analyze the complexities of its Model-Checking and Satisfiability problems.

As our next step, we plan to build a dynamic logic of *explicit* beliefs, adding to the present logic the actions that would model the communications between agents during the processes of argumentation and debate. This is not a trivial task. The standard announcements that describe changes of knowledge [4], sometimes called *hard announcements*, are too strong for our needs, since they are required to be *truthful* and not merely *sincere* (using such announcements without respecting the requisite that the announced formula should be true can generate logical inconsistencies). On the other hand, the standard announcements that describe changes of beliefs, called *belief upgrades* or *soft announcements* [13] are too weak, since, even tough they are only required to be *sincere*, they still make the agents receiving the announcement start believing in it, regardless of their current beliefs. In our desired framework, the announcement of a sentence should be accompanied by a justification as to why the agent performing the announcement believes in it. Then, each agent receiving the announcement should judge by himself whether he should start believing or not in the announced sentence, based on his current beliefs both about what was announced and about the justification that was given.

There are, in the literature, works that combine features from Justification Logics and Dynamic Epistemic Logics. Yavorskaya [14] developed the first proposal of a Justification Logic with communication between the agents. However, these communication actions were extremely simple. Later, the series of works [15–17] developed logics that add to Justification Logics a series of communication actions, some rather complex. However, those actions are all from the family of *hard announcements*, so they cannot be used for our purpose. Our combination of explicit evidence terms, inspired by the language of Justification Logics, and Plausibility Models, using these evidence terms to model explicit *beliefs* of the agents in such models, seems to be a novel approach. As was mentioned above, [12] also developed a logic of evidence-based beliefs, but, unlike our logic, it has no explicit evidence terms in the language and some of the modalities are non-normal. Besides that, also unlike our logic, no complete proof system for that logic is presented.

The rest of this paper is organized as follows. In Section 2, we introduce the necessary concepts that are used as building blocks for our logic: Justification Logic and Plausibility Models. The language and semantics of our logic, called Logic of Plausible Justifications (LPJ), is presented in Section 3, together with a sound and strongly complete axiomatic system for it. We also show that our logic has the finite model property and that the complexities of its Model-Checking and Satisfiability problems are the same as in the case of basic modal logics. In this section, we also present an extension of LPJ with a form of quantification over evidence terms, called LPJ^{Q}, similar to what [18] did in the context of traditional Justification Logics. Finally, in Section 4, we state our final remarks and point out potential further developments, including the one which originally motivated this work: the construction of a dynamic logic that can model argumentation and debate in multi-agent systems.

2. Background concepts

This section presents two important concepts for the construction of our logic: Justification Logic and Plausibility Models.

2.1. Justification Logic

In this section, we provide a brief account of Justification Logic. For more details, [9–11,19] can be consulted.