Artificial Intelligence 252 (2017) 1–21

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Resolving distributed knowledge ☆

Thomas Ågotnes [a,b], Yì N. Wáng [b,*]

[a] *University of Bergen, Norway*
[b] *Zhejiang University, China*

ARTICLE INFO

Article history:
Received 30 November 2015
Received in revised form 18 July 2017
Accepted 22 July 2017
Available online xxxx

Keywords:
Epistemic logic
Distributed knowledge
Expressivity
Axiomatization

ABSTRACT

In epistemic logic, a key formal theory for reasoning about knowledge in AI and other fields, different notions of group knowledge describe different ways in which knowledge can be associated with a group of agents. *Distributed knowledge* can be seen as the sum of the knowledge in a group; it is sometimes referred to as the potential knowledge of a group, or the joint knowledge they could obtain if they had unlimited means of communication. In epistemic logic, a formula of the form $D_G\varphi$ is intended to express the fact that group G has distributed knowledge of φ, that the total information in the group can be used to infer φ. In this paper we show that this is not the same as φ necessarily being true after *the members of the group actually share all their information with each other* – perhaps contrary to intuitive ideas about what distributed knowledge is. We furthermore introduce a new operator R_G, such that $R_G\varphi$ means that φ is true after G have shared all their information with each other – after G's distributed knowledge has been *resolved*. The R_G operators are called *resolution operators*. We study logics with different combinations of resolution operators and operators for common and distributed knowledge. Of particular interest is the relationship between distributed and common knowledge. The main results are characterizations of expressive power, and sound and complete axiomatizations. We also study the relationship to public announcement logic.

1. Introduction

A key formal theory for reasoning about knowledge in AI and other fields is epistemic logic based on modal logic [15,22,34]. When epistemic logic is used to reason about multi-agent systems, different notions of *group knowledge* describe different ways in which knowledge can be associated with a group of agents. *Common knowledge* is stronger than individual knowledge: that something is common knowledge requires not only that everybody in the group knows it, but that everybody knows that everybody knows it, and so on. *Distributed knowledge*, on the other hand, is weaker than individual knowledge: distributed knowledge is knowledge that is distributed throughout the group even if no individual knows it. Common and distributed knowledge have been characterized as "what every fool knows" and "what a wise man knows", respectively [15]. Given the important role of common knowledge for coordination and in social interactions in general, it is not strange that this type of group knowledge has received considerable attention from researchers. Distributed knowledge has received somewhat less attention, but should also be of great interest in AI and computer science as it is also closely

☆ A preliminary version of this paper can be found in the proceedings of TARK 2015.

* Corresponding author.
E-mail addresses: thomas.agotnes@uib.no (T. Ågotnes), ynw@xixilogic.org (Y.N. Wáng).

http://dx.doi.org/10.1016/j.artint.2017.07.002

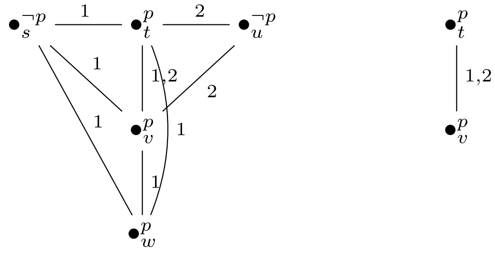

Fig. 1. Example taken from [30, p. 248]. Model M on the left, its communication core (for the set of all agents $\{1, 2\}$) on the right. Reflexivity is implicitly assumed.

related to several key problems, such as communication and aggregation of information. In this paper we make an attempt to rectify this, to better understand the theory of distributed knowledge.

Multi-agent epistemic logic contains unary operators K_i and D_G for each agent i and group of agents G, where a formula of the form $K_i\varphi$ is intended to mean that agent i knows φ and $D_G\varphi$ that φ is distributed knowledge in G. As an illustration, if we have that $K_1(p \to q)$ (agent 1 knows that p implies q) and that $K_2 p$ (agent 2 knows that p), then it would follow that $D_{\{1,2\}}q$ (it is distributed knowledge in group $\{1, 2\}$ that q). The formal semantics of this language is based on Kripke (possible worlds) models where each agent considers a number of states possible, among which the actual state of the world is always included. The formula $K_i\varphi$ is true if φ is true in all the states agent i considers possible. Distributed knowledge $D_G\varphi$ is then defined as holding whenever φ is true in each state that *every* agent in G considers possible. The intuition behind this definition, henceforth called the *standard semantics* of distributed knowledge, is that if at least one agent in the group knows that a state is *not* the actual state of the world, the group as a whole knows it. For example, in state t in the model M in Fig. 1, neither $K_1 p$ nor $K_2 p$ is true, but $D_{\{1,2\}} p$ is true. Intuitively, in this example, if the agents have unlimited means of communication then in state t agent 2 can tell agent 1 that s is not the actual state of the world and agent 1 can tell agent 2 that u us not the actual state of the world, so together they know that the actual state must be t. This information is distributed among them.

However, we should be careful about our intuitions here, because it turns out that there are many misconceptions about what the standard semantics of distributed knowledge actually entails, and many descriptions of what distributed knowledge actually is are inaccurate or even wrong. A particular misconception is that something is distributed knowledge in a group *if and only if* the agents in the group could get to know it after some (perhaps unlimited) communications between them,[1] as just illustrated in the example. To see that this interpretation must be incorrect, consider the formula $D_{\{1,2\}}(p \wedge \neg K_1 p)$. This formula says that it is distributed knowledge among agents 1 and 2 that p is true and that agent 1 does not know p. The formula is *consistent*: it is true in state t in model M in Fig. 1 ($p \wedge \neg K_1 p$ is true in states t and v). However, it is not possible that agents 1 and 2 both can get to know that p is true and that agent 1 does not know that p is true, no matter how much they communicate or "pool" their knowledge together. The "problem" here is that in a formula $D_G\psi$, ψ describes the possible states of the world as they were before any communication or other events took place, so a more accurate reading of $D_{\{1,2\}}(p \wedge \neg K_1 p)$ is that it follows from the combination of 1 and 2's knowledge that $p \wedge \neg K_1 p$ *were true before any communication or other events took place*. More technically, the "problem" is due to the compositional semantics of modal logic: in the evaluation of $D_G\varphi$, the D_G operator picks out a number of states considered possible by the group G (the states considered possible by *all* members of the group), and then φ is evaluated in each of these states *in the original model, without any effect of the D_G operator*.

However, we do not really consider this a problem. There are other interpretations of distributed knowledge where the consistency of the mentioned formula makes perfect sense, such that when distributed knowledge is the knowledge of a third party, someone "outside the system" who somehow has access to the epistemic states of all the group members. It shows, however, that it does not make sense to interpret distributed knowledge (with the standard semantics) as something that is true after the agents in the group have communicated *with each other*.

In this paper we introduce and study an alternative group modality R_G, where $R_G\varphi$ means (roughly speaking) that φ is true after the agents in the group have shared all their information with each other (and that it is common knowledge among the other agents in the system that they have). We call that *resolving* distributed knowledge, and the R_G operators are called *resolution operators*.

Semantically, we say that the expression $R_G\varphi$ is true iff φ is true in the model update obtained by removing links to states for members of G that are not linked by *all* members of G. This updated model is called G's *communication core* in [30, p. 249]. See Fig. 1 for an illustration. For example, the formula $R_{\{1,2\}}(p \wedge K_1 p)$ is true in state t in the model in Fig. 1,

[1] Some examples of informal descriptions of distributed knowledge from the literature include "A group has distributed knowledge of a fact φ if the knowledge of φ is distributed among its members, so that by pooling their knowledge together the members of the group can deduce φ" [15]; "... it should be possible for the members of the group to establish φ through communication" [33,28]; "... the knowledge that would result of the agents could somehow 'combine' their knowledge" [33]. These descriptions can at least give a reader the impression that distributed knowledge is about internal communication in the group of agents.