Modal and justification logics for multi-agent systems (extended abstract)

Christian Cachin[0000−−0001−−8967−−9213], David Lehnherr[0000−−0002−−4956−−4064], and Thomas Studer[0000−−0002−−0949−−3302]

Institute of Computer Science, University of Bern, Bern, Switzerland
{christian.cachin, david.lehnherr, thomas.studer}@unibe.ch

1 Introduction

Epistemic modal logic is an important tool in the area of distributed and multi-agent systems. An introduction and overview is given by the classic [10]. Recent work in this tradition includes, for instance, the study of the epistemic principles underlying blockchain mechanisms [5, 15, 21].

Another current line of research is concerned with combinatorial topological models of distributed systems and the development of corresponding modal logics [9, 17]. They give rise to new notions of group knowledge [7, 14] and new epistemic dynamic principles [13].

Modal logic cannot formalize the justifications underlying knowledge. This is remedied in justification logic. It replaces the $\Box$-modality from modal logic with explicit terms [3, 18]. That is, instead of formulas $\Box\phi$, meaning ϕ is known, justification logic includes formulas of the form $t : \phi$, meaning t represents a proof of ϕ. Originally, Artemov [1] introduced justification logic to give a classical provability semantics to intuitionistic logic. Since then, justification logic has been adapted to many epistemic and deontic use cases [2, 6, 11, 19, 22].

The first part of this talk deals with synergistic knowledge, a novel form of distributed knowledge. It is based on [7], which is joint work with Christian Cachin and David Lehnherr. The second part presents an epistemic model of zero-knowledge proofs in justification logic. It uses results from [20], which is joint work with David Lehnherr and Zoran Ognjanović.

2 Synergistic Knowledge

In modal logic, distributed knowledge of a group is usually defined as the knowledge that the group would have if all its members share their individual knowledge. This model, however, does not consider relations between agents. In this section, we discuss the notion of synergistic knowledge, which makes it possible to consider different relationships between the members of a group.

To do so, we use a novel semantics for modal logic that is based on simplicial complexes. With this semantics, a group of agents may know more than

just traditional distributed knowledge. Our logic features epistemic operators supporting a principle that could be paraphrased as *the sum is greater than its parts*, hence the name *synergistic knowledge*.

The following semantic definitions and the logic of synergistic knowledge have first been presented in [7].

Let Ag denote a set of finitely many agents and let

$$\mathsf{Agsi} = \{(A, i) \mid A \subseteq \mathsf{Ag} \text{ and } i \in \mathbb{N}\}.$$

The pair $(a, i) \in \mathsf{Agsi}$ represents agent a in local state i. Further, let $S \subseteq \mathsf{Agsi}$. An element $(A, i) \in S$ is *maximal in* S if and only if

$$\forall (B, j) \in S.|A| \geq |B|, \text{ where } |X| \text{ denotes the cardinality of the set } X.$$

Definition 1 (Simplex). *Let $\emptyset \neq S \subseteq \mathsf{Agsi}$. S is a* simplex *if and only if*

S1*: The maximal element is unique, i.e.*

$$\textit{if } (A, i) \in S \textit{ and } (B, j) \in S \textit{ are maximal in } S \textit{ then, } A = B \textit{ and } i = j.$$

The maximal element of S is denoted as $\max(S)$.

S2*: S is uniquely downwards closed, i.e. for all $(B, i) \in S$ and $\emptyset \neq C \subseteq B$*

$$\exists! j \in \mathbb{N}.(C, j) \in S, \textit{ where } !\exists j \textit{ means that there exists exactly one } j.$$

S3*: S contains nothing else, i.e.*

$$(B, i) \in S \textit{ and } (A, j) = \max(S) \textit{ implies } B \subseteq A.$$

Definition 2 (Complex). *Let $\mathbb{C}$ be a set of simplexes. $\mathbb{C}$ is a* complex *if and only if*

C*: For any $S, T \in \mathbb{C}$, if there exist A and i with $(A, i) \in S$ and $(A, i) \in T$, then*

$$\textit{for all } B \subseteq A \textit{ and all } j \quad (B, j) \in S \iff (B, j) \in T.$$

Definition 3 (Indistinguishability). *Let $S \subseteq \mathsf{Agsi}$, we define*

$$S^\circ = \{A \mid \exists i \in \mathbb{N} : (A, i) \in S\}.$$

An agent pattern *is a subset of $\mathsf{Pow}(\mathsf{Ag}) \setminus \{\emptyset\}$. For two simplicies S and T and an agent pattern G, we define $S \sim_G T$ if and only if $G \subseteq (S \cap T)^\circ$. In this case, we say that G* cannot distinguish *between S and T.*

Based on this indistinguishability relation we can define an epistemic logic. We start with a countable set of atomic propositions Prop. Formulas of the language $\mathcal{L}_{\mathsf{Syn}}$ are inductively defined by:

$$\phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid [G]\phi$$

where $p \in \mathsf{Prop}$ and G is an agent pattern. The remaining Boolean connectives are defined as usual. In particular, we set $\bot := p \wedge \neg p$ for some fixed $p \in \mathsf{Prop}$. Further, let G be an agent pattern. We define the formula $\mathsf{alive}(G)$ to be $\neg[G]\bot$. For a single agent a we write $\mathsf{alive}(a)$ instead of $\mathsf{alive}(\{a\})$.

Definition 4 (Model). *A model $\mathcal{M} = (\mathbb{C}, V)$ is a pair where*

1. *$\mathbb{C}$ is a complex and*
2. *$V : \mathbb{C} \to \mathsf{Pow}(\mathsf{Prop})$ is a valuation.*

Definition 5 (Truth). *Let $\mathcal{M} = (\mathbb{C}, V)$ be a model, $w \in \mathbb{C}$, and $\phi \in \mathcal{L}_{\mathsf{Syn}}$. We define $\mathcal{M}, w \Vdash \phi$ inductively by*

$$\begin{array}{lll} \mathcal{M}, w \Vdash p & \textit{iff} & p \in V(w) \\ \mathcal{M}, w \Vdash \neg\phi & \textit{iff} & \mathcal{M}, w \nVdash \phi \\ \mathcal{M}, w \Vdash \phi \wedge \psi & \textit{iff} & \mathcal{M}, w \Vdash \phi \textit{ and } \mathcal{M}, w \Vdash \psi \\ \mathcal{M}, w \Vdash [G]\phi & \textit{iff} & w \sim_G v \textit{ implies } \mathcal{M}, v \Vdash \phi \quad \textit{for all } v \in \mathbb{C}. \end{array}$$

We write $\mathcal{M} \Vdash \phi$ if $\mathcal{M}, w \Vdash \phi$ for all $w \in \mathbb{C}$. A formula ϕ is valid *if $\mathcal{M} \Vdash \phi$ for all models $\mathcal{M}$.*

The following formulas are valid:

$$[G](\phi \to \psi) \to ([G]\phi \to [G]\psi) \tag{K}$$
$$[G]\phi \to [G][G]\phi \tag{4}$$
$$\phi \to [G]\neg[G]\neg\phi \tag{B}$$
$$[G]\phi \to [H]\phi \quad \text{if } G \subseteq H \tag{Mono}$$
$$\mathsf{alive}(G) \wedge \mathsf{alive}(H) \to \mathsf{alive}(G \cup H) \tag{Union}$$
$$\mathsf{alive}(G) \to \mathsf{alive}(\{B\}) \quad \text{if there is } A \text{ with } A \in G \text{ and } B \subseteq A \tag{Sub}$$
$$\mathsf{alive}(G) \to \mathsf{alive}(\{A \cup B\}) \quad \text{if } A, B \in G \tag{Clo}$$
$$\mathsf{alive}(G) \to ([G]\phi \to \phi) \tag{T}$$

We finish this section with an example. Consider the complex given in Figure 1. It consists of two simplices, each being a tetrahedron. They share the vertices $a0$, $b0$, and $c0$. They also share the edges $(a0, b0)$, $(b0, c0)$, and $(c0, a0)$; but they do *not* share a face $(a0, b0, c0)$. Instead, there are two faces between these three vertices: one belonging to the upper tetrahedron, and one belonging to the lower tetrahedron.

Formally this complex is given by

$$\left\{ \left\{ \begin{array}{c} abcd0 \\ abc0, abd0, acd0, bcd0 \\ ab0, bc0, ac0, ad0, bd0, cd0 \\ a0, b0, c0, d0 \end{array} \right\}, \left\{ \begin{array}{c} abcd1 \\ abc1, abd1, acd1, bcd1 \\ ab0, bc0, ac0, ad1, bd1, cd1 \\ a0, b0, c0, d1 \end{array} \right\} \right\}.$$

We denote the two simplicies of this complex by $\langle abcd0 \rangle$ and $\langle abcd1 \rangle$. We find that

$$\langle abcd0 \rangle \sim_{\{\{a,b\},\{b,c\},\{a,c\}\}} \langle abcd1 \rangle \tag{1}$$

and

$$\langle abcd0 \rangle \not\sim_{\{\{a,b,c\}\}} \langle abcd1 \rangle. \tag{2}$$

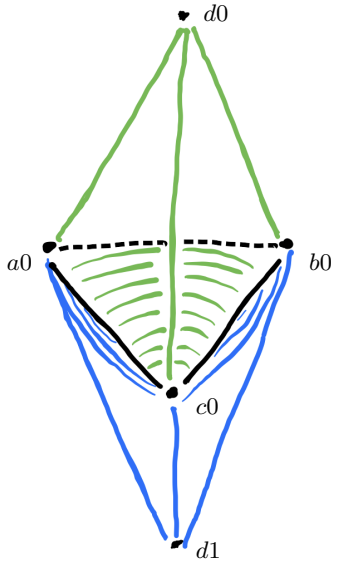

Fig. 1. Two tetrahedrons

(1) states that for the agents a, b, and c, having pairwise access to shared objects is not sufficient for knowing (as a group) whether d is in state 0 or in state 1. However, (2) models that if the agents a, b, and c have joint access to one shared object, then they can distinguish d being in state 0 from d being in state 1, i.e. they know in which state agent d is.

In distributed computing, such shared objects may be, for instance, shared-coin primitives or consensus objects. The notions of agent pattern and synergistic knowledge [7] can thus be used to analyze the concept of consensus number [16] or the problem of the dining cryptographers [8].

3 A logical model of zero-knowledge proofs

A recent application of justification logic [20] to multi-agent systems is to give an epistemic logic model of interactive proof systems and zero-knowledge proofs [4, 12]. These are protocols with the aim that an agent a has proof that an agent b knows ϕ without having the justification for b's knowledge. This may occur, for instance, if b wants to convince a that b knows a password without revealing the password.

An additional complication for a logical model of zero-knowledge proofs is that a cannot be fully convinced that b knows ϕ, but with a very high probability. Technically, this is done using the notion of negligible functions. For us, it suffices to add probability operators such that $P_{\approx r}\phi$ states that the probability of ϕ is almost r, i.e. infinitesimally close to r.

We will not present the logic and its semantics here. Instead, we only mention the key formulas in order to give an example of what can be expressed in this framework.

The formula

$$t :_b \phi \to P_{\approx 1}(f :_a \Box_b \phi)$$

states that if t is b's justification for knowing ϕ, then the protocol yields an f such that with almost certainty, f proves to a that b knows ϕ. Now this does not yet represent a zero-knowledge proof as b could simply transmit the justification t to a. A zero-knowledge proof additionally satisfies

$$t :_b \phi \to P_{\approx 0}(f :_a t :_b \phi).$$

If t is b's justification for knowing ϕ, then the probability that f proves to a that t is b's justification for knowing ϕ is negligible (where f is the result of the protocol).

This formalization is not only interesting from the perspective of computer science. It also that shows that for formal epistemology, it is important to have both the implicit $\Box_a$ operator and the explicit $t :_a$ modalities. The zero-knowledge protocol yields a proof f such that the probability of $f :_a \Box_b \phi$ is almost 1 whereas the probability of $f :_a t :_b \phi$ is almost 0. It also formalizes the fact that a can have higher-order knowledge of b knowing ϕ without knowing b's justification for that knowledge.

References

1. Artemov, S.: Explicit provability and constructive semantics. Bulletin of Symbolic Logic **7**(1), 1–36 (Mar 2001)
2. Artemov, S.: The logic of justification. The Review of Symbolic Logic **1**(4), 477–513 (Dec 2008). https://doi.org/10.1017/S1755020308090060
3. Artemov, S.N., Fitting, M.: Justification Logic: Reasoning with Reasons. Cambridge University Press (2019)
4. Babai, L.: Trading group theory for randomness. In: Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing. pp. 421—429. STOC '85, Association for Computing Machinery (1985). https://doi.org/10.1145/22145.22192
5. Brünnler, K., Flumini, D., Studer, T.: A logic of blockchain updates. Journal of Logic and Computation **30**(8), 1469–1485 (2020). https://doi.org/10.1093/logcom/exaa045
6. Bucheli, S., Kuznets, R., Studer, T.: Justifications for common knowledge. Journal of Applied Non-Classical Logics **21**(1), 35–60 (Jan–Mar 2011). https://doi.org/10.3166/JANCL.21.35-60
7. Cachin, C., Lehnherr, D., Studer, T.: Synergistic knowledge (submitted)
8. Chaum, D.: The dining cryptographers problem: Unconditional sender and recipient untraceability. J. Cryptol. **1**(1), 65–75 (1988). https://doi.org/10.1007/BF00206326, https://doi.org/10.1007/BF00206326
9. van Ditmarsch, H., Goubault, É., Ledent, J., Rajsbaum, S.: Knowledge and simplicial complexes. In: Lundgren, B., Nuñez Hernández, N.A. (eds.) Philosophy of Computing. pp. 1–50. Springer International Publishing, Cham (2022)

10. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: Reasoning About Knowledge. MIT Press (1995). https://doi.org/10.7551/mitpress/5803.001.0001, https://doi.org/10.7551/mitpress/5803.001.0001
11. Faroldi, F.L.G., Ghari, M., Lehmann, E., Studer, T.: Consistency and permission in deontic justification logic. Journal of Logic and Computation (2022). https://doi.org/10.1093/logcom/exac045
12. Goldwasser, S., Micali, S., Rackoff, C.: The knowledge complexity of interactive proof-systems. In: Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing. pp. 291—304. STOC '85, Association for Computing Machinery (1985). https://doi.org/10.1145/22145.22178
13. Goubault, É., Ledent, J., Rajsbaum, S.: A simplicial complex model for dynamic epistemic logic to study distributed task computability. Inf. Comput. **278**, 104597 (2021). https://doi.org/10.1016/j.ic.2020.104597, https://doi.org/10.1016/j.ic.2020.104597
14. Goubault, É.G., Kniazev, R., Ledent, J., Rajsbaum, S.: Semi-simplicial set models for distributed knowledge (2023)
15. Halpern, J.H., Pass, R.: A knowledge-based analysis of the blockchain protocol. In: Lang, K. (ed.) TARK 2017. pp. 324–335. No. 251 in EPTCS (2017). https://doi.org/10.4204/EPTCS.251.22
16. Herlihy, M.: Wait-free synchronization. ACM Trans. Program. Lang. Syst. **13**(1), 124–149 (1991). https://doi.org/10.1145/114005.102808, https://doi.org/10.1145/114005.102808
17. Herlihy, M., Kozlov, D.N., Rajsbaum, S.: Distributed Computing Through Combinatorial Topology. Morgan Kaufmann (2013), https://store.elsevier.com/product.jsp?isbn=9780124045781
18. Kuznets, R., Studer, T.: Logics of Proofs and Justifications. College Publications (2019)
19. Lehmann, E., Studer, T.: Subset models for justification logic. In: De Queiroz, R., Iemhoff, R., Moortgat, M. (eds.) WoLLIC 2019, Proceedings. Springer (2019)
20. Lehnherr, D., Ognjanović, Z., Studer, T.: A logic of interactive proofs. Journal of Logic and Computation **32**(8), 1645–1658 (2022). https://doi.org/10.1093/logcom/exac071
21. Marinković, B., Glavan, P., Ognjanović, Z., Studer, T.: A temporal epistemic logic with a non-rigid set of agents for analyzing the blockchain protocol. Journal of Logic and Computation **29**(5), 803–830 (2019). https://doi.org/10.1093/logcom/exz007
22. Xu, C., Wang, Y., Studer, T.: A logic of knowing why. Synthese **198**, 1259–1285 (2021)